CAMBRIDGE UNIVERSITY PRESS

ENDORSED BY
CAMBRIDGE International Examinations

David Sang, Graham Jones,
Gurinder Chadha and Richard Woodside

Cambridge International AS and A Level Physics

Coursebook

Second Edition



Completely Cambridge
Cambridge resources for Cambridge qualifications

David Sang, Graham Jones,
Gurinder Chadha and Richard Woodside

Cambridge International AS and A Level

Physics

Coursebook

Second Edition

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org

© Cambridge University Press 2010, 2014

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2010
Second edition 2014

Printed in the United Kingdom by Latimer Trend

A catalogue record for this publication is available from the British Library

ISBN 978-1-107-69769-0 Paperback with CD-ROM for Windows® and MAC®

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate. Information regarding prices, travel timetables, and other factual information given in this work is correct at the time of first printing but Cambridge University Press does not guarantee the accuracy of such information thereafter.

NOTICE TO TEACHERS IN THE UK
It is illegal to reproduce any part of this book in material form (including photocopying and electronic storage) except under the following circumstances:
(i) where you are abiding by a licence granted to your school or institution by the Copyright Licensing Agency;
(ii) where no such licence exists, or where you wish to exceed the terms of a licence, and you have gained the written permission of Cambridge University Press;
(iii) where you are allowed to reproduce without permission under the provisions of Chapter 3 of the Copyright, Designs and Patents Act 1988, which covers, for example, the reproduction of short passages within certain types of educational anthology and reproduction for the purposes of setting examination questions.

Example answers and all other end-of-chapter questions were written by the authors.

Contents

Introduction vii

How to use this book viii

Chapter 1: Kinematics – describing motion 1

- Speed 2
- Distance and displacement, scalar and vector 4
- Speed and velocity 5
- Displacement–time graphs 6
- Combining displacements 8
- Combining velocities 10

Chapter 2: Accelerated motion 14

- The meaning of acceleration 15
- Calculating acceleration 15
- Units of acceleration 16
- Deducing acceleration 17
- Deducing displacement 17
- Measuring velocity and acceleration 18
- Determining velocity and acceleration in the laboratory 18
- The equations of motion 20
- Deriving the equations of motion 22
- Uniform and non-uniform acceleration 24
- Acceleration caused by gravity 25
- Determining *g* 25
- Motion in two dimensions – projectiles 28
- Understanding projectiles 29

Chapter 3: Dynamics – explaining motion 37

- Calculating the acceleration 38
- Understanding SI units 39
- The pull of gravity 41
- Mass and inertia 43
- Top speed 44
- Moving through fluids 45
- Identifying forces 47
- Newton's third law of motion 49

Chapter 4: Forces – vectors and moments 53

- Combining forces 54
- Components of vectors 56
- Centre of gravity 59
- The turning effect of a force 59
- The torque of a couple 63

Chapter 5: Work, energy and power 69

- Doing work, transferring energy 71
- Gravitational potential energy 75
- Kinetic energy 76
- g.p.e.–k.e. transformations 76
- Down, up, down – energy changes 77
- Energy transfers 78
- Power 80

Chapter 6: Momentum 85

- The idea of momentum 86
- Modelling collisions 86
- Understanding collisions 89
- Explosions and crash-landings 91
- Collisions in two dimensions 93
- Momentum and Newton's laws 95
- Understanding motion 96

Chapter 7: Matter and materials 101

- Density 102
- Pressure 102
- Compressive and tensile forces 104
- Stretching materials 105
- Elastic potential energy 108

Chapter 8: Electric fields 116

- Attraction and repulsion 117
- The concept of an electric field 118
- Electric field strength 119
- Force on a charge 122

Chapter 9: Electric current, potential difference and resistance 127

- Circuit symbols and diagrams 128
- Electric current 129
- An equation for current 132
- The meaning of voltage 134
- Electrical resistance 135
- Electrical power 136

Chapter 10: Kirchhoff's laws 143

- Kirchhoff's first law 144
- Kirchhoff's second law 145
- Applying Kirchhoff's laws 146
- Resistor combinations 148

Chapter 11: Resistance and resistivity 156

The *I–V* characteristic for a metallic conductor 157
Ohm's law 158
Resistance and temperature 159
Resistivity 162

Chapter 12: Practical circuits 168

Internal resistance 169
Potential dividers 172
Potentiometer circuits 172

Chapter 13: Waves 178

Describing waves 179
Longitudinal and transverse waves 181
Wave energy 182
Wave speed 183
The Doppler effect 184
Electromagnetic waves 185
Electromagnetic radiation 186
Orders of magnitude 187
The nature of electromagnetic waves 188

Chapter 14: Superposition of waves 192

The principle of superposition of waves 193
Diffraction of waves 194
Interference 196
The Young double-slit experiment 200
Diffraction gratings 203

Chapter 15: Stationary waves 210

From moving to stationary 211
Nodes and antinodes 212
Formation of stationary waves 212
Determining the wavelength and speed of sound 216

Chapter 16: Radioactivity 222

Looking inside the atom 223
Alpha-particle scattering and the nucleus 223
A simple model of the atom 225
Nucleons and electrons 226
Forces in the nucleus 229
Fundamental particles? 229
Families of particles 230
Discovering radioactivity 231
Radiation from radioactive substances 231
Discovering neutrinos 232
Fundamental families 232
Fundamental forces 232
Properties of ionising radiation 233

P1: Practical skills at AS level 239

Practical work in physics 240
Using apparatus and following instructions 240
Gathering evidence 241
Precision, accuracy, errors and uncertainties 241
Finding the value of an uncertainty 243
Percentage uncertainty 245
Recording results 246
Analysing results 246
Testing a relationship 248
Identifying limitations in procedures and suggesting improvements 250

Chapter 17: Circular motion 258

Describing circular motion 259
Angles in radians 260
Steady speed, changing velocity 261
Angular velocity 261
Centripetal forces 262
Calculating acceleration and force 264
The origins of centripetal forces 265

Chapter 18: Gravitational fields 271

Representing a gravitational field 272
Gravitational field strength *g* 274
Energy in a gravitational field 276
Gravitational potential 276
Orbiting under gravity 277
The orbital period 278
Orbiting the Earth 279

Chapter 19: Oscillations 285

Free and forced oscillations 286
Observing oscillations 287
Describing oscillations 288
Simple harmonic motion 289
Representing s.h.m. graphically 291
Frequency and angular frequency 292
Equations of s.h.m. 293
Energy changes in s.h.m. 296
Damped oscillations 297
Resonance 299

Chapter 20: Communications systems 309

Radio waves 310
Analogue and digital signals 314
Channels of communication 317
Comparison of different channels 319

Chapter 21: Thermal physics 327
Changes of state 328
Energy changes 329
Internal energy 331
The meaning of temperature 332
Thermometers 334
Calculating energy changes 336

Chapter 22: Ideal gases 345
Particles of a gas 346
Explaining pressure 348
Measuring gases 348
Boyle's law 349
Changing temperature 350
Ideal gas equation 351
Modelling gases – the kinetic model 352
Temperature and molecular kinetic energy 354

Chapter 23: Coulomb's law 359
Electric fields 360
Coulomb's law 360
Electric field strength for a radial field 362
Electric potential 363
Comparing gravitational and electric fields 366

Chapter 24: Capacitance 372
Capacitors in use 373
Energy stored in a capacitor 375
Capacitors in parallel 377
Capacitors in series 378
Comparing capacitors and resistors 379
Capacitor networks 380

Chapter 25: Electronics 386
Components of an electronic sensing system 387
The operational amplifier (op-amp) 393
The inverting amplifier 397
The non-inverting amplifier 398
Output devices 398

Chapter 26: Magnetic fields and electromagnetism 406
Producing and representing magnetic fields 407
Magnetic force 409
Magnetic flux density 411
Measuring magnetic flux density 411
Currents crossing fields 413
Forces between currents 415
Relating SI units 416
Comparing forces in magnetic, electric and gravitational fields 417

Chapter 27: Charged particles 422
Observing the force 423
Orbiting charges 423
Electric and magnetic fields 427
The Hall effect 428
Discovering the electron 429

Chapter 28: Electromagnetic induction 435
Observing induction 436
Explaining electromagnetic induction 437
Faraday's law of electromagnetic induction 441
Lenz's law 443
Using induction: eddy currents, generators and transformers 445

Chapter 29: Alternating currents 451
Sinusoidal current 452
Alternating voltages 453
Power and a.c. 455
Why use a.c. for electricity supply? 457
Transformers 458
Rectification 460

Chapter 30: Quantum physics 466
Modelling with particles and waves 467
Particulate nature of light 468
The photoelectric effect 471
Line spectra 475
Explaining the origin of line spectra 476
Photon energies 477
Electron energies in solids 478
The nature of light – waves or particles? 480
Electron waves 480

Chapter 31: Nuclear physics 489
Balanced equations 490
Mass and energy 491
Energy released in radioactive decay 494
Binding energy and stability 494
Randomness and decay 496
The mathematics of radioactive decay 497
Decay graphs and equations 499
Decay constant and half-life 501

Chapter 32: Medical imaging 506

The nature and production of X-rays 507
X-ray attenuation 509
Improving X-ray images 511
Computerised axial tomography 513
Using ultrasound in medicine 516
Echo sounding 518
Ultrasound scanning 520
Magnetic resonance imaging 522

P2: Planning, analysis and evaluation 529

Planning 530
Analysis of the data 532
Treatment of uncertainties 536
Conclusions and evaluation of results 538

Appendix 1: Physical quantities and units 542

Prefixes 542
Estimation 542

Appendix 2: Data, formulae and relationships 543

Data 543
Conversion factors 543
Mathematical equations 544
Formulae and relationships 544

Appendix 3: The Periodic Table 545

Glossary 546

Index 555

Acknowledgements 564

CD-ROM CD1

Advice on how to revise for and approach examinations CD1
Introduction to the examination and changes to the syllabus CD9
Answers to self-assessment questions CD12
Answers to end-of-chapter questions CD70
Recommended resources CD136

Introduction

This book covers the entire syllabus of Cambridge International Examinations AS and A Level Physics. It is designed to work with the syllabus that will be examined from 2016. It is in three parts:

- Chapters 1–16 and P1: the AS level content, covered in the first year of the course, including a chapter (P1) dedicated to the development of your practical skills
- Chapters 17–32 and P2: the remaining A level content, including a chapter (P2) dedicated to developing your ability to plan, analyse and evaluate practical investigations
- Appendices of useful formulae, a Glossary and an Index.

The main tasks of a textbook like this are to explain the various concepts of physics that you need to understand and to provide you with questions that will help you to test your understanding and prepare for your examinations. You will find a visual guide to the structure of each chapter and the features of this book on the next two pages.

When tackling questions, it is a good idea to make a first attempt without referring to the explanations in this Coursebook or to your notes. This will help to reveal any gaps in your understanding. By working out which concepts you find most challenging, and by spending more time to understand these concepts at an early stage, you will progress faster as the course continues.

The CD-ROM that accompanies this Coursebook includes answers with workings for all the questions in the book, as well as suggestions for revising and preparing for any examinations you take. There are also lists of recommended further reading, which in many cases will take you beyond the requirements of the syllabus, but which will help you deepen your knowledge and explain more of the background to the physics concepts covered in this Coursebook.

In your studies, you will find that certain key concepts come up again and again, and that these concepts form 'themes' that link the different areas of physics together. It will help you to progress and gain confidence in tackling problems if you take note of these themes. For this Coursebook, these key concepts include:

- Models of physical systems
- Testing predictions against evidence
- Mathematics as a language and problem-solving tool
- Matter, energy and waves
- Forces and fields

In this Coursebook, the mathematics has been kept to the minimum required by the Cambridge International Examinations AS and A Level Physics syllabus. If you are also studying mathematics, you may find that more advanced techniques such as calculus will help you with many aspects of physics.

Studying physics can be a stimulating and worthwhile experience. It is an international subject; no single country has a monopoly on the development of the ideas. It can be a rewarding exercise to discover how men and women from many countries have contributed to our knowledge and well-being, through their research into and application of the concepts of physics. We hope not only that this book will help you to succeed in your future studies and career, but also that it will stimulate your curiosity and fire your imagination. Today's students become the next generation of physicists and engineers, and we hope that you will learn from the past to take physics to ever-greater heights.

How to use this book

Each chapter begins with a short list of the facts and concepts that are explained in it.

Chapter 1: Kinematics – describing motion

Learning outcomes

You should be able to:

- define displacement, speed and velocity
- draw and interpret displacement–time graphs
- describe laboratory methods for determining speed
- use vector addition to add two or more vectors

There is a short context at the beginning of each chapter, containing an example of how the material covered in the chapter relates to the 'real world'.

Describing movement

Our eyes are good at detecting movement. We notice even quite small movements out of the corners of our eyes. It's important for us to be able to judge movement – think about crossing the road, cycling or driving, or catching a ball.

Figure 1.1 shows a way in which movement can be recorded on a photograph. This is a stroboscopic photograph of a boy juggling three balls. As he juggles, a bright lamp flashes several times a second so that the camera records the positions of the balls at equal intervals of time.

If we knew the time between flashes, we could measure the photograph and calculate the speed of a ball as it moves through the air.

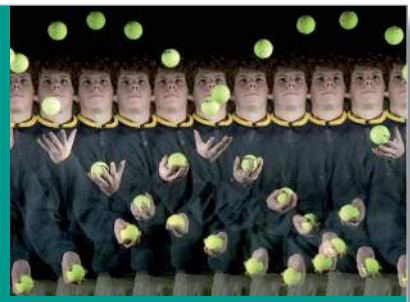


Figure 1.1 This boy is juggling three balls. A stroboscopic lamp flashes at regular intervals; the camera is moved to one side at a steady rate to show separate images of the boy.

The text and illustrations describe and explain all of the facts and concepts that you need to know. The chapters, and often the content within them as well, are arranged in a similar sequence to your syllabus, but with AS and A Level content clearly separated into the two halves of the book.

Figure 13.3 or a similar graph of displacement against time illustrates the following important definitions about waves and wave motion:

- The distance of a point on the wave from its undisturbed position or equilibrium position is called the **displacement** *x*.
- The maximum displacement of any point on the wave from its undisturbed position is called the **amplitude** *A*. The amplitude of a wave on the sea is measured in units of distance, e.g. metres. The greater the amplitude of the wave, the louder the sound or the rougher the sea!
- The distance from any point on a wave to the next exactly similar point (e.g. crest to crest) is called the **wavelength** λ (the Greek letter lambda). The wavelength of a wave on the sea is measured in units of distance, e.g. metres.
- The time taken for one complete oscillation of a point in a wave is called the **period** *T*. It is the time taken for a point to move from one particular position and return to that same position, moving in the same direction. It is measured in units of time, e.g. seconds..
- The number of oscillations per unit time of a point in a wave is called its **frequency** *f*. For sound waves, the higher the frequency of a musical note, the higher is its pitch. Frequency is measured in hertz (Hz), where 1 Hz = one oscillation per second (1 kHz = 10^3 Hz and 1 MHz = 10^6 Hz). The frequency *f* of a wave is the reciprocal of the period *T*:

$$f = \frac{1}{T}$$

Waves are called **mechanical waves** if they need a substance (medium) through which to travel. Sound is one example of such a wave. Other cases are waves on strings, seismic waves and water waves (Figure 13.4).

Some properties of typical waves are given on page 183 in Table 13.1.



Figure 13.4 The impact of a droplet on the surface of a liquid creates a vibration, which in turn gives rise to waves on the surface.

QUESTION

1 Determine the wavelength and amplitude of each of the two waves shown in Figure 13.5.

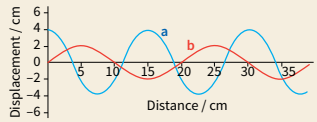


Figure 13.5 Two waves – for Question 1.

Questions throughout the text give you a chance to check that you have understood the topic you have just read about. You can find the answers to these questions on the CD-ROM.

BOX 13.1: Measuring frequency

You can measure the frequency of sound waves using a cathode-ray oscilloscope (c.r.o.). Figure 13.6 shows how.

A microphone is connected to the input of the c.r.o. Sound waves are captured by the microphone and converted into a varying voltage which has the same frequency as the sound waves. This voltage is displayed on the c.r.o. screen.

It is best to think of a c.r.o. as a voltmeter which is capable of displaying a rapidly varying voltage. To do this, its spot moves across the screen at a steady speed, set by the time-base control. At the same time, the spot moves up and down according to the voltage of the input.

Hence the display on the screen is a graph of the varying voltage, with time on the (horizontal) *x*-axis. If we know the horizontal scale, we can determine the period and hence the frequency of the sound wave. Worked example **1** shows how to do this. (In Chapter 15 we will look at one method of measuring the wavelength of sound waves.)

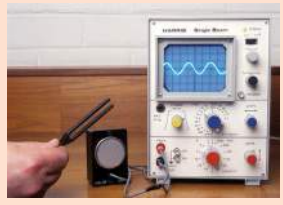


Figure 13.6 Measuring the frequency of sound waves from a tuning fork.

This book does not contain detailed instructions for doing particular experiments, but you will find background information about the practical work you need to do in these Boxes. There are also two chapters, P1 and P2, which provide detailed information about the practical skills you need to develop during your course.

Important equations and other facts are shown in highlight boxes.

For an object of mass *m* travelling at a speed *v*, we have:

kinetic energy = $\frac{1}{2}$ × mass × speed^2

$$E_k = \tfrac{1}{2}mv^2$$

Wherever you need to know how to use a formula to carry out a calculation, there are worked example boxes to show you how to do this.

WORKED EXAMPLE

1 In Figure 6.5, trolley A of mass 0.80 kg travelling at a velocity of 3.0 m s^{-1} collides head-on with a stationary trolley B. Trolley B has twice the mass of trolley A. The trolleys stick together and have a common velocity of 1.0 m s^{-1} after the collision. Show that momentum is conserved in this collision.

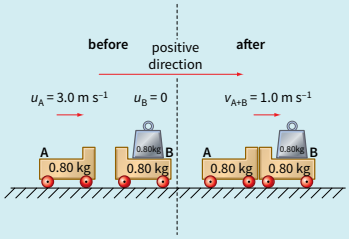


Figure 6.5 The state of trolleys A and B, before and after the collision.

Step 1 Make a sketch using the information given in the question. Notice that we need two diagrams to show the situations, one before and one after the collision. Similarly, we need two calculations – one for the momentum of the trolleys before the collision and one for their momentum after the collision.

Step 2 Calculate the momentum before the collision:

momentum of trolleys before collision

$= m_A \times u_A + m_B \times u_B$

$= (0.80 \times 3.0) + 0$

$= 2.4$ kg m s^{-1}

Trolley B has no momentum before the collision, because it is not moving.

Step 3 Calculate the momentum after the collision:

momentum of trolleys after collision

$= (m_A + m_B) \times v_{A+B}$

$= (0.80 + 1.60) \times 1.0$

$= 2.4$ kg m s^{-1}

So, both before and after the collision, the trolleys have a combined momentum of 2.4 kg m s^{-1}. Momentum has been conserved.

Key words are highlighted in the text when they are first introduced.

The metre, kilogram and second are three of the seven SI **base units**. These are defined with great precision so that every standards laboratory can reproduce them correctly.

You will also find definitions of these words in the Glossary.

base units Defined units of the SI system from which all other units are derived.

There is a summary of key points at the end of each chapter. You might find this helpful when you are revising.

Summary

- Forces are vector quantities that can be added by means of a vector triangle. Their resultant can be determined using trigonometry or by scale drawing.
- Vectors such as forces can be resolved into components. Components at right angles to one another can be treated independently of one another. For a force F at an angle θ to the x-direction, the components are:

  x-direction: $F\cos\theta$

  y-direction: $F\sin\theta$
- The moment of a force = force × perpendicular distance of the pivot from the line of action of the force.
- The principle of moments states that, for any object that is in equilibrium, the sum of the clockwise moments about any point provided by the forces acting on the object equals the sum of the anticlockwise moments about that same point.
- A couple is a pair of equal, parallel but opposite forces whose effect is to produce a turning effect on a body without giving it linear acceleration.

  torque of a couple = one of the forces × perpendicular distance between the forces
- For an object to be in equilibrium, the resultant force acting on the object must be zero and the resultant moment must be zero.

Questions at the end of each chapter begin with shorter answer questions, then move on to more demanding exam-style questions, some of which may require use of knowledge from previous chapters. Answers to these questions can be found on the CD–ROM.

1 Figure 15.19 shows a stationary wave on a string.

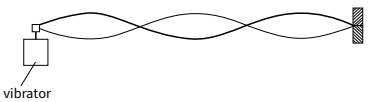


Figure 15.19 For End-of-chapter Question 1.

a On a copy of Figure 15.19, label one **node** (N) and one **antinode** (A). [1]

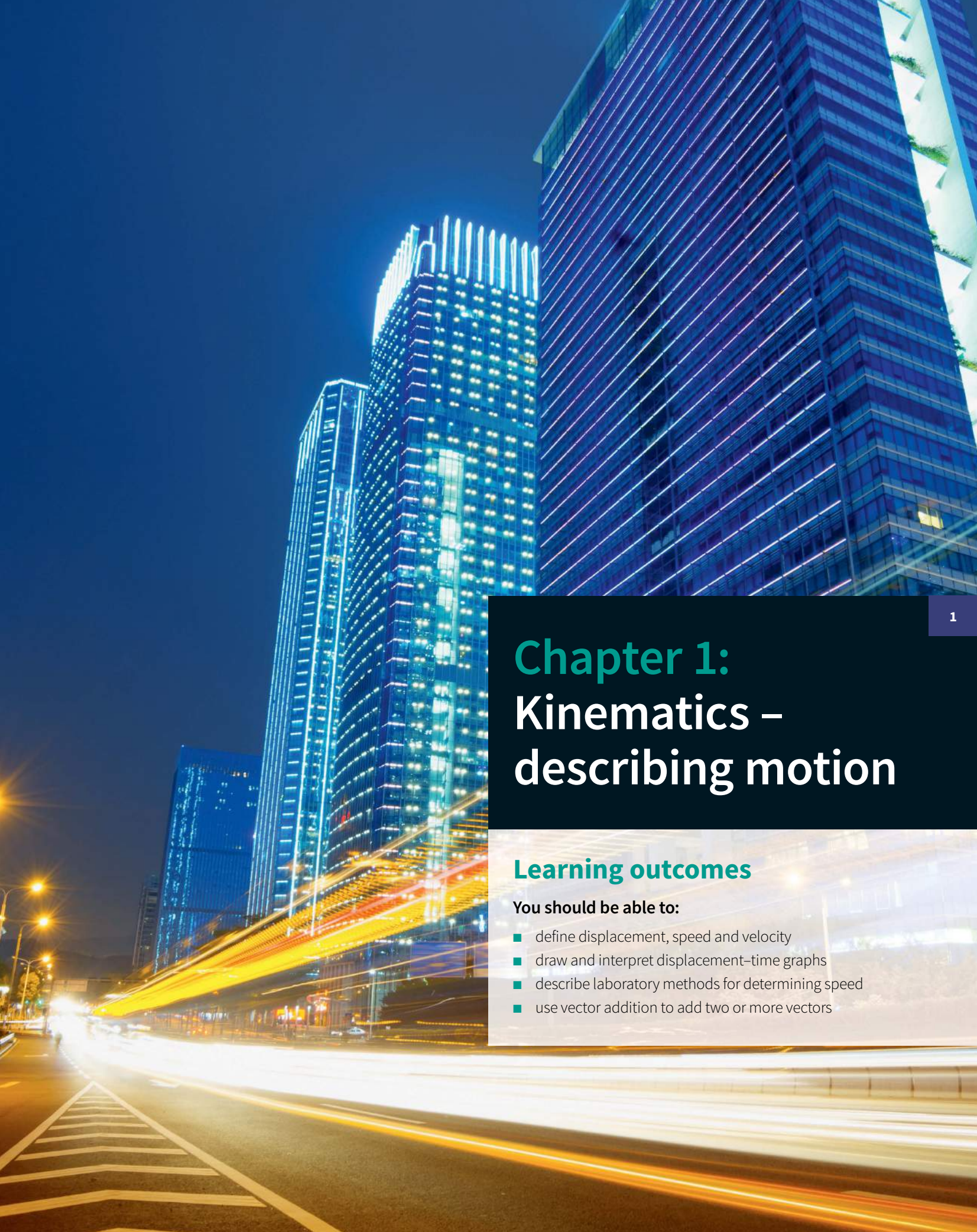
b Mark on your diagram the wavelength of the standing wave and label it λ. [1]

c The frequency of the vibrator is doubled. Describe the changes in the standing wave pattern. [1]

2 A tuning fork which produces a note of 256 Hz is placed above a tube which is nearly filled with water. The water level is lowered until resonance is first heard.

a Explain what is meant by the term **resonance**. [1]

b The length of the column of air above the water when resonance is first heard is 31.2 cm. Calculate the speed of the sound wave. [2]



Chapter 1: Kinematics – describing motion

Learning outcomes

You should be able to:

- define displacement, speed and velocity
- draw and interpret displacement–time graphs
- describe laboratory methods for determining speed
- use vector addition to add two or more vectors

Describing movement

Our eyes are good at detecting movement. We notice even quite small movements out of the corners of our eyes. It's important for us to be able to judge movement – think about crossing the road, cycling or driving, or catching a ball.

Figure 1.1 shows a way in which movement can be recorded on a photograph. This is a stroboscopic photograph of a boy juggling three balls. As he juggles, a bright lamp flashes several times a second so that the camera records the positions of the balls at equal intervals of time.

If we knew the time between flashes, we could measure the photograph and calculate the speed of a ball as it moves through the air.

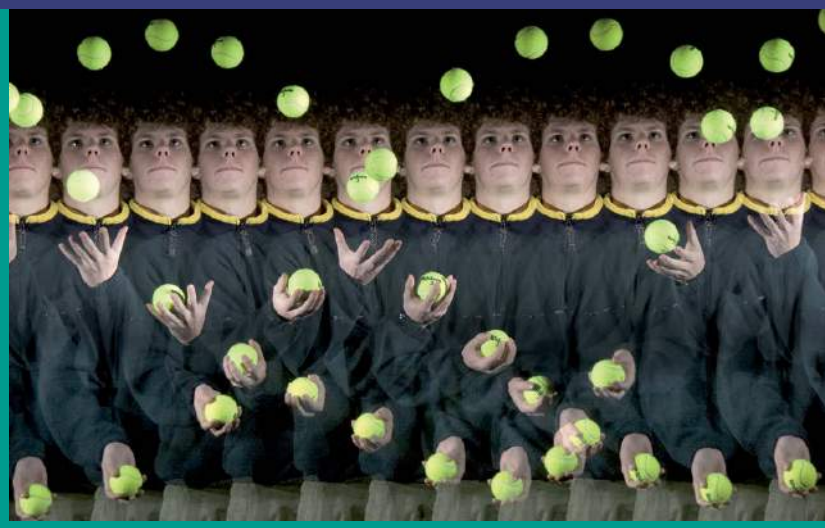


Figure 1.1 This boy is juggling three balls. A stroboscopic lamp flashes at regular intervals; the camera is moved to one side at a steady rate to show separate images of the boy.

Speed

We can calculate the average speed of something moving if we know the distance it moves and the time it takes:

$$\text{average speed} = \frac{\text{distance}}{\text{time}}$$

In symbols, this is written as:

$$v = \frac{d}{t}$$

where v is the average speed and d is the distance travelled in time t. The photograph (Figure 1.2) shows Ethiopia's Kenenisa Bekele posing next to the scoreboard after breaking the world record in a men's 10 000 metres race. The time on the clock in the photograph enables us to work out his average speed.

If the object is moving at a constant speed, this equation will give us its speed during the time taken. If its speed is changing, then the equation gives us its **average speed**. Average speed is calculated over a period of time.



Figure 1.2 Ethiopia's Kenenisa Bekele set a new world record for the 10 000 metres race in 2005.

If you look at the speedometer in a car, it doesn't tell you the car's average speed; rather, it tells you its speed at the instant when you look at it. This is the car's **instantaneous speed**.

QUESTION

1 Look at Figure 1.2. The runner ran 10 000 m, and the clock shows the total time taken. Calculate his average speed during the race.

Units

In the Système Internationale d'Unités (the SI system), distance is measured in metres (m) and time in seconds (s). Therefore, speed is in metres per second. This is written as m s^{-1} (or as m/s). Here, s^{-1} is the same as 1/s, or 'per second'.

There are many other units used for speed. The choice of unit depends on the situation. You would probably give the speed of a snail in different units from the speed of a racing car. Table 1.1 includes some alternative units of speed.

Note that in many calculations it is necessary to work in SI units (m s^{-1}).

m s^{-1}	metres per second
cm s^{-1}	centimetres per second
km s^{-1}	kilometres per second
km h^{-1} or km/h	kilometres per hour
mph	miles per hour

Table 1.1 Units of speed.

QUESTIONS

2 Here are some units of speed:

$m\,s^{-1}$ $mm\,s^{-1}$ $km\,s^{-1}$ $km\,h^{-1}$

Which of these units would be appropriate when stating the speed of each of the following?

a a tortoise

b a car on a long journey

c light

d a sprinter.

3 A snail crawls 12 cm in one minute. What is its average speed in $mm\,s^{-1}$?

Determining speed

You can find the speed of something moving by measuring the time it takes to travel between two fixed points. For example, some motorways have emergency telephones every 2000 m. Using a stopwatch you can time a car over this distance. Note that this can only tell you the car's average speed between the two points. You cannot tell whether it was increasing its speed, slowing down, or moving at a constant speed.

BOX 1.1: Laboratory measurements of speed

Here we describe four different ways to measure the speed of a trolley in the laboratory as it travels along a straight line. Each can be adapted to measure the speed of other moving objects, such as a glider on an air track, or a falling mass.

Measuring speed using two light gates

The leading edge of the card in Figure 1.3 breaks the light beam as it passes the first light gate. This starts the timer. The timer stops when the front of the card breaks the second beam. The trolley's speed is calculated from the time interval and the distance between the light gates.

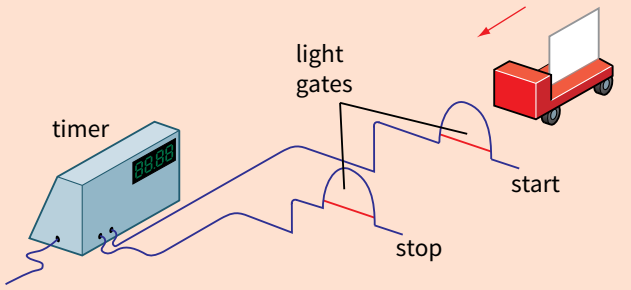


Figure 1.3 Using two light gates to find the average speed of a trolley.

Measuring speed using one light gate

The timer in Figure 1.4 starts when the leading edge of the card breaks the light beam. It stops when the trailing edge passes through. In this case, the time shown is the time taken for the trolley to travel a distance equal to the length of the card. The computer software can calculate the speed directly by dividing the distance by the time taken.

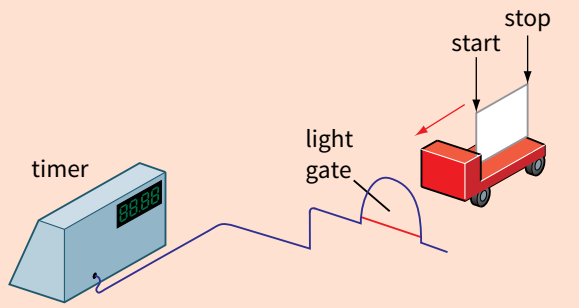


Figure 1.4 Using a single light gate to find the average speed of a trolley.

Measuring speed using a ticker-timer

The ticker-timer (Figure 1.5) marks dots on the tape at regular intervals, usually s (i.e. 0.02 s). (This is because it works with alternating current, and in most countries the frequency of the alternating mains is 50 Hz.)
The pattern of dots acts as a record of the trolley's movement.

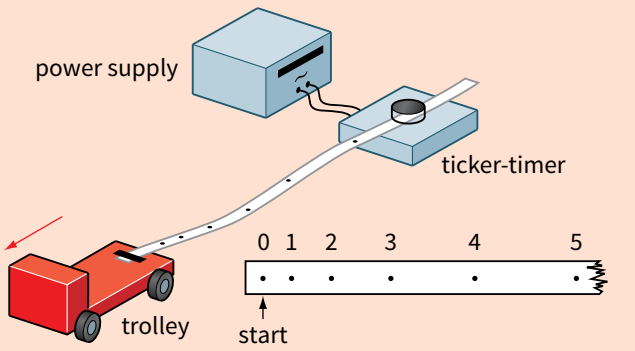


Figure 1.5 Using a ticker-timer to investigate the motion of a trolley.

BOX 1.1: Laboratory measurements of speed (continued)

Start by inspecting the tape. This will give you a description of the trolley's movement. Identify the start of the tape. Then look at the spacing of the dots:

- even spacing – constant speed
- increasing spacing – increasing speed.

Now you can make some measurements. Measure the distance of every fifth dot from the start of the tape. This will give you the trolley's distance at intervals of 0.1 s. Put the measurements in a table and draw a distance–time graph.

Measuring speed using a motion sensor

The motion sensor (Figure 1.6) transmits regular pulses of ultrasound at the trolley. It detects the reflected waves and determines the time they took for the trip to the trolley and back. From this, the computer can deduce the distance to the trolley from the motion sensor. It can generate a distance–time graph. You can determine the speed of the trolley from this graph.

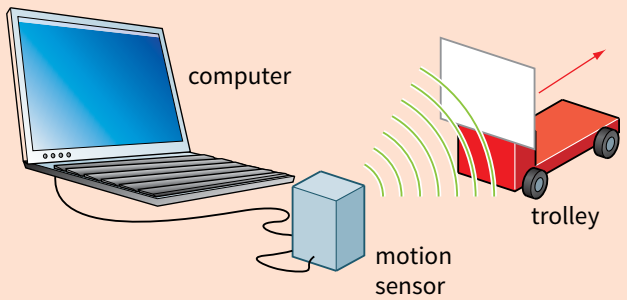


Figure 1.6 Using a motion sensor to investigate the motion of a trolley.

Choosing the best method

Each of these methods for finding the speed of a trolley has its merits. In choosing a method, you might think about the following points:

- Does the method give an average value of speed or can it be used to give the speed of the trolley at different points along its journey?
- How precisely does the method measure time – to the nearest millisecond?
- How simple and convenient is the method to set up in the laboratory?

QUESTIONS

4 A trolley with a 5.0 cm long card passed through a single light gate. The time recorded by a digital timer was 0.40 s. What was the average speed of the trolley in $m\,s^{-1}$?

5 Figure 1.7 shows two ticker-tapes. Describe the motion of the trolleys which produced them.

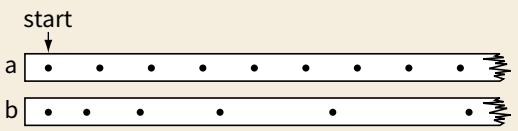


Figure 1.7 Two ticker-tapes; for Question 5.

6 Four methods for determining the speed of a moving trolley have been described. Each could be adapted to investigate the motion of a falling mass. Choose two methods which you think would be suitable, and write a paragraph for each to say how you would adapt it for this purpose.

Distance and displacement, scalar and vector

In physics, we are often concerned with the distance moved by an object in a particular direction. This is called its **displacement**. Figure 1.8 illustrates the difference between distance and displacement. It shows the route followed by walkers as they went from town A to town C. Their winding route took them through town B, so that they covered a total distance of 15 km. However, their displacement was much less than this. Their finishing position was just 10 km from where they started. To give a complete statement of their displacement, we need to give both distance and direction:

displacement = 10 km 30° E of N

Displacement is an example of a **vector quantity**. A vector quantity has both magnitude (size) and direction. Distance, on the other hand, is a **scalar quantity**. Scalar quantities have magnitude only.

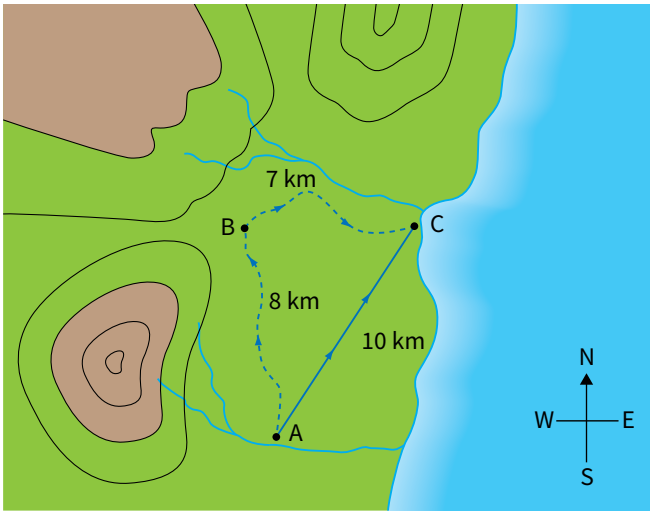


Figure 1.8 If you go on a long walk, the distance you travel will be greater than your displacement. In this example, the walkers travel a distance of 15 km, but their displacement is only 10 km, because this is the distance from the start to the finish of their walk.

Speed and velocity

It is often important to know both the speed of an object and the direction in which it is moving. Speed and direction are combined in another quantity, called **velocity**. The velocity of an object can be thought of as its speed in a particular direction. So, like displacement, velocity is a **vector** quantity. Speed is the corresponding scalar quantity, because it does not have a direction. So, to give the velocity of something, we have to state the direction in which it is moving. For example, an aircraft flies with a velocity of 300 m s^{-1} due north. Since velocity is a vector quantity, it is defined in terms of displacement:

$$\text{velocity} = \frac{\text{change in displacement}}{\text{time taken}}$$

Alternatively, we can say that velocity is the rate of change of an object's displacement. From now on, you need to be clear about the distinction between velocity and speed, and between displacement and distance. Table 1.2 shows the standard symbols and units for these quantities.

Quantity	Symbol for quantity	Symbol for unit
distance	d	m
displacement	s, x	m
time	t	s
speed, velocity	v	m s^{-1}

Table 1.2 Standard symbols and units. (Take care not to confuse italic *s* for displacement with s for seconds. Notice also that *v* is used for both speed and velocity.)

QUESTION

7 Which of these gives speed, velocity, distance or displacement? (Look back at the definitions of these quantities.)

- **a** The ship sailed south-west for 200 miles.
- **b** I averaged 7 mph during the marathon.
- **c** The snail crawled at 2 mm s^{-1} along the straight edge of a bench.
- **d** The sales representative's round trip was 420 km.

Speed and velocity calculations

We can write the equation for velocity in symbols:

$$v = \frac{s}{t}$$

$$v = \frac{\Delta s}{\Delta t}$$

The word equation for velocity is:

$$\text{velocity} = \frac{\text{change in displacement}}{\text{time taken}}$$

Note that we are using Δs to mean 'change in displacement *s*'. The symbol Δ, Greek letter delta, means 'change in'. It does not represent a quantity (in the way that *s* does); it is simply a convenient way of representing a change in a quantity. Another way to write Δs would be $s_2 - s_1$, but this is more time-consuming and less clear.

The equation for velocity, $v = \frac{\Delta s}{\Delta t}$, can be rearranged as follows, depending on which quantity we want to determine:

$$\text{change in displacement } \Delta s = v \times \Delta t$$

$$\text{change in time } \Delta t = \frac{\Delta s}{v}$$

Note that each of these equations is balanced in terms of units. For example, consider the equation for displacement. The units on the right-hand side are m s^{-1} × s, which simplifies to m, the correct unit for displacement.

Note also that we can, of course, use the same equations to find speed and distance, that is:

$$v = \frac{d}{t}$$

$$\text{distance } d = v \times t$$

$$\text{time } t = \frac{d}{v}$$

WORKED EXAMPLES

1 A car is travelling at 15 m s^{-1}. How far will it travel in 1 hour?

Step 1 It is helpful to start by writing down what you know and what you want to know:

$v = 15\,\text{m s}^{-1}$

$t = 1\,\text{h} = 3600\,\text{s}$

$d = ?$

Step 2 Choose the appropriate version of the equation and substitute in the values. Remember to include the units:

$d = v \times t$

$= 15 \times 3600$

$= 5.4 \times 10^4\,\text{m}$

$= 54\,\text{km}$

The car will travel 54 km in 1 hour.

2 The Earth orbits the Sun at a distance of 150 000 000 km. How long does it take light from the Sun to reach the Earth?
(Speed of light in space = $3.0 \times 10^8\,\text{m s}^{-1}$.)

Step 1 Start by writing what you know. Take care with units; it is best to work in m and s. You need to be able to express numbers in scientific notation (using powers of 10) and to work with these on your calculator.

$v = 3.0 \times 10^8\,\text{m s}^{-1}$

$d = 150\,000\,000\,\text{km}$

$= 150\,000\,000\,000\,\text{m}$

$= 1.5 \times 10^{11}\,\text{m}$

Step 2 Substitute the values in the equation for time:

$$t = \frac{d}{v} = \frac{1.5 \times 10^{11}}{3.0 \times 10^8} = 500\,\text{s}$$

Light takes 500 s (about 8.3 minutes) to travel from the Sun to the Earth.

Hint: *When using a calculator, to calculate the time t, you press the buttons in the following sequence:*

[1.5] [EXP] [11] [÷] [3] [EXP] [8]

or

[1.5] [×10^n] [11] [÷] [3] [×10^n] [8]

Making the most of units

In Worked example 1 and Worked example 2, units have been omitted in intermediate steps in the calculations. However, at times it can be helpful to include units as this can be a way of checking that you have used the correct equation; for example, that you have not divided one quantity by another when you should have multiplied them. The units of an equation must be balanced, just as the numerical values on each side of the equation must be equal.

If you take care with units, you should be able to carry out calculations in non-SI units, such as kilometres per hour, without having to convert to metres and seconds.

For example, how far does a spacecraft travelling at 40 000 km h^{-1} travel in one day? Since there are 24 hours in one day, we have:

distance travelled = $40\,000\,\text{km h}^{-1} \times 24\,\text{h}$

$= 960\,000\,\text{km}$

QUESTIONS

8 A submarine uses sonar to measure the depth of water below it. Reflected sound waves are detected 0.40 s after they are transmitted. How deep is the water? (Speed of sound in water = 1500 m s^{-1}.)

9 The Earth takes one year to orbit the Sun at a distance of 1.5×10^{11} m. Calculate its speed. Explain why this is its average speed and not its velocity.

Displacement–time graphs

We can represent the changing position of a moving object by drawing a displacement–time graph. The gradient (slope) of the graph is equal to its velocity (Figure 1.9). The steeper the slope, the greater the velocity. A graph like this can also tell us if an object is moving forwards or backwards. If the gradient is negative, the object's velocity is negative – it is moving backwards.

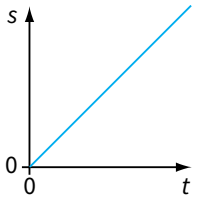
Deducing velocity from a displacement–time graph

A toy car moves along a straight track. Its displacement at different times is shown in Table 1.3. This data can be used to draw a displacement–time graph from which we can deduce the car's velocity.

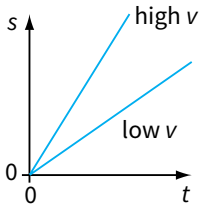
Displacement / m	1.0	3.0	5.0	7.0	7.0	7.0
Time / s	0.0	1.0	2.0	3.0	4.0	5.0

Table 1.3 Displacement (s) and time (t) data for a toy car.

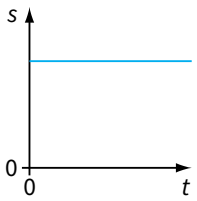
The straight line shows that the object's velocity is constant.



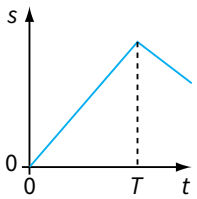
The slope shows which object is moving faster. The steeper the slope, the greater the velocity.



The slope of this graph is 0. The displacement s is not changing. Hence the velocity $v = 0$. The object is stationary.



The slope of this graph suddenly becomes negative. The object is moving back the way it came. Its velocity v is negative after time T.



This displacement–time graph is curved. The slope is changing. This means that the object's velocity is changing – this is considered in Chapter 2.

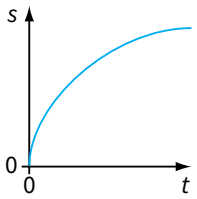


Figure 1.9 The slope of a displacement–time (s–t) graph tells us how fast an object is moving.

It is useful to look at the data first, to see the pattern of the car's movement. In this case, the displacement increases steadily at first, but after 3.0 s it becomes constant. In other words, initially the car is moving at a steady velocity, but then it stops.

Now we can plot the displacement–time graph (Figure 1.11).

We want to work out the velocity of the car over the first 3.0 seconds. We can do this by working out the gradient of the graph, because:

velocity = gradient of displacement–time graph

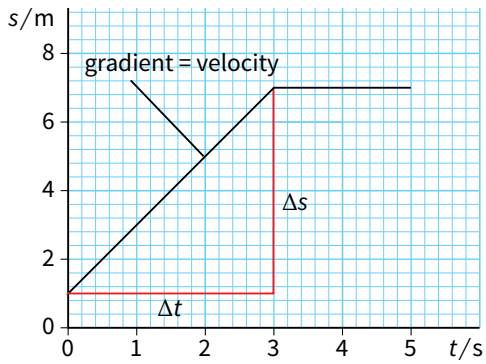


Figure 1.11 Displacement–time graph for a toy car; data as shown in Table 1.3.

We draw a right-angled triangle as shown. To find the car's velocity, we divide the change in displacement by the change in time. These are given by the two sides of the triangle labelled Δs and Δt.

$$\text{velocity } v = \frac{\text{change in displacement}}{\text{time taken}}$$

$$v = \frac{\Delta s}{\Delta t}$$

$$v = \frac{(7.0 - 1.0)}{(3.0 - 0)} = \frac{6.0}{3.0} = 2.0\,\text{ms}^{-1}$$

If you are used to finding the gradient of a graph, you may be able to reduce the number of steps in this calculation.

QUESTIONS

10 The displacement–time sketch graph in Figure 1.10 represents the journey of a bus. What does the graph tell you about the journey?

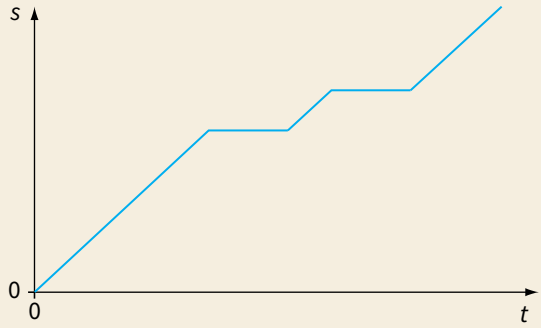


Figure 1.10 For Question 10.

11 Sketch a displacement–time graph to show your motion for the following event. You are walking at a constant speed across a field after jumping off a gate. Suddenly you see a bull and stop. Your friend says there's no danger, so you walk on at a reduced constant speed. The bull bellows, and you run back to the gate. Explain how each section of the walk relates to a section of your graph.

12 Table 1.4 shows the displacement of a racing car at different times as it travels along a straight track during a speed trial.

a Determine the car's velocity.

b Draw a displacement–time graph and use it to find the car's velocity.

Displacement / m	0	85	170	255	340
Time / s	0	1.0	2.0	3.0	4.0

Table 1.4 Displacement (s) and time (t) data for Question 12.

13 An old car travels due south. The distance it travels at hourly intervals is shown in Table 1.5.

a Draw a distance–time graph to represent the car's journey.

b From the graph, deduce the car's speed in km h^{-1} during the first three hours of the journey.

c What is the car's average speed in km h^{-1} during the whole journey?

Time / h	**Distance / km**
0	0
1	23
2	46
3	69
4	84

Table 1.5 Data for Question 13.

Combining displacements

The walkers shown in Figure 1.12 are crossing difficult ground. They navigate from one prominent point to the next, travelling in a series of straight lines. From the map, they can work out the distance that they travel and their displacement from their starting point:

distance travelled = 25 km

(Lay thread along route on map; measure thread against map scale.)

displacement = 15 km north-east

(Join starting and finishing points with straight line; measure line against scale.)

A map is a scale drawing. You can find your displacement by measuring the map. But how can you **calculate** your displacement? You need to use ideas from geometry and trigonometry. Worked examples 3 and 4 show how.

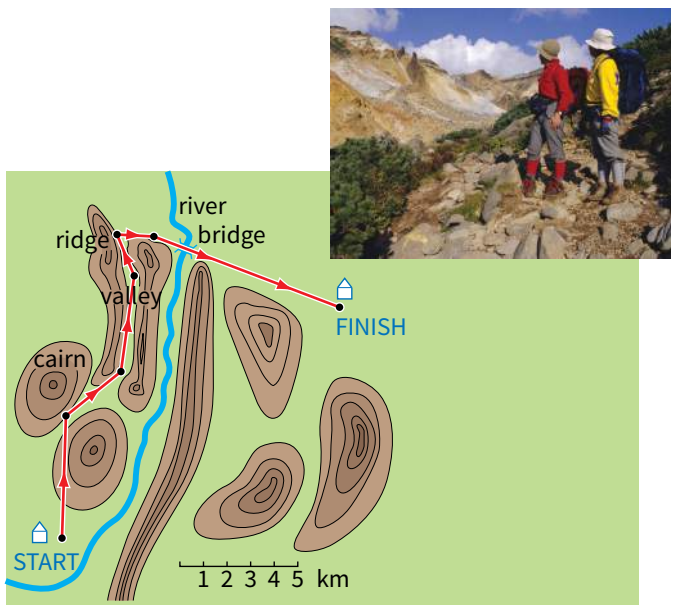


Figure 1.12 In rough terrain, walkers head straight for a prominent landmark.

WORKED EXAMPLES

3 A spider runs along two sides of a table (Figure 1.13). Calculate its final displacement.

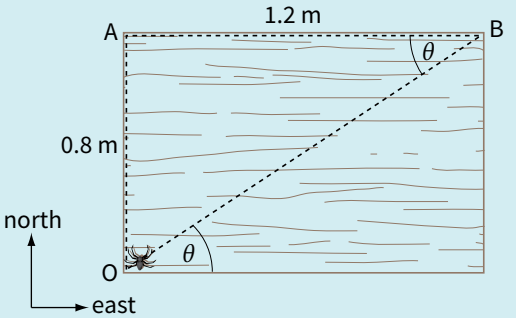


Figure 1.13 The spider runs a distance of 2.0 m, but what is its displacement?

Step 1 Because the two sections of the spider's run (OA and AB) are at right angles, we can **add** the two displacements using Pythagoras's theorem:

$OB^2 = OA^2 + AB^2$

$= 0.8^2 + 1.2^2 = 2.08$

$OB = \sqrt{2.08} = 1.44\,\text{m} \approx 1.4\,\text{m}$

Step 2 Displacement is a vector. We have found the **magnitude** of this vector, but now we have to find its direction. The angle θ is given by:

$$\tan\theta = \frac{\text{opp}}{\text{adj}} = \frac{0.8}{1.2}$$

$= 0.667$

$\theta = \tan^{-1}(0.667)$

$= 33.7° \approx 34°$

So the spider's displacement is 1.4 m at an angle of 34° north of east.

4 An aircraft flies 30 km due east and then 50 km north-east (Figure 1.14). Calculate the final displacement of the aircraft.

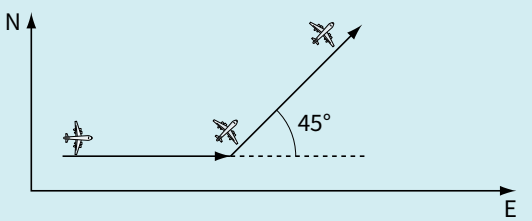


Figure 1.14 What is the aircraft's final displacement?

Here, the two displacements are not at 90° to one another, so we can't use Pythagoras's theorem. We can solve this problem by making a scale drawing, and measuring the final displacement. (However, you could solve the same problem using trigonometry.)

Step 1 Choose a suitable scale. Your diagram should be reasonably large; in this case, a scale of 1 cm to represent 5 km is reasonable.

Step 2 Draw a line to represent the first vector. North is at the top of the page. The line is 6 cm long, towards the east (right).

Step 3 Draw a line to represent the second vector, starting at the end of the first vector. The line is 10 cm long, and at an angle of 45° (Figure 1.15).

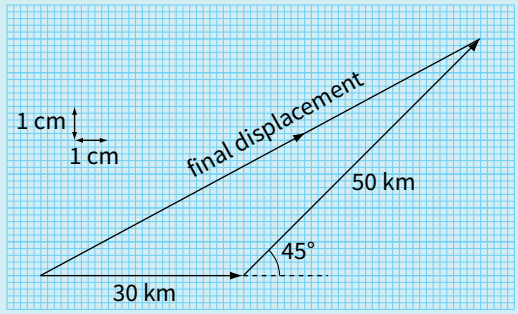


Figure 1.15 Scale drawing for Worked example 4. Using graph paper can help you to show the vectors in the correct directions.

Step 4 To find the final displacement, join the start to the finish. You have created a **vector triangle**. Measure this displacement vector, and use the scale to convert back to kilometres:

length of vector = 14.8 cm

final displacement = 14.8 × 5 = 74 km

Step 5 Measure the angle of the final displacement vector:

angle = 28° N of E

Therefore the aircraft's final displacement is 74 km at 28° north of east.

QUESTIONS

14 You walk 3.0 km due north, and then 4.0 km due east.

- **a** Calculate the total distance in km you have travelled.
- **b** Make a scale drawing of your walk, and use it to find your final displacement. Remember to give both the magnitude and the direction.
- **c** Check your answer to part **b** by calculating your displacement.

15 A student walks 8.0 km south-east and then 12 km due west.

- **a** Draw a vector diagram showing the route. Use your diagram to find the total displacement. Remember to give the scale on your diagram and to give the direction as well as the magnitude of your answer.
- **b** Calculate the resultant displacement. Show your working clearly.

This process of adding two displacements together (or two or more of any type of vector) is known as **vector addition**. When two or more vectors are added together, their combined effect is known as the **resultant** of the vectors.

Combining velocities

Velocity is a vector quantity and so two velocities can be combined by vector addition in the same way that we have seen for two or more displacements.

Imagine that you are attempting to swim across a river. You want to swim directly across to the opposite bank, but the current moves you sideways at the same time as you are swimming forwards. The outcome is that you will end up on the opposite bank, but downstream of your intended landing point. In effect, you have two velocities:

- the velocity due to your swimming, which is directed straight across the river
- the velocity due to the current, which is directed downstream, at right angles to your swimming velocity.

These combine to give a **resultant** (or net) velocity, which will be diagonally downstream. In order to swim directly across the river, you would have to aim upstream. Then your resultant velocity could be directly across the river.

WORKED EXAMPLE

5 An aircraft is flying due north with a velocity of 200 m s^{-1}. A side wind of velocity 50 m s^{-1} is blowing due east. What is the aircraft's resultant velocity (give the magnitude and direction)?

Here, the two velocities are at 90°. A sketch diagram and Pythagoras's theorem are enough to solve the problem.

Step 1 Draw a sketch of the situation – this is shown in Figure 1.16a.

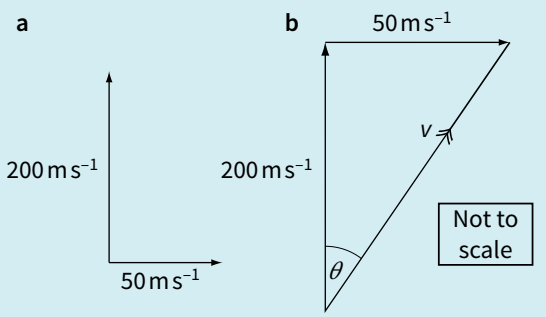


Figure 1.16 Finding the resultant of two velocities – for Worked example 5.

Step 2 Now sketch a vector triangle. Remember that the second vector starts where the first one ends. This is shown in Figure 1.16b.

Step 3 Join the start and end points to complete the triangle.

Step 4 Calculate the magnitude of the resultant vector v (the hypotenuse of the right-angled triangle).

$$v^2 = 200^2 + 50^2 = 40\,000 + 2500 = 42\,500$$

$$v = \sqrt{42\,500} \approx 206 \text{ m s}^{-1}$$

Step 5 Calculate the angle θ:

$$\tan\theta = \frac{50}{200}$$

$$= 0.25$$

$$\theta = \tan^{-1}(0.25) \approx 14°$$

So the aircraft's resultant velocity is 206 m s^{-1} at 14° east of north.

QUESTIONS

16 A swimmer can swim at 2.0 m s^{-1} in still water. She aims to swim directly across a river which is flowing at 0.80 m s^{-1}. Calculate her resultant velocity. (You must give both the magnitude and the direction.)

17 A stone is thrown from a cliff and strikes the surface of the sea with a vertical velocity of 18 m s^{-1} and a horizontal velocity *v*. The resultant of these two velocities is 25 m s^{-1}.

a Draw a vector diagram showing the two velocities and the resultant.

b Use your diagram to find the value of *v*.

c Use your diagram to find the angle between the stone and the vertical as it strikes the water.

Summary

- Displacement is the distance travelled in a particular direction.
- Velocity is defined by the word equation

$$\text{velocity} = \frac{\text{change in displacement}}{\text{time taken}}$$

The gradient of a displacement–time graph is equal to velocity:

$$\text{velocity} = \frac{\Delta s}{\Delta t}$$

- Distance and speed are scalar quantities. A scalar quantity has only magnitude.
- Displacement and velocity are vector quantities. A vector quantity has both magnitude and direction.
- Vector quantities may be combined by vector addition to find their resultant.

End-of-chapter questions

1 A car travels one complete lap around a circular track at a constant speed of 120 km h^{-1}.

a If one lap takes 2.0 minutes, show that the length of the track is 4.0 km. [2]

b Explain why values for the **average speed** and **average velocity** are different. [1]

c Determine the magnitude of the displacement of the car in a time of 1.0 minute. [2]

(The circumference of a circle = $2\pi R$, where R is the radius of the circle.)

2 A boat leaves point A and travels in a straight line to point B (Figure 1.17). The journey takes 60 s.

Calculate:

a the distance travelled by the boat [2]

b the total displacement of the boat [2]

c the average velocity of the boat. [2]

Remember that each vector quantity must be given a direction as well as a magnitude.

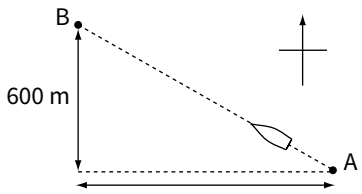


Figure 1.17 For End-of-chapter Question 2.

3 A boat travels at 2.0 m s^{-1} east towards a port, 2.2 km away. When the boat reaches the port, the passengers travel in a car due north for 15 minutes at 60 km h^{-1}.

Calculate:

a the total distance travelled [2]

b the total displacement [2]

c the total time taken [2]

d the average speed in m s^{-1} [2]

e the magnitude of the average velocity. [2]

4 A river flows from west to east with a constant velocity of 1.0 m s^{-1}. A boat leaves the south bank heading due north at 2.40 m s^{-1}. Find the resultant velocity of the boat. [2]

5 **a** Define **displacement**. [1]

b Use the definition of displacement to explain how it is possible for an athlete to run round a track yet have no displacement. [2]

6 A girl is riding a bicycle at a constant velocity of 3.0 m s^{-1} along a straight road. At time t = 0, she passes a boy sitting on a stationary bicycle. At time t = 0, the boy sets off to catch up with the girl. His velocity increases from time t = 0 until t = 5.0 s, when he has covered a distance of 10 m. He then continues at a constant velocity of 4.0 m s^{-1}.

a Draw the displacement–time graph for the girl from t = 0 to 12 s. [1]

b On the same graph axes, draw the displacement–time graph for the boy. [2]

c Using your graph, determine the value of t when the boy catches up with the girl. [1]

7 A student drops a small black sphere alongside a vertical scale marked in centimetres. A number of flash photographs of the sphere are taken at 0.1 s intervals, as shown in Figure 1.18. The first photograph is taken with the sphere at the top at time $t = 0$ s.

a Explain how Figure 1.18 shows that the sphere reaches a constant speed. [2]

b Determine the constant speed reached by the sphere. [2]

c Determine the distance that the sphere has fallen when $t = 0.8$ s. [2]

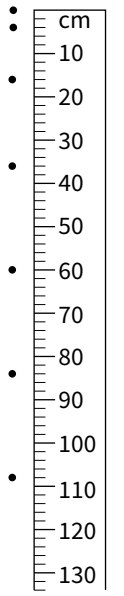


Figure 1.18 For End-of-chapter Question 7.

8 **a** State **one** difference between a scalar quantity and a vector quantity and give an example of each. [3]

b A plane has an air speed of 500 km h^{-1} due north. A wind blows at 100 km h^{-1} from east to west. Draw a vector diagram to calculate the resultant velocity of the plane. Give the direction of travel of the plane with respect to north. [4]

c The plane flies for 15 minutes. Calculate the displacement of the plane in this time. [1]

9 A small aircraft for one person is used on a short horizontal flight. On its journey from A to B, the resultant velocity of the aircraft is 15 m s^{-1} in a direction 60° east of north and the wind velocity is 7.5 m s^{-1} due north (Figure 1.19).

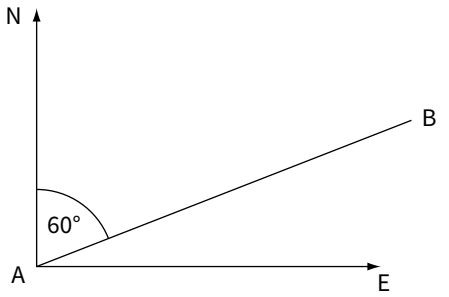


Figure 1.19 For End-of-chapter Question 9.

a Show that for the aircraft to travel from A to B it should be pointed due east. [2]

b After flying 5 km from A to B, the aircraft returns along the same path from B to A with a resultant velocity of 13.5 m s^{-1}. Assuming that the time spent at B is negligible, calculate the average speed for the complete journey from A to B and back to A. [3]

Chapter 2: Accelerated motion

Learning outcomes

You should be able to:

- define acceleration
- draw and interpret velocity–time graphs
- derive and use the equations of uniformly accelerated motion
- describe a method for determining the acceleration due to gravity, *g*
- explain projectile motion in terms of horizontal and vertical components of motion

Quick off the mark

The cheetah (Figure 2.1) has a maximum speed of over 30 m s^{-1} (108 km/h). From a standing start a cheetah can reach 20 m s^{-1} in just three or four strides, taking only two seconds.

A car cannot increase its speed as rapidly but on a long straight road it can easily travel faster than a cheetah.



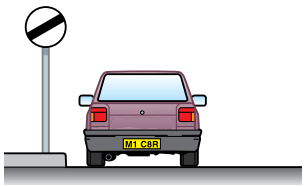
Figure 2.1 The cheetah is the world's fastest land animal. Its acceleration is impressive, too.

The meaning of acceleration

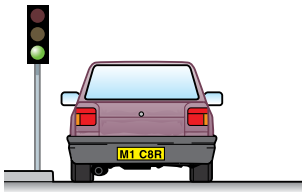
In everyday language, the term **accelerating** means 'speeding up'. Anything whose speed is increasing is accelerating. Anything whose speed is decreasing is decelerating.

To be more precise in our definition of acceleration, we should think of it as **changing velocity**. Any object whose speed is changing or which is changing its **direction** has acceleration. Because acceleration is linked to velocity in this way, it follows that it is a **vector** quantity.

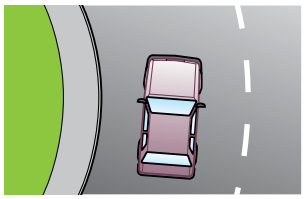
Some examples of objects accelerating are shown in Figure 2.2.



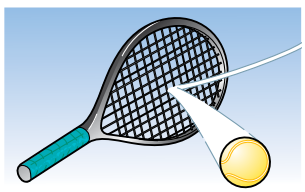
A car speeding up as it leaves the town. The driver presses on the accelerator pedal to increase the car's velocity.



A car setting off from the traffic lights. There is an instant when the car is both stationary **and** accelerating. Otherwise it would not start moving.



A car travelling round a bend at a steady speed. The car's speed is constant, but its velocity is changing as it changes direction.



A ball being hit by a tennis racket. Both the ball's speed and direction are changing. The ball's velocity changes.



A stone dropped over a cliff. Gravity makes the stone go faster and faster. The stone accelerates as it falls.

Figure 2.2 Examples of objects accelerating.

Calculating acceleration

The acceleration of something indicates the rate at which its velocity is changing. Language can get awkward here. Looking at the sprinter in Figure 2.3, we might say, 'The sprinter accelerates **faster** than the car.' However, 'faster' really means 'greater speed'. It is better to say, 'The sprinter has a greater acceleration than the car.'

Acceleration is defined as follows:

$$\text{acceleration} = \text{rate of change of velocity}$$

$$\text{average acceleration} = \frac{\text{change in velocity}}{\text{time taken}}$$

So to calculate acceleration a, we need to know two quantities – the change in velocity Δv and the time taken Δt:

$$a = \frac{\Delta v}{\Delta t}$$

Sometimes this equation is written differently. We write u for the **initial velocity** and v for the **final velocity** (because u comes before v in the alphabet). The moving object

accelerates from u to v in a time t (this is the same as the time represented by Δt above). Then the acceleration is given by the equation:

$$a = \frac{v-u}{t}$$

You must learn the definition of acceleration. It can be put in words or symbols. If you use symbols you must state what those symbols mean.

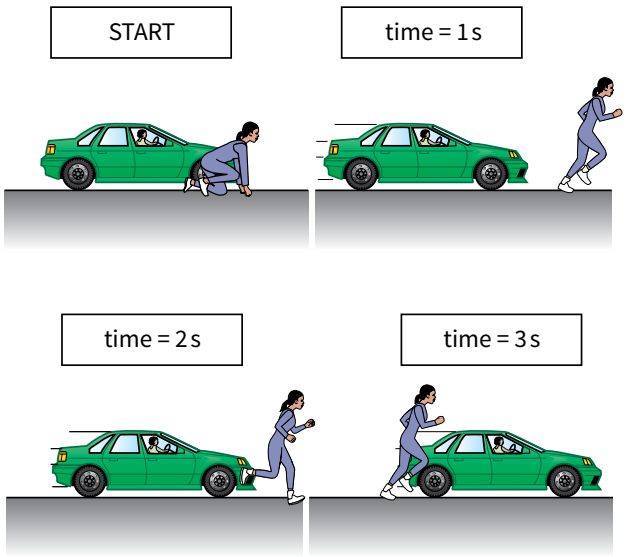


Figure 2.3 The sprinter has a greater acceleration than the car, but her top speed is less.

Units of acceleration

The unit of acceleration is $m\,s^{-2}$ (metres per second squared). The sprinter might have an acceleration of $5\,m\,s^{-2}$; her velocity increases by $5\,m\,s^{-1}$ every second. You could express acceleration in other units. For example, an advertisement might claim that a car accelerates from 0 to 60 miles per hour (mph) in 10 s. Its acceleration would then be $6\,mph\,s^{-1}$ (6 miles per hour per second). However, mixing together hours and seconds is not a good idea, and so acceleration is almost always given in the standard SI unit of $m\,s^{-2}$.

QUESTIONS

1 A car accelerates from a standing start and reaches a velocity of $18\,m\,s^{-1}$ after 6.0 s. Calculate its acceleration.

2 A car driver brakes gently. Her car slows down from $23\,m\,s^{-1}$ to $11\,m\,s^{-1}$ in 20 s. Calculate the magnitude (size) of her deceleration. (Note that, because she is slowing down, her acceleration is negative.)

3 A stone is dropped from the top of a cliff. Its acceleration is $9.81\,m\,s^{-2}$. How fast is it moving:

a after 1 s?

b after 3 s?

WORKED EXAMPLES

1 Leaving a bus stop, a bus reaches a velocity of $8.0\,m\,s^{-1}$ after 10 s. Calculate the acceleration of the bus.

Step 1 Note that the bus's initial velocity is $0\,m\,s^{-1}$. Therefore:

change in velocity $\Delta v = (8.0-0)\,m\,s^{-1}$

time taken $\Delta t = 10\,s$

Step 2 Substitute these values in the equation for acceleration:

$$\text{acceleration} = \frac{\Delta v}{\Delta t} = \frac{8.0}{10}$$

$$= 0.80\,m\,s^{-2}$$

2 A sprinter starting from rest has an acceleration of $5.0\,m\,s^{-2}$ during the first 2.0 s of a race. Calculate her velocity after 2.0 s.

Step 1 Rearranging the equation $a = \frac{v-u}{t}$ gives:

$v = u + at$

Step 2 Substituting the values and calculating gives:

$v = 0 + (5.0 \times 2.0) = 10\,m\,s^{-1}$

3 A train slows down from $60\,m\,s^{-1}$ to $20\,m\,s^{-1}$ in 50 s. Calculate the magnitude of the deceleration of the train.

Step 1 Write what you know:

$u = 60\,m\,s^{-1}$ $v = 20\,m\,s^{-1}$ $t = 50\,s$

Step 2 Take care! Here the train's final velocity is less than its initial velocity. To ensure that we arrive at the correct answer, we will use the alternative form of the equation to calculate a.

$$a = \frac{v-u}{t}$$

$$= \frac{20-60}{50} = \frac{-40}{50} = -0.80\,m\,s^{-2}$$

The minus sign (negative acceleration) indicates that the train is slowing down. It is decelerating. The magnitude of the deceleration is $0.80\,m\,s^{-2}$.

Deducing acceleration

The gradient of a velocity–time graph tells us whether the object's velocity has been changing at a high rate or a low rate, or not at all (Figure 2.4). We can deduce the value of the acceleration from the gradient of the graph:

acceleration = gradient of velocity–time graph

The graph (Figure 2.5) shows how the velocity of a cyclist changed during the start of a sprint race. We can find his acceleration during the first section of the graph (where the line is straight) using the triangle as shown.

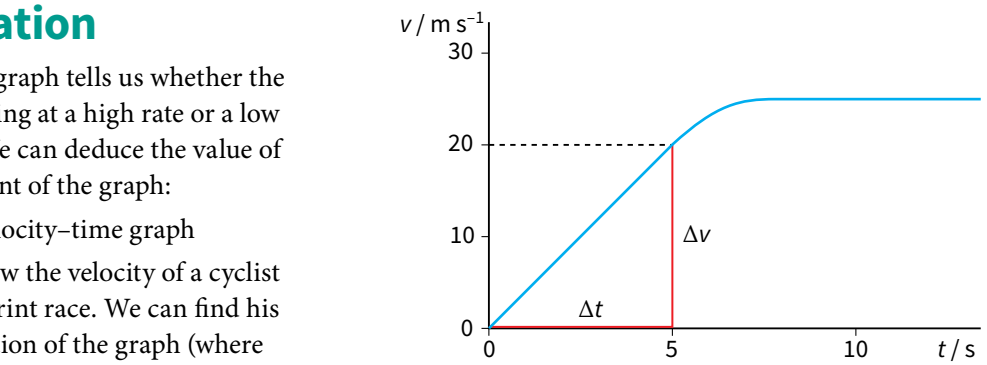


Figure 2.5 Deducing acceleration from a velocity–time graph.

The change in velocity Δv is given by the vertical side of the triangle. The time taken Δt is given by the horizontal side.

$$\text{acceleration} = \frac{\text{change in displacement}}{\text{time taken}}$$

$$= \frac{20 - 0}{5}$$

$$= 4.0\ \text{m s}^{-2}$$

A more complex example where the velocity–time graph is curved is shown on page 24.

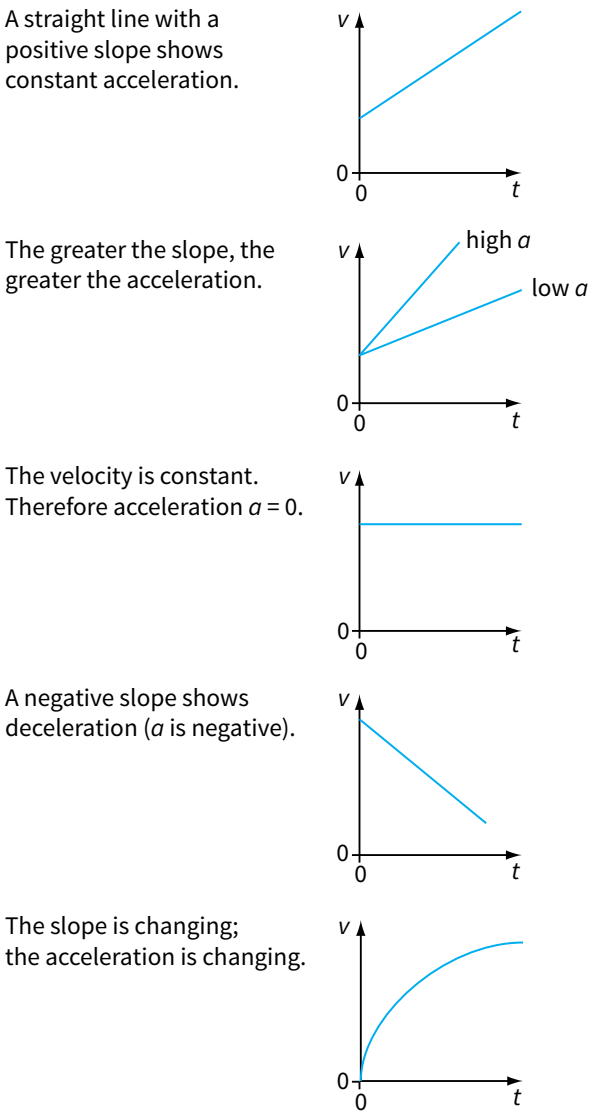


Figure 2.4 The gradient of a velocity–time graph is equal to acceleration.

Deducing displacement

We can also find the displacement of a moving object from its velocity–time graph. This is given by the area under the graph:

displacement = area under velocity–time graph

It is easy to see why this is the case for an object moving at a constant velocity. The displacement is simply velocity × time, which is the area of the shaded rectangle (Figure 2.6a).

For changing velocity, again the area under the graph gives displacement (Figure 2.6b). The area of each square of the graph represents a distance travelled: in this case, 1 m s^{-1} × 1 s, or 1 m. So, for this simple case in which the area is a triangle, we have:

$$\text{displacement} = \tfrac{1}{2} \times \text{base} \times \text{height}$$

$$= \tfrac{1}{2} \times 5.0 \times 10 = 25\ \text{m}$$

It is easy to confuse displacement–time graphs and velocity–time graphs. Check by looking at the quantity marked on the vertical axis.

For more complex graphs, you may have to use other techniques such as counting squares to deduce the area, but this is still equal to the displacement.

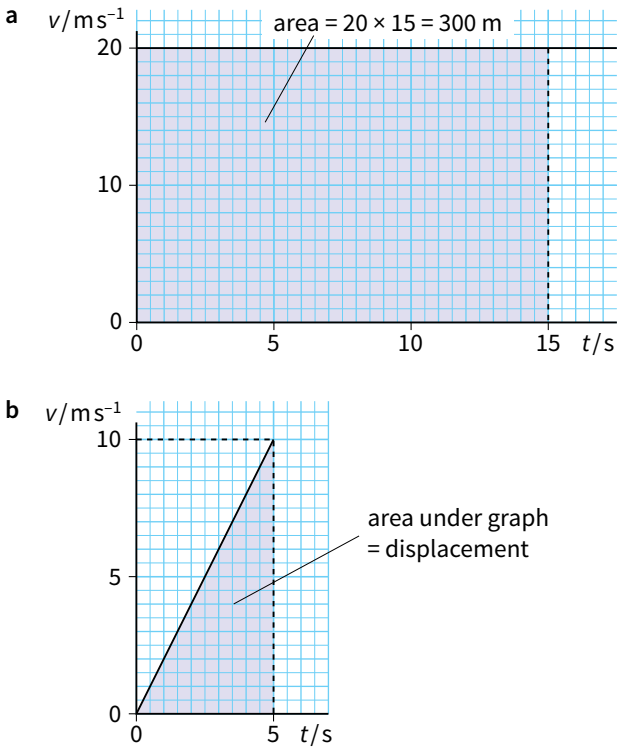


Figure 2.6 The area under the velocity–time graph is equal to the displacement of the object.

QUESTIONS

4 A lorry driver is travelling at the speed limit on a motorway. Ahead, he sees hazard lights and gradually slows down. He sees that an accident has occurred, and brakes suddenly to a halt. Sketch a velocity–time graph to represent the motion of this lorry.

5 Table 2.1 shows how the velocity of a motorcyclist changed during a speed trial along a straight road.

a Draw a velocity–time graph for this motion.

b From the table, deduce the motorcyclist's acceleration during the first 10 s.

c Check your answer by finding the gradient of the graph during the first 10 s.

d Determine the motorcyclist's acceleration during the last 15 s.

e Use the graph to find the total distance travelled during the speed trial.

Velocity / m s^{-1}	0	15	30	30	20	10	0
Time / s	0	5	10	15	20	25	30

Table 2.1 Data for a motorcyclist.

(Take care when counting squares: it is easiest when the sides of the squares stand for one unit. Check the axes, as the sides may represent 2 units, or 5 units, or some other number.)

Measuring velocity and acceleration

In a car crash, the occupants of the car may undergo a very rapid deceleration. This can cause them serious injury, but can be avoided if an air-bag is inflated within a fraction of a second. Figure 2.7 shows the tiny accelerometer at the heart of the system, which detects large accelerations and decelerations.

The acceleration sensor consists of two rows of interlocking teeth. In the event of a crash, these move relative to one another, and this generates a voltage which triggers the release of the air-bag.

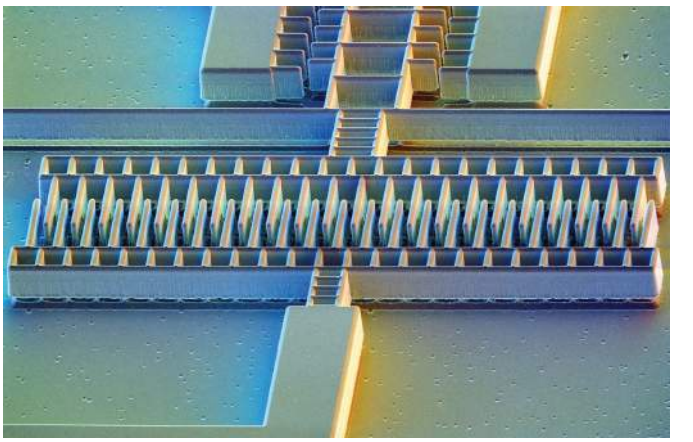


Figure 2.7 A micro-mechanical acceleration sensor is used to detect sudden accelerations and decelerations as a vehicle travels along the road. This electron microscope image shows the device magnified about 1000 times.

At the top of the photograph, you can see a second sensor which detects sideways accelerations. This is important in the case of a side impact.

These sensors can also be used to detect when a car swerves or skids, perhaps on an icy road. In this case, they activate the car's stability-control systems.

Determining velocity and acceleration in the laboratory

In Chapter 1, we looked at ways of finding the velocity of a trolley moving in a straight line. These involved measuring distance and time, and deducing velocity. Box 2.1 below shows how these techniques can be extended to find the acceleration of a trolley.

BOX 2.1: Laboratory measurements of acceleration

Measurements using light gates

The computer records the time for the first 'interrupt' section of the card to pass through the light beam of the light gate (Figure 2.8). Given the length of the interrupt, it can work out the trolley's initial velocity u. This is repeated for the second interrupt to give final velocity v. The computer also records the time interval $t_3 - t_1$ between these two velocity measurements. Now it can calculate the acceleration a as shown below:

$$u = \frac{l_1}{t_2 - t_1}$$

(l_1 = length of first section of the interrupt card)

and

$$v = \frac{l_2}{t_4 - t_3}$$

(l_2 = length of second section of the interrupt card)

Therefore:

$$a = \frac{\text{change in velocity}}{\text{time taken}}$$

$$= \frac{v - u}{t_3 - t_1}$$

(Note that this calculation gives only an approximate value for a. This is because u and v are **average** speeds over a period of time; for an accurate answer we would need to know the speeds at times t_1 and t_3.)

Sometimes two light gates are used with a card of length l. The computer can still record the times as shown above and calculate the acceleration in the same way, with $l_1 = l_2 = l$.

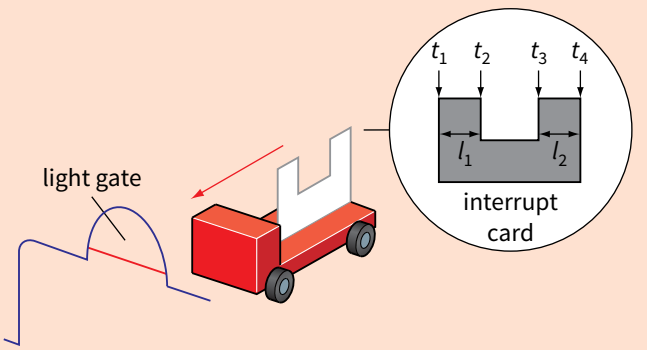


Figure 2.8 Determining acceleration using a single light gate.

Measurements using a ticker-timer

The practical arrangement is the same as for measuring velocity. Now we have to think about how to interpret the tape produced by an accelerating trolley (Figure 2.9).

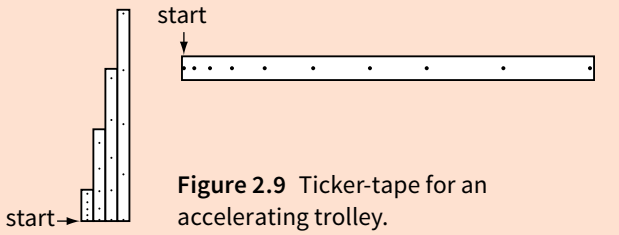


Figure 2.9 Ticker-tape for an accelerating trolley.

The tape is divided into sections, as before, every five dots. Remember that the time interval between adjacent dots is 0.02 s. Each section represents 0.10 s.

By placing the sections of tape side by side, you can picture the velocity–time graph.

The length of each section gives the trolley's displacement in 0.10 s, from which the average velocity during this time can be found. This can be repeated for each section of the tape, and a velocity–time graph drawn. The gradient of this graph is equal to the acceleration. Table 2.2 and Figure 2.10 show some typical results.

The acceleration is calculated to be:

$$a = \frac{\Delta v}{\Delta t} = \frac{0.93}{0.20} \approx 4.7\ \text{m s}^{-2}$$

Section of tape	Time at start / s	Time interval / s	Length of section / cm	Velocity / m s^{-1}
1	0.0	0.10	2.3	0.23
2	0.10	0.10	7.0	0.70
3	0.20	0.10	11.6	1.16

Table 2.2 Data for Figure 2.10.

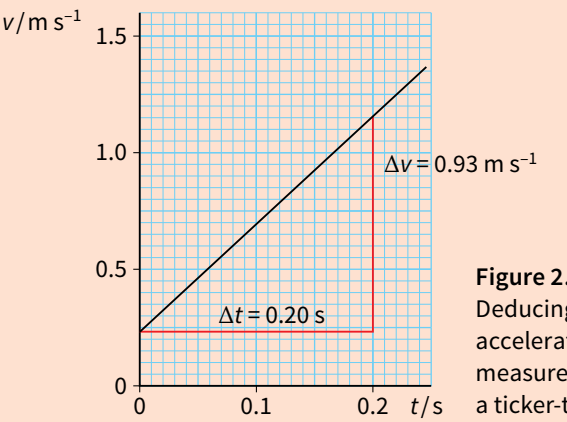


Figure 2.10 Deducing acceleration from measurements of a ticker-tape.

BOX 2.1: Laboratory measurements of acceleration (continued)

Measurements using a motion sensor

The computer software which handles the data provided by the motion sensor can calculate the acceleration of a trolley. However, because it deduces velocity from measurements of position, and then calculates acceleration from values of velocity, its precision is relatively poor.

QUESTIONS

6 Sketch a section of ticker-tape for a trolley which travels at a steady velocity and which then decelerates.

7 Figure 2.11 shows the dimensions of an interrupt card, together with the times recorded as it passed through a light gate. Use these measurements to calculate the acceleration of the card. (Follow the steps outlined on page 19.)

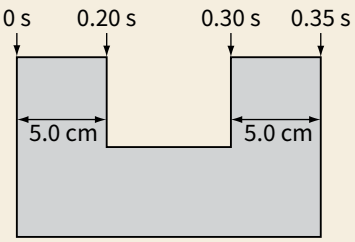


Figure 2.11 For Question 7.

8 Two adjacent five-dot sections of a ticker-tape measure 10 cm and 16 cm, respectively. The interval between dots is 0.02 s. Deduce the acceleration of the trolley which produced the tape.

The equations of motion

As a space rocket rises from the ground, its velocity steadily increases. It is accelerating (Figure 2.12).

Eventually it will reach a speed of several kilometres per second. Any astronauts aboard find themselves pushed back into their seats while the rocket is accelerating.

The engineers who planned the mission must be able to calculate how fast the rocket will be travelling and where it will be at any point in its journey. They have sophisticated computers to do this, using more elaborate versions of the equations given below.



Figure 2.12 A rocket accelerates as it lifts off from the ground.

There is a set of equations which allows us to calculate the quantities involved when an object is moving with a constant acceleration. The quantities we are concerned with are:

s	displacement	a	acceleration
u	initial velocity	t	time taken
v	final velocity		

Here are the four **equations of motion**.

equation 1: $v = u + at$

equation 2: $s = \frac{(u+v)}{2} \times t$

equation 3: $s = ut + \frac{1}{2}at^2$

equation 4: $v^2 = u^2 + 2as$

Take care using these equations. They can only be used:

- for motion in a straight line
- for an object with constant acceleration.

To get a feel for how to use these equations, we will consider some worked examples. In each example, we will follow the same procedure:

Step 1 We write down the quantities which we know, and the quantity we want to find.

Step 2 Then we choose the equation which links these quantities, and substitute in the values.

Step 3 Finally, we calculate the unknown quantity.

We will look at where these equations come from in the next section.

WORKED EXAMPLES

4 The rocket shown in Figure 2.12 lifts off from rest with an acceleration of 20 m s^{-2}. Calculate its velocity after 50 s.

Step 1 What we know: $u = 0\,\text{m s}^{-1}$, $a = 20\,\text{m s}^{-2}$, $t = 50\,\text{s}$

and what we want to know: $v = ?$

Step 2 The equation linking u, a, t and v is equation 1:

$v = u + at$

Substituting gives:

$v = 0 + (20 \times 50)$

Step 3 Calculation then gives:

$v = 1000\,\text{m s}^{-1}$

So the rocket will be travelling at 1000 m s^{-1} after 50 s. This makes sense, since its velocity increases by 20 m s^{-1} every second, for 50 s.

You could use the same equation to work out how long the rocket would take to reach a velocity of 2000 m s^{-1}, or the acceleration it must have to reach a speed of 1000 m s^{-1} in 40 s, and so on.

5 The car shown in Figure 2.13 is travelling along a straight road at 8.0 m s^{-1}. It accelerates at 1.0 m s^{-2} for a distance of 18 m. How fast is it then travelling?

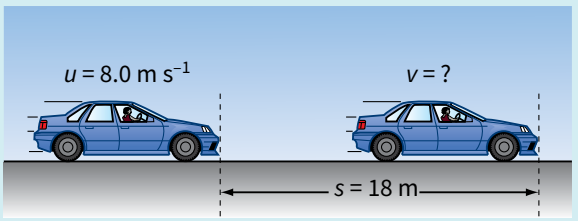


Figure 2.13 For Worked example 5. This car accelerates for a short distance as it travels along the road.

In this case, we will have to use a different equation, because we know the distance during which the car accelerates, not the time.

Step 1 What we know: $u = 8.0\,\text{m s}^{-1}$, $a = 1.0\,\text{m s}^{-2}$, $s = 18\,\text{m}$

and what we want to know: $v = ?$

Step 2 The equation we need is equation 4:

$v^2 = u^2 + 2as$

Substituting gives:

$v^2 = 8.0^2 + (2 \times 1.0 \times 18)$

Step 3 Calculation then gives:

$v^2 = 64 + 36 = 100\,\text{m}^2\,\text{s}^{-2}$

$v = 10\,\text{m s}^{-1}$

So the car will be travelling at 10 m s^{-1} when it stops accelerating.

(You may find it easier to carry out these calculations without including the units of quantities when you substitute in the equation. However, including the units can help to ensure that you end up with the correct units for the final answer.)

6 A train (Figure 2.14) travelling at 20 m s^{-1} accelerates at 0.50 m s^{-2} for 30 s. Calculate the distance travelled by the train in this time.

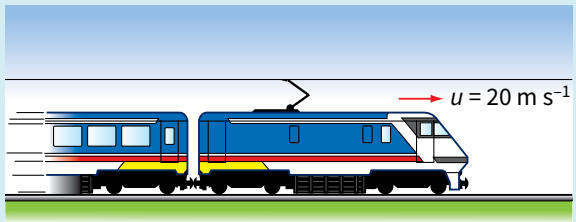


Figure 2.14 For Worked example 6. This train accelerates for 30 s.

Step 1 What we know: $u = 20\,\text{m s}^{-1}$, $t = 30\,\text{s}$, $a = 0.50\,\text{m s}^{-2}$

and what we want to know: $s = ?$

Step 2 The equation we need is equation 3:

$s = ut + \frac{1}{2}at^2$

Substituting gives:

$s = (20 \times 30) + \frac{1}{2} \times 0.5 \times (30)^2$

Step 3 Calculation then gives:

$s = 600 + 225 = 825\,\text{m}$

So the train will travel 825 m while it is accelerating.

WORKED EXAMPLES (continued)

7 The cyclist in Figure 2.15 is travelling at 15 m s^{-1}. She brakes so that she doesn't collide with the wall. Calculate the magnitude of her deceleration.

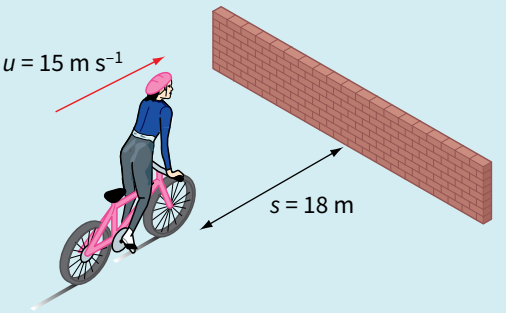


Figure 2.15 For Worked example 7. The cyclist brakes to stop herself colliding with the wall.

This example shows that it is sometimes necessary to rearrange an equation, to make the unknown quantity its subject. It is easiest to do this before substituting in the values.

Step 1 What we know: $u = 15\,\text{m s}^{-1}$
$v = 0\,\text{m s}^{-1}$
$s = 18\,\text{m}$

and what we want to know: $a = ?$

Step 2 The equation we need is equation 4:

$v^2 = u^2 + 2as$

Rearranging gives:

$$a = \frac{v^2 - u^2}{2s}$$

$$a = \frac{0^2 - 15^2}{2 \times 18} = \frac{-225}{36}$$

Step 3 Calculation then gives:

$a = -6.25\,\text{m s}^{-2} \approx -6.3\,\text{m s}^{-2}$

So the cyclist will have to brake hard to achieve a deceleration of magnitude 6.3 m s^{-2}. The minus sign shows that her acceleration is negative, i.e. a deceleration.

QUESTIONS

9 A car is initially stationary. It has a constant acceleration of 2.0 m s^{-2}.

- **a** Calculate the velocity of the car after 10 s.
- **b** Calculate the distance travelled by the car at the end of 10 s.
- **c** Calculate the time taken by the car to reach a velocity of 24 m s^{-1}.

10 A train accelerates steadily from 4.0 m s^{-1} to 20 m s^{-1} in 100 s.

- **a** Calculate the acceleration of the train.
- **b** From its initial and final velocities, calculate the average velocity of the train.
- **c** Calculate the distance travelled by the train in this time of 100 s.

11 A car is moving at 8.0 m s^{-1}. The driver makes it accelerate at 1.0 m s^{-2} for a distance of 18 m. What is the final velocity of the car?

Deriving the equations of motion

On the previous pages, we have seen how to make use of the equations of motion. But where do these equations come from? They arise from the definitions of velocity and acceleration.

We can find the first two equations from the velocity–time graph shown in Figure 2.16. The graph represents the motion of an object. Its initial velocity is u. After time t, its final velocity is v.

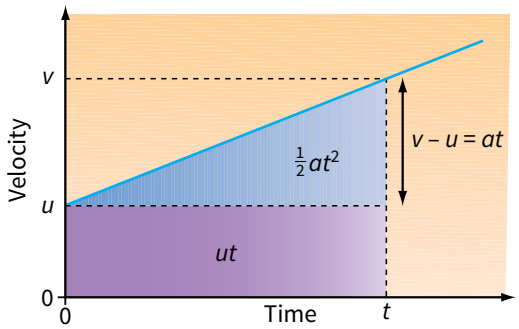


Figure 2.16 This graph shows the variation of velocity of an object with time. The object has constant acceleration.

Equation 1

The graph of Figure 2.16 is a straight line, therefore the object's acceleration a is constant. The gradient (slope) of the line is equal to acceleration.

The acceleration is defined as:

$$a = \frac{(v-u)}{t}$$

which is the gradient of the line. Rearranging this gives the first equation of motion:

$$v = u + at \qquad \text{(equation 1)}$$

Equation 2

Displacement is given by the area under the velocity–time graph. Figure 2.17 shows that the object's average velocity is half-way between u and v. So the object's average velocity, calculated by averaging its initial and final velocities, is given by:

$$\frac{(u-v)}{2}$$

The object's displacement is the shaded area in Figure 2.17. This is a rectangle, and so we have:

displacement = average velocity × time taken

and hence:

$$s = \frac{(u+v)}{2} \times t \qquad \text{(equation 2)}$$

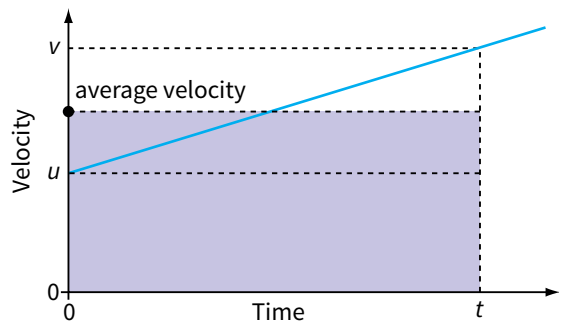


Figure 2.17 The average velocity is half-way between u and v.

Equation 3

From equations 1 and 2, we can derive equation 3:

$$v = u + at \qquad \text{(equation 1)}$$

$$s = \frac{(u+v)}{2} \times t \qquad \text{(equation 2)}$$

Substituting v from equation 1 gives:

$$s = \frac{(u+u+at)}{2} \times t$$

$$s = \frac{2ut}{2} + \frac{at^2}{2}$$

So

$$s = ut + \tfrac{1}{2}at^2 \qquad \text{(equation 3)}$$

Looking at Figure 2.16, you can see that the two terms on the right of the equation correspond to the areas of the rectangle and the triangle which make up the area under the graph. Of course, this is the same area as the rectangle in Figure 2.17.

Equation 4

Equation 4 is also derived from equations 1 and 2:

$$v = u + at \qquad \text{(equation 1)}$$

$$s = \frac{(u+u)}{2} \times t \qquad \text{(equation 2)}$$

Substituting for time t from equation 1 gives:

$$s = \frac{(u+v)}{2} + \frac{(v+u)}{a} \qquad \text{(equation 2)}$$

Rearranging this gives:

$$2as = (u+v)(v-u)$$

$$= v^2 - u^2$$

or simply:

$$v^2 = u^2 + 2as \qquad \text{(equation 4)}$$

Investigating road traffic accidents

The police frequently have to investigate road traffic accidents. They make use of many aspects of physics, including the equations of motion. The next two questions will help you to apply what you have learned to situations where police investigators have used evidence from skid marks on the road.

QUESTIONS

12 Trials on the surface of a new road show that, when a car skids to a halt, its acceleration is $-7.0\,m\,s^{-2}$. Estimate the skid-to-stop distance of a car travelling at a speed limit of $30\,m\,s^{-1}$ (approx. $110\,km\,h^{-1}$ or 70 mph).

13 At the scene of an accident on a country road, police find skid marks stretching for 50 m. Tests on the road surface show that a skidding car decelerates at $6.5\,m\,s^{-2}$. Was the car which skidded exceeding the speed limit of $25\,m\,s^{-1}$ ($90\,km\,h^{-1}$) on this road?

Uniform and non-uniform acceleration

It is important to note that the equations of motion only apply to an object which is moving with a constant acceleration. If the acceleration a was changing, you wouldn't know what value to put in the equations. Constant acceleration is often referred to as uniform acceleration.

The velocity–time graph in Figure 2.18 shows **non-uniform** acceleration. It is not a straight line; its gradient is changing (in this case, decreasing).

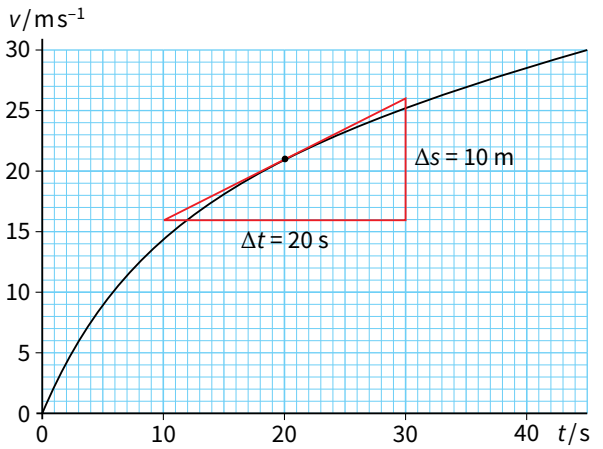


Figure 2.18 This curved velocity–time graph cannot be analysed using the equations of motion.

The acceleration at any instant in time is given by the gradient of the velocity–time graph. The triangle in Figure 2.18 shows how to find the acceleration at $t = 20$ seconds:

- At the time of interest, mark a point on the graph.
- Draw a **tangent** to the curve at that point.
- Make a large right-angled triangle, and use it to find the gradient.

You can find the change in displacement of the body as it accelerates by determining the area under the velocity–time graph.

To find the displacement of the object in Figure 2.18 between $t = 0$ and 20 s, the most straightforward, but lengthy, method is just to count the number of small squares.

In this case up to $t = 20$ s, there are about 250 small squares. This is tedious to count but you can save yourself a lot of time by drawing a line from the origin to the point at 20 s. The area of the triangle is easy to find (200 small squares) and then you only have to count the number of small squares between the line you have drawn and the curve on the graph (about 50 squares)

In this case each square is 1 m s^{-1} on the y-axis by 1 s on the x-axis, so the area of each square is $1 \times 1 = 1$ m and the displacement is 250 m. In other cases note carefully the value of each side of the square you have chosen.

QUESTIONS

14 The graph in Figure 2.19 represents the motion of an object moving with varying acceleration. Lay your ruler on the diagram so that it is tangential to the graph at point P.

a What are the values of time and velocity at this point?

b Estimate the object's acceleration at this point.

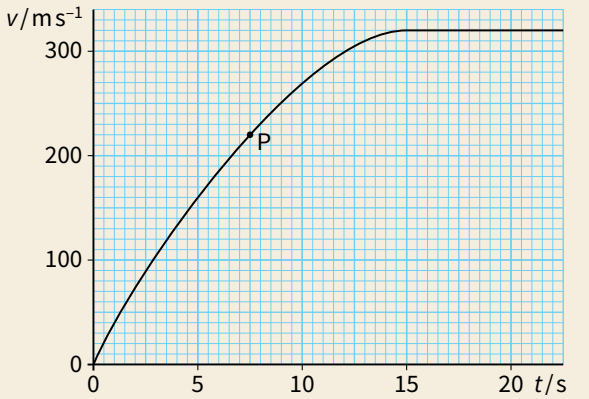


Figure 2.19 For Question 14.

15 The velocity–time graph (Figure 2.20) represents the motion of a car along a straight road for a period of 30 s.

a Describe the motion of the car.

b From the graph, determine the car's initial and final velocities over the time of 30 s.

c Determine the acceleration of the car.

d By calculating the area under the graph, determine the displacement of the car.

e Check your answer to part **d** by calculating the car's displacement using $s = ut + \frac{1}{2}at^2$.

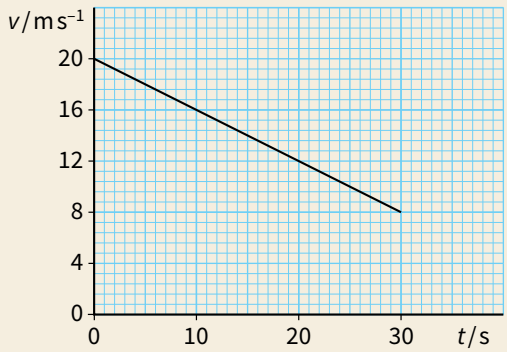


Figure 2.20 For Question 15.

Acceleration caused by gravity

If you drop a ball or stone, it falls to the ground. Figure 2.21, based on a multiflash photograph, shows the ball at equal intervals of time. You can see that the ball's velocity increases as it falls because the spaces between the images of the ball increase steadily. The ball is accelerating.

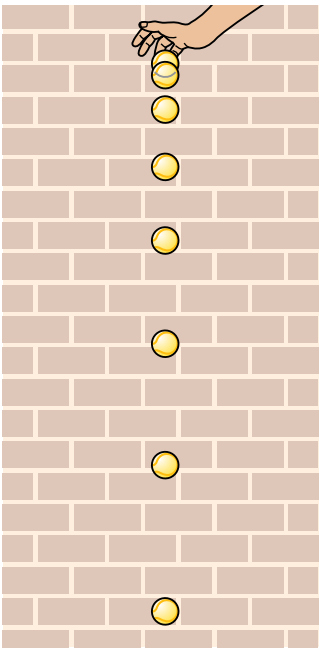


Figure 2.21 This diagram of a falling ball, based on a multiflash photo, clearly shows that the ball's velocity increases as it falls.

A multiflash photograph is useful to demonstrate that the ball accelerates as it falls. Usually, objects fall too quickly for our eyes to be able to observe them speeding up. It is easy to imagine that the ball moves quickly as soon as you let it go, and falls at a steady speed to the ground. Figure 2.21 shows that this is not the case.

If we measure the acceleration of a freely falling object on the surface of the Earth, we find a value of about $9.81\ \text{m s}^{-2}$. This is known as the **acceleration of free fall**, and is given the symbol g:

acceleration of free fall, $g = 9.81\ \text{m s}^{-2}$

The value of g depends on where you are on the Earth's surface, but we usually take $g = 9.81\ \text{m s}^{-2}$.

If we drop an object, its initial velocity $u = 0$. How far will it fall in time t? Substituting in $s = ut + \frac{1}{2}at^2$ gives displacement s:

$$s = \tfrac{1}{2} \times 9.81 \times t^2$$

$$= 4.9 \times t^2$$

Hence, by timing a falling object, we can determine g.

QUESTIONS

16 If you drop a stone from the edge of a cliff, its initial velocity $u = 0$, and it falls with acceleration $g = 9.81\ \text{m s}^{-2}$. You can calculate the distance s it falls in a given time t using an equation of motion.

- **a** Copy and complete Table 2.3, which shows how s depends on t.
- **b** Draw a graph of s against t.
- **c** Use your graph to find the distance fallen by the stone in 2.5 s.
- **d** Use your graph to find how long it will take the stone to fall to the bottom of a cliff 40 m high. Check your answer using the equations of motion.

Time / s	0	1.0	2.0	3.0	4.0
Displacement / m	0	4.9			

Table 2.3 Time (t) and displacement (s) data for Question 16.

17 An egg falls off a table. The floor is 0.8 m from the table-top.

- **a** Calculate the time taken to reach the ground.
- **b** Calculate the velocity of impact with the ground.

Determining g

One way to measure the acceleration of free fall g would be to try bungee-jumping (Figure 2.22). You would need to carry a stopwatch, and measure the time between jumping from the platform and the moment when the elastic rope begins to slow your fall. If you knew the length of the unstretched rope, you could calculate g.

There are easier methods for finding g which can be used in the laboratory. These are described in Box 2.2.



Figure 2.22 A bungee-jumper falls with initial acceleration g.

BOX 2.2: Laboratory measurements of *g*

Measuring *g* using an electronic timer

In this method, a steel ball-bearing is held by an electromagnet (Figure 2.23). When the current to the magnet is switched off, the ball begins to fall and an electronic timer starts. The ball falls through a trapdoor, and this breaks a circuit to stop the timer. This tells us the time taken for the ball to fall from rest through the distance *h* between the bottom of the ball and the trapdoor.

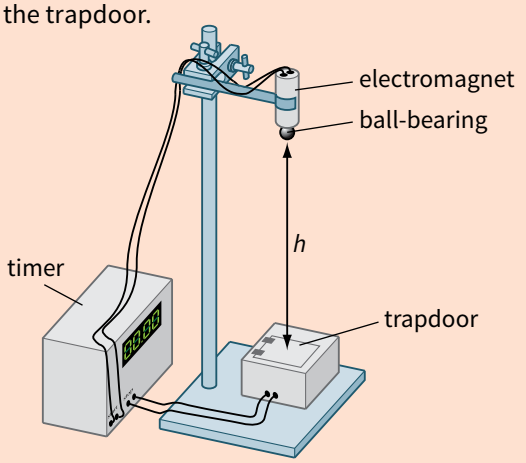


Figure 2.23 The timer records the time for the ball to fall through the distance *h*.

Here is how we can use one of the equations of motion to find *g*:

displacement $s = h$
time taken $= t$
initial velocity $u = 0$
acceleration $a = g$

Substituting in $s = ut + \frac{1}{2}at^2$ gives:

$$h = \tfrac{1}{2}gt^2$$

and for any values of *h* and *t* we can calculate a value for *g*.

A more satisfactory procedure is to take measurements of *t* for several different values of *h*. The height of the ball bearing above the trapdoor is varied systematically, and the time of fall measured several times to calculate an average for each height. Table 2.4 and Figure 2.24 show some typical results. We can deduce g from the gradient of the graph of *h* against t^2. The equation for a straight line through the origin is:

$$y = mx$$

In our experiment we have:

$h = \frac{1}{2}g\,t^2$
$y = m\,x$

h/m	t/s	t^2/s^2
0.27	0.25	0.063
0.39	0.30	0.090
0.56	0.36	0.130
0.70	0.41	0.168
0.90	0.46	0.212

Table 2.4 Data for Figure 2.24. These are mean values.

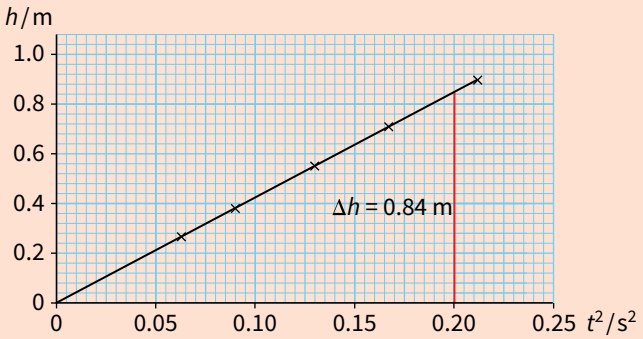


Figure 2.24 The acceleration of free fall can be determined from the gradient.

The gradient of the straight line of a graph of *h* against t^2 is equal to $\frac{g}{2}$. Therefore:

$$\text{gradient} = \frac{g}{2} = \frac{0.84}{0.20} = 4.2$$

$$g = 4.2 \times 2 = 8.4\,\text{m s}^{-2}$$

Sources of uncertainty

The electromagnet may retain some magnetism when it is switched off, and this may tend to slow the ball's fall. Consequently, the time *t* recorded by the timer may be longer than if the ball were to fall completely freely. From $h = \frac{1}{2}gt^2$, it follows that, if *t* is too great, the experimental value of *g* will be too small. This is an example of a systematic error – all the results are systematically distorted so that they are too great (or too small) as a consequence of the experimental design.

Measuring the height *h* is awkward. You can probably only find the value of *h* to within ±1 mm at best. So there is a random error in the value of *h*, and this will result in a slight scatter of the points on the graph, and a degree of uncertainty in the final value of *g*. For more about errors, see P1: Practical skills for AS.

BOX 2.2: Laboratory measurements of *g* (continued)

Measuring *g* using a ticker-timer

Figure 2.25 shows a weight falling. As it falls, it pulls a tape through a ticker-timer. The spacing of the dots on the tape increases steadily, showing that the weight is accelerating. You can analyse the tape to find the acceleration, as discussed on page 19.

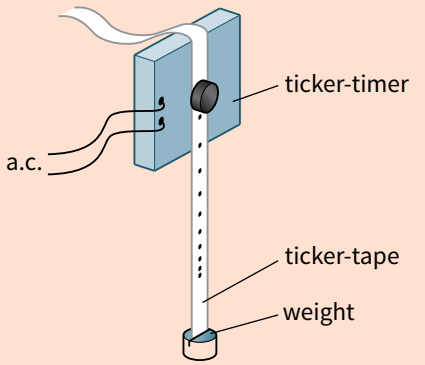


Figure 2.25 A falling weight pulls a tape through a ticker-timer.

This is not a very satisfactory method of measuring *g*. The main problem arises from friction between the tape and the ticker-timer. This slows the fall of the weight and so its acceleration is less than *g*. (This is another example of a systematic error.)

The effect of friction is less of a problem for a large weight, which falls more freely. If measurements are made for increasing weights, the value of acceleration gets closer and closer to the true value of *g*.

Measuring *g* using a light gate

Figure 2.26 shows how a weight can be attached to a card 'interrupt'. The card is designed to break the light beam twice as the weight falls. The computer can then calculate the velocity of the weight twice as it falls, and hence find its acceleration:

$$\text{initial velocity } u = \frac{x}{t_2 - t_1}$$

$$\text{final velocity } v = \frac{x}{t_4 - t_3}$$

Therefore:

$$\text{acceleration } a = \frac{v - u}{t_3 - t_1}$$

The weight can be dropped from different heights above the light gate. This allows you to find out whether its acceleration is the same at different points in its fall. This is an advantage over Method 1, which can only measure the acceleration from a stationary start.

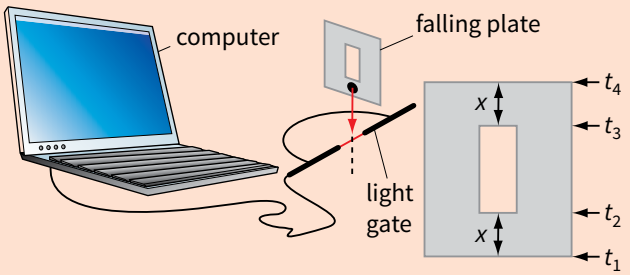


Figure 2.26 The weight accelerates as it falls. The upper section of the card falls more quickly through the light gate.

WORKED EXAMPLE

8 To get a rough value for *g*, a student dropped a stone from the top of a cliff. A second student timed the stone's fall using a stopwatch. Here are their results:

estimated height of cliff = 30 m

time of fall = 2.6 s

Use the results to estimate a value for *g*.

Step 1 Calculate the average speed of the stone:

$$\text{average speed of stone during fall} = \frac{30}{2.6} = 11.5\,\text{m s}^{-1}$$

Step 2 Find the values of *v* and *u*:

final speed $v = 2 \times 11.5\,\text{m s}^{-1} = 23.0\,\text{m s}^{-1}$

initial speed $u = 0\,\text{m s}^{-1}$

Step 3 Substitute these values into the equation for acceleration:

$$a = \frac{v - u}{t} = \frac{23.0}{2.6} = 8.8\,\text{m s}^{-2}$$

Note that you can reach the same result more directly using $s = ut + \frac{1}{2}at^2$, but you may find it easier to follow what is going on using the method given here. We should briefly consider why the answer is less than the expected value of $g = 9.81\,\text{m s}^{-2}$. It might be that the cliff was higher than the student's estimate. The timer may not have been accurate in switching the stopwatch on and off. There will have been air resistance which slowed the stone's fall.

QUESTIONS

18 A steel ball falls from rest through a height of 2.10 m. An electronic timer records a time of 0.67 s for the fall.

a Calculate the average acceleration of the ball as it falls.

b Suggest reasons why the answer is not exactly 9.81 m s^{-2}.

19 In an experiment to determine the acceleration due to gravity, a ball was timed electronically as it fell from rest through a height *h*. The times *t* shown in Table 2.5 were obtained.

a Plot a graph of *h* against t^2.

b From the graph, determine the acceleration of free fall, *g*.

c Comment on your answer.

Height / m	0.70	1.03	1.25	1.60	1.99
Time / s	0.99	1.13	1.28	1.42	1.60

Table 2.5 Height (*h*) and time (*t*) data for Question 19.

20 In Chapter 1, we looked at how to use a motion sensor to measure the speed and position of a moving object. Suggest how a motion sensor could be used to determine *g*.

Motion in two dimensions – projectiles

A curved trajectory

A multiflash photograph can reveal details of the path, or trajectory, of a projectile. Figure 2.27 shows the trajectories of a projectile – a bouncing ball. Once the ball has left the child's hand and is moving through the air, the only force acting on it is its weight.

The ball has been thrown at an angle to the horizontal. It speeds up as it falls – you can see that the images of the ball become further and further apart. At the same time, it moves steadily to the right. You can see this from the even spacing of the images across the picture. The ball's path has a mathematical shape known as a **parabola**. After it bounces, the ball is moving more slowly. It slows down, or decelerates, as it rises – the images get closer and closer together.

We interpret this picture as follows. The vertical motion of the ball is affected by the force of gravity, that is, its weight. When it rises it has a vertical deceleration

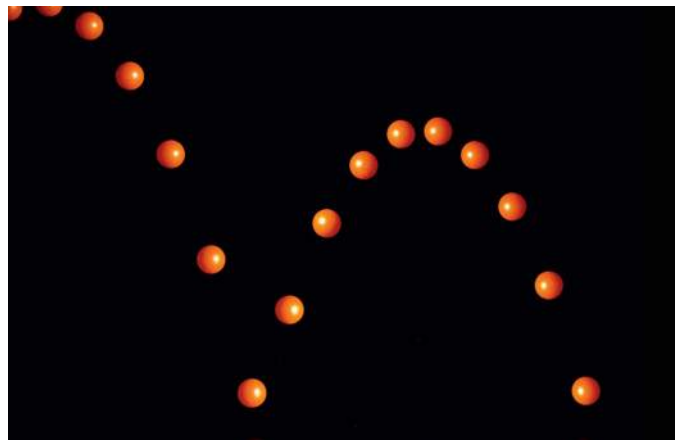


Figure 2.27 A bouncing ball is an example of a projectile. This multiflash photograph shows details of its motion which would escape the eye of an observer.

of magnitude *g*, which slows it down, and when it falls it has an acceleration of *g*, which speeds it up. The ball's horizontal motion is unaffected by gravity. In the absence of air resistance, the ball has a constant velocity in the horizontal direction. We can treat the ball's vertical and horizontal motions separately, because they are independent of one another.

Components of a vector

In order to understand how to treat the velocity in the vertical and horizontal directions separately we start by considering a constant velocity.

If an aeroplane has a constant velocity *v* at an angle θ as shown in Figure 2.28, then we say that this velocity has two effects or components, v_N in a northerly direction and v_E in an easterly direction. These two components of velocity add up to make the actual velocity *v*.

This process of taking a velocity and determining its effect along another direction is known as **resolving** the velocity along a different direction. In effect splitting the velocity into two components at right angles is the reverse

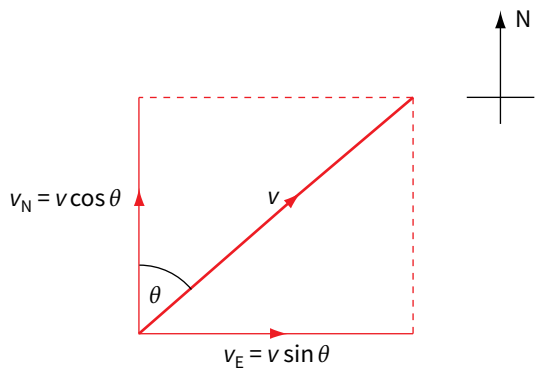


Figure 2.28 Components of a velocity. The component due north is $v_N = v\cos\theta$ and the component due east is $v_E = v\sin\theta$.

of adding together two vectors – it is splitting one vector into two vectors along convenient directions.

To find the component of any vector (e.g. displacement, velocity, acceleration) in a particular direction, we can use the following strategy:

Step 1 Find the angle θ between the vector and the direction of interest.

Step 2 Multiply the vector by the cosine of the angle θ.

So the component of an object's velocity v at angle θ to v is equal to $v\cos\theta$ (Figure 2.28).

QUESTION

21 Find the x- and y-components of each of the vectors shown in Figure 2.29. (You will need to use a protractor to measure angles from the diagram.)

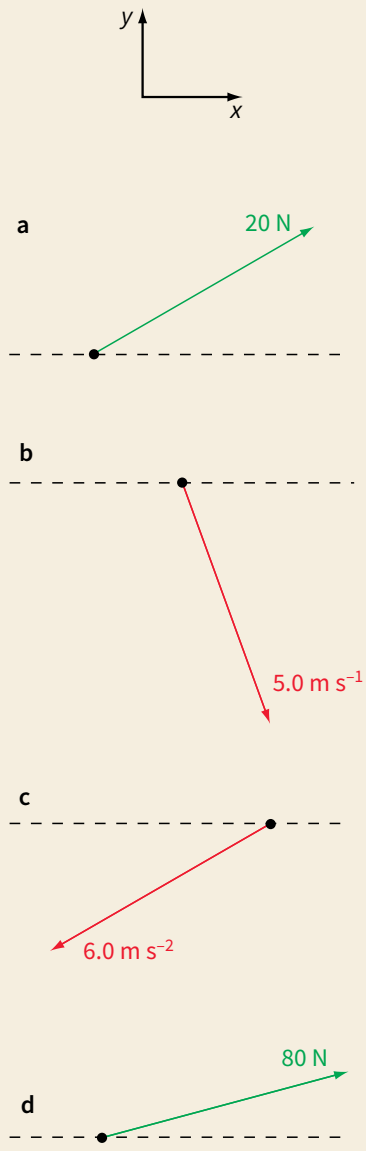


Figure 2.29 The vectors for Question 21.

Understanding projectiles

We will first consider the simple case of a projectile thrown straight up in the air, so that it moves vertically. Then we will look at projectiles which move horizontally and vertically at the same time.

Up and down

A stone is thrown upwards with an initial velocity of $20\,\text{m s}^{-1}$. Figure 2.30 shows the situation.

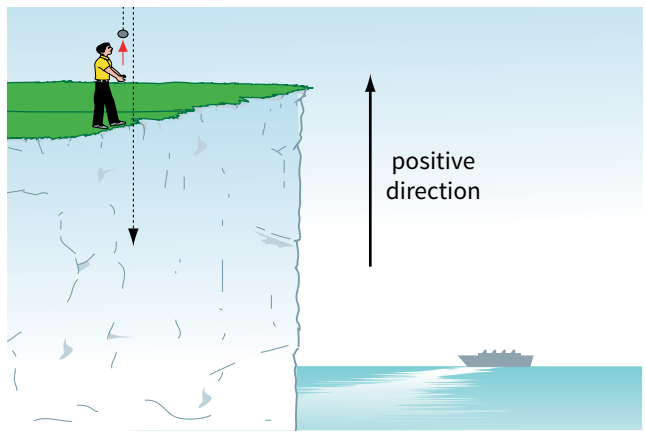


Figure 2.30 Standing at the edge of the cliff, you throw a stone vertically upwards. The height of the cliff is 25 m.

It is important to use a consistent sign convention here. We will take upwards as positive, and downwards as negative. So the stone's initial velocity is positive, but its acceleration g is negative. We can solve various problems about the stone's motion by using the equations of motion.

How high?

How high will the stone rise above ground level of the cliff?

As the stone rises upwards, it moves more and more slowly – it decelerates, because of the force of gravity. At its highest point, the stone's velocity is zero. So the quantities we know are:

initial velocity = $u = 20\,\text{m s}^{-1}$

final velocity = $v = 0\,\text{m s}^{-1}$

acceleration = $a = -9.81\,\text{m s}^{-2}$

displacement = s = ?

The relevant equation of motion is $v^2 = u^2 + 2as$.
Substituting values gives:

$$0^2 = 20^2 + 2\times(-9.81)\times s$$

$$0 = 400 - 19.62s$$

$$s = \frac{400}{19.62} = 20.4\,\text{m} \approx 20\,\text{m}$$

The stone rises 20 m upwards, before it starts to fall again.

How long?

How long will it take from leaving your hand for the stone to fall back to the clifftop?

When the stone returns to the point from which it was thrown, its displacement s is zero. So:

$s = 0 \quad u = 20\,\text{m s}^{-1} \quad a = -9.81\,\text{m s}^{-2} \quad t = ?$

Substituting in $s = ut + \frac{1}{2}at^2$ gives:

$0 = 20t \times \frac{1}{2}(-9.81) \times t^2$

$= 20t - 4.905t^2 = (20 - 4.905t) \times t$

There are two possible solutions to this:

- $t = 0$ s, i.e. the stone had zero displacement at the instant it was thrown
- $t = 4.1$ s, i.e. the stone returned to zero displacement after 4.1 s, which is the answer we are interested in.

Falling further

The height of the cliff is 25 m. How long will it take the stone to reach the foot of the cliff?

This is similar to the last example, but now the stone's final displacement is 25 m below its starting point. By our sign convention, this is a negative displacement, and $s = -25$ m.

QUESTIONS

22 In the example above (Falling further), calculate the time it will take for the stone to reach the foot of the cliff.

23 A ball is fired upwards with an initial velocity of 30 m s^{-1}. Table 2.6 shows how the ball's velocity changes. (Take $g = 9.81\,\text{m s}^{-2}$.)

a Copy and complete the table.

b Draw a graph to represent the data.

c Use your graph to deduce how long the ball took to reach its highest point.

Velocity / m s^{-1}	30	20.19				
Time / s	0	1.0	2.0	3.0	4.0	5.0

Table 2.6 For Question 23.

Vertical and horizontal at the same time

Here is an example to illustrate what happens when an object travels vertically and horizontally at the same time.

In a toy, a ball-bearing is fired horizontally from a point 0.4 m above the ground. Its initial velocity is 2.5 m s^{-1}. Its positions at equal intervals of time have been calculated and are shown in Table 2.7. These results are also shown in Figure 2.31. Study the table and the graph. You should notice the following:

- The horizontal distance increases steadily. This is because the ball's horizontal motion is unaffected by the force of gravity. It travels at a steady velocity horizontally so we can use $v = \frac{s}{t}$.
- The vertical distances do not show the same pattern. The ball is accelerating downwards so we must use the equations of motion. (These figures have been calculated using $g = 9.81\,\text{m s}^{-2}$.)

Time / s	**Horizontal distance / m**	**Vertical distance / m**
0.00	0.00	0.000
0.04	0.10	0.008
0.08	0.20	0.031
0.12	0.30	0.071
0.16	0.40	0.126
0.20	0.50	0.196
0.24	0.60	0.283
0.28	0.70	0.385

Table 2.7 Data for the example of a moving ball, as shown in Figure 2.31.

You can calculate the distance s fallen using the equation of motion $s = ut + \frac{1}{2}at^2$. (The initial vertical velocity $u = 0$.)

The horizontal distance is calculated using:

horizontal distance $= 2.5 \times t$

The vertical distance is calculated using:

vertical distance $= \frac{1}{2} \times 9.81 \times t^2$

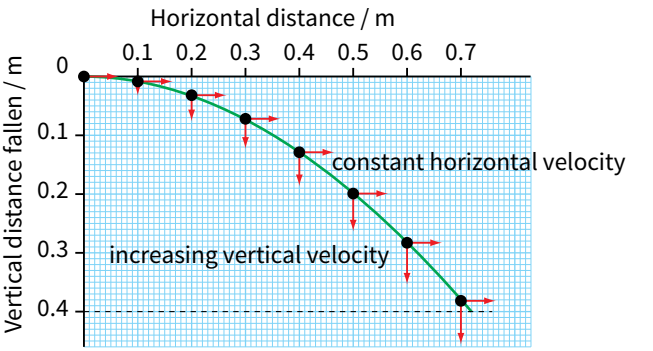


Figure 2.31 This sketch shows the path of the ball projected horizontally. The arrows represent the horizontal and vertical components of its velocity.

WORKED EXAMPLES

9 A stone is thrown horizontally with a velocity of 12 m s^{-1} from the top of a vertical cliff.

Calculate how long the stone takes to reach the ground 40 m below and how far the stone lands from the base of the cliff.

Step 1 Consider the ball's vertical motion. It has zero initial speed vertically and travels 40 m with acceleration 9.81 m s^{-2} in the same direction.

$s = ut + \frac{1}{2}at^2$

$40 = 0 + \frac{1}{2} \times 9.81 \times t^2$

Thus $t = 2.86$ s.

Step 2 Consider the ball's horizontal motion. The ball travels with a constant horizontal velocity, 12 m s^{-1}, as long as there is no air resistance.

distance travelled = $u \times t = 12 \times 2.86 = 34.3$ m

Hint: *You may find it easier to summarise the information like this:*

vertically	$s = 40$	$u = 0$	$a = 9.81$	$t = ?$	$v = ?$
horizontally	$u = 12$	$v = 12$	$a = 0$	$t = ?$	$s = ?$

10 A ball is thrown with an initial velocity of 20 m s^{-1} at an angle of 30° to the horizontal (Figure 2.32). Calculate the horizontal distance travelled by the ball (its **range**).

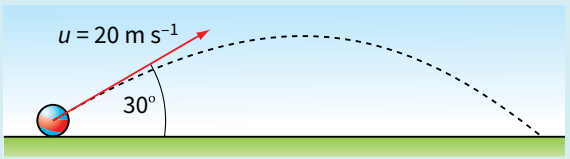


Figure 2.32 Where will the ball land?

Step 1 Split the ball's initial velocity into horizontal and vertical components:

initial velocity = $u = 20$ m s^{-1}

horizontal component of initial velocity

$= u\cos\theta = 20 \times \cos 30° = 17.3$ m s^{-1}

vertical component of initial velocity

$= u\sin\theta = 20 \times \sin 30° = 10$ m s^{-1}

Step 2 Consider the ball's vertical motion. How long will it take to return to the ground? In other words, when will its displacement return to zero?

$u = 10$ m s^{-1} $a = -9.81$ m s^{-2} $s = 0$ $t = ?$

Using $s = ut + \frac{1}{2}at^2$, we have:

$0 = 10t - 4.905t^2$

This gives $t = 0$ s or $t = 2.04$ s. So the ball is in the air for 2.04 s.

Step 3 Consider the ball's horizontal motion. How far will it travel horizontally in the 2.04 s before it lands? This is simple to calculate, since it moves with a constant horizontal velocity of 17.3 m s^{-1}.

horizontal displacement $s = 17.3 \times 2.04$

$= 35.3$ m

Hence the horizontal distance travelled by the ball (its range) is about 35 m.

QUESTIONS

24 A stone is thrown horizontally from the top of a vertical cliff and lands 4.0 s later at a distance 12.0 m from the base of the cliff. Ignore air resistance.

- **a** Calculate the horizontal speed of the stone.
- **b** Calculate the height of the cliff.

25 A stone is thrown with a velocity of 8 m s^{-1} into the air at an angle of 40° to the horizontal.

- **a** Calculate the vertical component of the velocity.
- **b** State the value of the vertical component of the velocity when the stone reaches its highest point. Ignore air resistance.
- **c** Use your answers to **a** and **b** to calculate the time the stone takes to reach it highest point.
- **d** Calculate the horizontal component of the velocity.
- **e** Use your answers to **c** and **d** to find the horizontal distance travelled by the stone as it climbs to its highest point.

26 The range of a projectile is the horizontal distance it travels before it reaches the ground. The greatest range is achieved if the projectile is thrown at 45° to the horizontal.

A ball is thrown with an initial velocity of 40 m s^{-1}. Calculate its greatest possible range when air resistance is considered to be negligible.

Summary

- Acceleration is equal to the rate of change of velocity.
- Acceleration is a vector quantity.
- The gradient of a velocity–time graph is equal to acceleration:

  $a = \frac{\Delta v}{\Delta t}$
- The area under a velocity–time graph is equal to displacement (or distance travelled).
- The equations of motion (for constant acceleration in a straight line) are:

  $v = u + at$ $\quad s = ut + \frac{1}{2}at^2$

  $s = \frac{(u + v)}{2}$ $\quad v^2 = u^2 + 2as$
- Vectors such as forces can be resolved into components. Components at right angles to one another can be treated independently of one another. For a velocity v at an angle θ to the x-direction, the components are:

  x-direction: $v\cos\theta$

  y-direction: $v\sin\theta$
- For projectiles, the horizontal and vertical components of velocity can be treated independently. In the absence of air resistance, the horizontal component of velocity is constant while the vertical component of velocity downwards increases at a rate of $9.81\,\text{m s}^{-2}$.

End-of-chapter questions

1 A motorway designer can assume that cars approaching a motorway enter a slip road with a velocity of $10\,\text{m s}^{-1}$ and reach a velocity of $30\,\text{m s}^{-1}$ before joining the motorway. Calculate the minimum length for the slip road, assuming that vehicles have an acceleration of $4.0\,\text{m s}^{-2}$. [4]

2 A train is travelling at $50\,\text{m s}^{-1}$ when the driver applies the brakes and gives the train a constant deceleration of magnitude $0.50\,\text{m s}^{-2}$ for 100 s. Describe what happens to the train. Calculate the distance travelled by the train in 100 s. [7]

3 A boy stands on a cliff edge and throws a stone vertically upwards at time $t = 0$. The stone leaves his hand at $20\,\text{m s}^{-1}$. Take the acceleration of the ball as $9.81\,\text{m s}^{-2}$.

a Show that the equation for the displacement of the ball is:
$s = 20t - 4.9t^2$ [2]

b What is the height of the stone 2.0 s after release and 6.0 s after release? [3]

c When does the stone return to the level of the boy's hand? Assume the boy's hand does not move vertically after the ball is released. [4]

4 The graph in Figure 2.33 shows the variation of velocity with time of two cars, A and B, which are travelling in the same direction over a period of time of 40 s. Car A, travelling at a constant velocity of 40 m s^{-1}, overtakes car B at time $t = 0$. In order to catch up with car A, car B immediately accelerates uniformly for 20 s to reach a constant velocity of 50 m s^{-1}. Calculate:

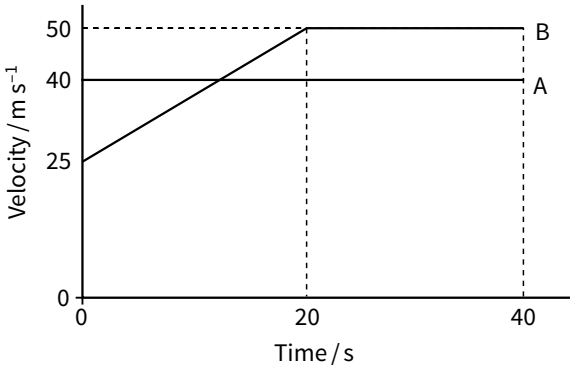


Figure 2.33 Velocity–time graphs for two cars, A and B. For End-of-chapter Question 4.

a how far A travels during the first 20 s [2]

b the acceleration and distance of travel of B during the first 20 s [5]

c the additional time taken for B to catch up with A [2]

d the distance each car will have then travelled since $t = 0$. [2]

5 An athlete competing in the long jump leaves the ground with a velocity of 5.6 m s^{-1} at an angle of 30° to the horizontal.

a Determine the vertical component of the velocity and use this value to find the time between leaving the ground and landing. [2]

b Determine the horizontal component of the velocity and use this value to find the horizontal distance travelled. [4]

6 Figure 2.34 shows an arrangement used to measure the acceleration of a metal plate as it falls vertically. The metal plate is released from rest and falls a distance of 0.200 m before breaking light beam 1. It then falls a further 0.250 m before breaking light beam 2.

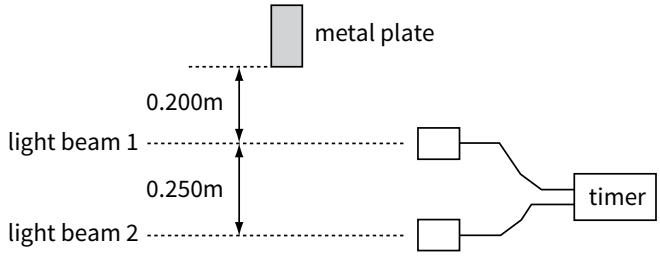


Figure 2.34 For End-of-chapter Question 6.

a Calculate the time taken for the plate to fall 0.200 m from rest. (You may assume that the metal plate falls with an acceleration equal to the acceleration of free fall.) [2]

b The timer measures the speed of the metal plate as it falls through each light beam. The speed as it falls through light beam 1 is 1.92 m s^{-1} and the speed as it falls through light beam 2 is 2.91 m s^{-1}.

i Calculate the acceleration of the plate between the two light beams. [2]

ii State and explain **one** reason why the acceleration of the plate is not equal to the acceleration of free fall. [2]

7 Figure 2.35 shows the velocity–time graph for a vertically bouncing ball. The ball is released at A and strikes the ground at B. The ball leaves the ground at D and reaches its maximum height at E. The effects of air resistance can be neglected.

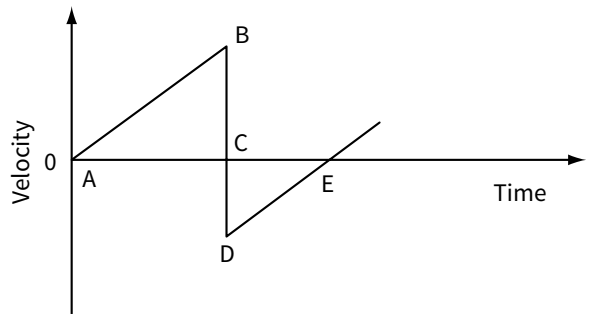


Figure 2.35 For End-of-chapter Question 7.

a State:

i why the velocity at D is negative [1]

ii why the gradient of the line AB is the same as the gradient of line DE [1]

iii what is represented by the area between the line AB and the time axis [1]

iv why the area of triangle ABC is greater than the area of triangle CDE. [1]

b The ball is dropped from rest from an initial height of 1.2 m. After hitting the ground the ball rebounds to a height of 0.80 m. The ball is in contact with the ground between B and D for a time of 0.16 s.

Using the acceleration of free fall, calculate:

i the speed of the ball immediately before hitting the ground [2]

ii the speed of the ball immediately after hitting the ground [2]

iii the acceleration of the ball while it is in contact with the ground. State the direction of this acceleration. [3]

8 A student measures the speed v of a trolley as it moves down a slope. The variation of v with time t is shown in the graph in Figure 2.36.

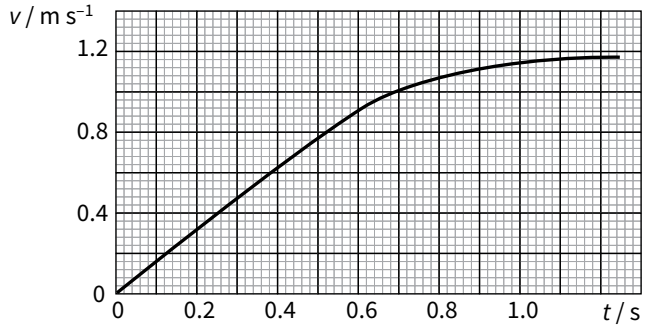


Figure 2.36 For End-of-chapter Question 8.

a Use the graph to find the acceleration of the trolley when $t = 0.7$ s. [2]

b State how the acceleration of the trolley varies between $t = 0$ and $t = 1.0$ s. Explain your answer by reference to the graph. [3]

c Determine the distance travelled by the trolley between $t = 0.6$ and $t = 0.8$ s. [3]

d The student obtained the readings for v using a motion sensor. The readings may have random errors and systematic errors. Explain how these two types of error affect the velocity–time graph. [2]

9 A car driver is travelling at speed v on a straight road. He comes over the top of a hill to find a fallen tree on the road ahead. He immediately brakes hard but travels a distance of 60 m at speed v before the brakes are applied. The skid marks left on the road by the wheels of the car are of length 140 m (Figure 2.37). The police investigate whether the driver was speeding and establish that the car decelerates at 2.0 m s^{-2} during the skid.

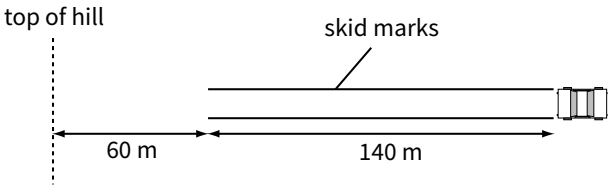


Figure 2.37 For End-of-chapter Question 9.

a Determine the initial speed v of the car before the brakes are applied. [2]

b Determine the time taken between the driver coming over the top of the hill and applying the brakes. Suggest whether this shows whether the driver was alert to the danger. [2]

c The speed limit on the road is 100 km/h. Determine whether the driver was breaking the speed limit. [2]

10 A hot-air balloon rises vertically. At time $t = 0$, a ball is released from the balloon. Figure 2.38 shows the variation of the ball's velocity v with t. The ball hits the ground at $t = 4.1$ s.

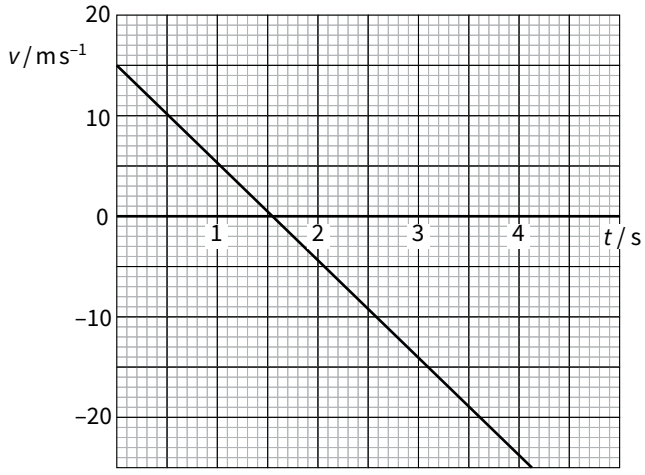


Figure 2.38 For End-of-chapter Question 10.

a Explain how the graph shows that the acceleration of the ball is constant. [1]

b Use the graph to:

i determine the time at which the ball reaches its highest point [1]

ii show that the ball rises for a further 12 m between release and its highest point [2]

iii determine the distance between the highest point reached by the ball and the ground. [2]

c The equation relating v and t is $v = 15 - 9.81t$. Explain the significance in the equation of:

i the number 15 [1]

ii the negative sign. [1]

11 An aeroplane is travelling horizontally at a speed of 80 m s^{-1} and drops a crate of emergency supplies (Figure 2.39). To avoid damage, the maximum vertical speed of the crate on landing is 20 m s^{-1}.
You may assume air resistance is negligible.

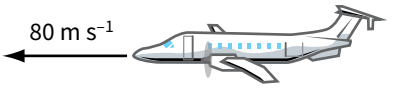


Figure 2.39 For End-of-chapter Question 11.

a Calculate the maximum height of the aeroplane when the crate is dropped. [2]

b Calculate the time taken for the crate to reach the ground from this height. [2]

c The aeroplane is travelling at the maximum permitted height. Calculate the horizontal distance travelled by the crate after it is released from the aeroplane. [1]



Chapter 3: Dynamics – explaining motion

Learning outcomes

You should be able to:

- state Newton's laws of motion
- identify the forces acting on a body in different situations
- describe how the motion of a body is affected by the forces acting on it
- solve problems using $F = ma$
- relate derived units to base units in the SI system

Force and acceleration

If you have ever flown in an aeroplane you will know how the back of the seat pushes you forwards when the aeroplane accelerates down the runway (Figure 3.1). The pilot must control many forces on the aeroplane to ensure a successful take-off.

In Chapters 1 and 2 we saw how motion can be **described** in terms of displacement, velocity, acceleration and so on. This is known as kinematics. Now we are going to look at how we can **explain** how an object moves in terms of the forces which change its motion. This is known as dynamics.



Figure 3.1 An aircraft takes off – the force provided by the engines causes the aircraft to accelerate.

Calculating the acceleration

Figure 3.2a shows how we represent the force which the motors on a train provide to cause it to accelerate. The resultant force is represented by a green arrow. The direction of the arrow shows the direction of the resultant force. The magnitude (size) of the resultant force of 20 000 N is also shown.

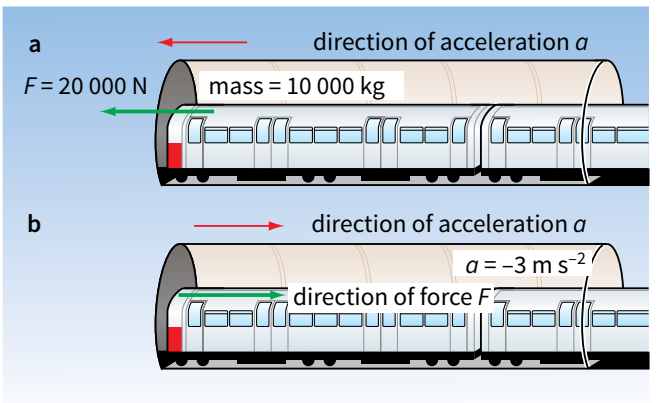


Figure 3.2 A force is needed to make the train **a** accelerate, and **b** decelerate.

To calculate the acceleration a of the train produced by the resultant force F, we must also know the train's mass m (Table 3.1). These quantities are related by:

$$a = \frac{F}{m} \quad \text{or} \quad F = ma$$

Quantity	Symbol	Unit
resultant force	F	N (newtons)
mass	m	kg (kilograms)
acceleration	a	$m\,s^{-2}$ (metres per second squared)

Table 3.1 The quantities related by $F = ma$.

In this example we have $F = 20\,000\,N$ and $m = 10\,000\,kg$, and so:

$$a = \frac{F}{m} = \frac{10\,000}{10\,000} = 2\,m\,s^{-2}$$

In Figure 3.2b, the train is decelerating as it comes into a station. Its acceleration is $-3.0\,m\,s^{-2}$. What force must be provided by the braking system of the train?

$$F = ma = 10\,000 \times -3 = -30\,000\,N$$

The minus sign shows that the force must act towards the right in the diagram, in the opposite direction to the motion of the train.

Force, mass and acceleration

The equation we used above, $F = ma$, is a simplified version of Newton's second law of motion.

> For a body of constant mass, its acceleration is directly proportional to the resultant force applied to it.

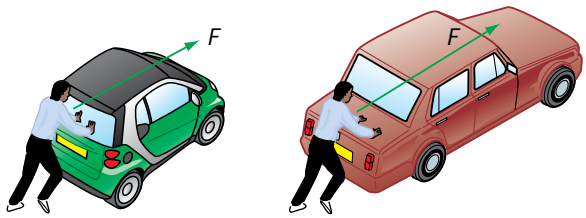
An alternative form of Newton's second law is given in Chapter 6 when you have studied momentum. Since Newton's second law holds for objects that have a constant mass, this equation can be applied to a train whose mass remains constant during its journey. The equation $a = \frac{F}{m}$ relates acceleration, resultant force and mass. In particular, it shows that the bigger the force, the greater the acceleration it produces. You will probably feel that this is an unsurprising result. For a given object, the acceleration is directly proportional to the resultant force:

$$a \propto F$$

The equation also shows that the acceleration produced by a force depends on the mass of the object. The **mass** of an object is a measure of its **inertia**, or its ability to resist any change in its motion. The greater the mass, the smaller the acceleration which results. If you push your hardest against a small car (which has a small mass), you will have a greater effect than if you push against a more massive car (Figure 3.3). So, for a constant force, the acceleration is inversely proportional to the mass:

$$a \propto \frac{1}{m}$$

The train driver knows that, when the train is full during the rush hour, it has a smaller acceleration. This is because its mass is greater when it is full of people. Similarly, it is more difficult to stop the train once it is moving. The brakes must be applied earlier to avoid the train overshooting the platform at the station.



mass m = 700 kg mass m = 2600 kg

Figure 3.3 It is easier to make a small mass accelerate than a large mass.

WORKED EXAMPLES

1 A cyclist of mass 60 kg rides a bicycle of mass 20 kg. When starting off, the cyclist provides a force of 200 N. Calculate the initial acceleration.

Step 1 This is a straightforward example. First, we must calculate the combined mass m of the bicycle and its rider:

$m = 20 + 60 = 80$ kg

We are given the force F:

force causing acceleration $F = 200$ N

Step 2 Substituting these values gives:

$$a = \frac{F}{m} = \frac{200}{80} = 2.5\,\text{m s}^{-2}$$

So the cyclist's acceleration is $2.5\,\text{m s}^{-2}$.

2 A car of mass 500 kg is travelling at $20\,\text{m s}^{-1}$. The driver sees a red traffic light ahead, and slows to a halt in 10 s. Calculate the braking force provided by the car.

Step 1 In this example, we must first calculate the acceleration required. The car's final velocity is $0\,\text{m s}^{-1}$, so its change in velocity $\Delta v = 0 - 20 = -20\,\text{m s}^{-1}$

$$\text{acceleration } a = \frac{\text{change in velocity}}{\text{time taken}}$$

$$= \frac{\Delta v}{\Delta t} = \frac{-20}{10} = -2\,\text{m s}^{-2}$$

Step 2 To calculate the force, we use:

$F = ma = 500 \times -2 = -1000$ N

So the brakes must provide a force of 1000 N. (The minus sign shows a force decreasing the velocity of the car.)

QUESTIONS

1 Calculate the force needed to give a car of mass 800 kg an acceleration of $2.0\,\text{m s}^{-2}$.

2 A rocket has a mass of 5000 kg. At a particular instant, the resultant force acting on the rocket is 200 000 N. Calculate its acceleration.

3 (In this question, you will need to make use of the equations of motion which you studied in Chapter 2.) A motorcyclist of mass 60 kg rides a bike of mass 40 kg. As she sets off from the lights, the forward force on the bike is 200 N. Assuming the resultant force on the bike remains constant, calculate the bike's velocity after 5.0 s.

Understanding SI units

Any quantity that we measure or calculate consists of a value and a unit. In physics, we mostly use units from the SI system. These units are all defined with extreme care, and for a good reason. In science and engineering, every measurement must be made on the same basis, so that measurements obtained in different laboratories can be compared. This is important for commercial reasons, too. Suppose an engineering firm in Taiwan is asked to produce a small part for the engine of a car which is to be assembled in India. The dimensions are given in millimetres and the part must be made with an accuracy of a tiny fraction of a millimetre. All concerned must know that the part will fit correctly – it wouldn't be acceptable to use a different millimetre scale in Taiwan and India.

Engineering measurements, as well as many other technical measurements, are made using SI units to ensure that customers get what they expected (and can complain if they don't). So governments around the world have set up standards laboratories to ensure that measuring instruments are as accurate as is required – scales weigh correctly, police speed cameras give reliable measurements, and so on. (Other, non-SI, units such as the foot, pound or hour, are defined in terms of SI units.)

Base units, derived units

The metre, kilogram and second are three of the seven SI **base units**. These are defined with great precision so that every standards laboratory can reproduce them correctly.

Other units, such as units of speed ($m\,s^{-1}$) and acceleration ($m\,s^{-2}$) are known as **derived units** because they are combinations of base units. Some derived units, such as the newton and the joule, have special names which are more convenient to use than giving them in terms of base units. The definition of the newton will show you how this works.

Defining the newton

Isaac Newton (1642–1727) played a significant part in developing the scientific idea of force. Building on Galileo's earlier thinking, he explained the relationship between force, mass and acceleration, which we now write as $F = ma$. For this reason, the SI unit of force is named after him.

We can use the equation $F = ma$ to define the **newton** (N).

One newton is the force that will give a 1 kg mass an acceleration of $1\,m\,s^{-2}$ in the direction of the force.

$1\,N = 1\,kg \times 1\,m\,s^{-2}$ or $1\,N = 1\,kg\,m\,s^{-2}$

The seven base units

In mechanics (the study of forces and motion), the units we use are based on three base units: the metre, kilogram and second. As we move into studying electricity, we will need to add another base unit, the ampere. Heat requires another base unit, the kelvin (the unit of temperature).

Table 3.2 shows the seven base units of the SI system. Remember that all other units can be derived from these seven. The equations that relate them are the equations that you will learn as you go along (just as $F = ma$ relates the newton to the kilogram, metre and second). The unit of luminous intensity is not part of the A/AS course.

Base unit	Symbol	Base unit
length	*x*, *l*, *s* etc.	m (metre)
mass	*m*	kg (kilogram)
time	*t*	s (second)
electric current	*I*	A (ampere)
thermodynamic temperature	*T*	K (kelvin)
amount of substance	*n*	mol (mole)
luminous intensity	*I*	cd (candela)

Table 3.2 SI base quantities and units. In this course, you will learn about all of these except the candela.

QUESTION

4 The pull of the Earth's gravity on an apple (its weight) is about 1 newton. We could devise a new international system of units by defining our unit of force as the weight of an apple. State as many reasons as you can why this would not be a very useful definition.

Other SI units

Using only seven base units means that only this number of quantities have to be defined with great precision. There would be confusion and possible contradiction if more units were also defined. For example, if the density of water were **defined** as exactly $1\,g\,cm^{-3}$, then $1000\,cm^3$ of a sample of water would have a mass of exactly 1 kg. However, it is unlikely that the mass of this volume of water would equal exactly the mass of the standard kilogram. The standard kilogram, which is kept in France, is the one standard from which all masses can ultimately be measured.

All other units can be derived from the base units. This is done using the definition of the quantity. For example, speed is defined as $\frac{\text{distance}}{\text{time}}$, and so the base units of speed in the SI system are $\mathbf{m\,s^{-1}}$.

Since the defining equation for force is $F = ma$, the base units for force are $\mathbf{kg\,m\,s^{-2}}$.

Equations that relate different quantities must have the same base units on each side of the equation. If this does not happen the equation must be wrong.

When each term in an equation has the same base units the equation is said to be **homogeneous**.

QUESTIONS

5 Determine the base units of:

a pressure $\left(= \frac{\text{force}}{\text{area}}\right)$

b energy (= force × distance)

c density $\left(= \frac{\text{mass}}{\text{volume}}\right)$

6 Use base units to prove that the following equations are homogeneous.

a pressure
= density × acceleration due to gravity × depth

b distance travelled
= initial speed × time + $\frac{1}{2}$ acceleration × time^2
($s = ut + \frac{1}{2}at^2$)

WORKED EXAMPLE

3 It is suggested that the time T for one oscillation of a swinging pendulum is given by the equation $T^2 = 4\pi^2(l/g)$ where l is the length of the pendulum and g is the acceleration due to gravity. Show that this equation is homogeneous.

For the equation to be homogeneous, the term on the left-hand side must have the same base units as all the terms on the right-hand side.

Step 1 The base unit of time T is **s**. The base unit of the left-hand side of the equation is therefore $\mathbf{s^2}$.

Step 2 The base unit of l is **m**. The base units of g are $\mathbf{m\,s^{-2}}$. Therefore the base unit of the right-hand side is $\frac{\mathbf{m}}{(\mathbf{m\,s^{-2}})} = \mathbf{s^2}$. (Notice that the constant $4\pi^2$ has no units.)

Since the base units on the left-hand side of the equation are the same as those on the right, the equation is homogeneous.

Prefixes

Each unit in the SI system can have **multiples** and **sub-multiples** to avoid using very high or low numbers. For example 1 millimetre (mm) is one thousandth of a metre and 1 micrometre (μm) is one millionth of a metre.

The **prefix** comes before the unit. In the unit mm, the first m is the prefix milli and the second m is the unit metre. You will need to recognise a number of prefixes for the A/AS course, as shown in Table 3.3.

Multiples			Sub-multiples		
Multiple	**Prefix**	**Symbol**	**Multiple**	**Prefix**	**Symbol**
10^3	kilo	k	10^{-1}	deci	d
10^6	mega	M	10^{-2}	centi	c
10^9	giga	G	10^{-3}	mill	m
10^{12}	tera	T	10^{-6}	micro	μ
10^{15}	peta	P	10^{-9}	nano	n
			10^{-12}	pico	p

Table 3.3 Multiples and sub-multiples.

You must take care when using prefixes.

- Squaring or cubing prefixes – for example:
  $1\,\text{cm} = 10^{-2}\,\text{m}$
  so $1\,\text{cm}^2 = (10^{-2}\,\text{m})^2 = 10^{-4}\,\text{m}^2$
  and $1\,\text{cm}^3 = (10^{-2}\,\text{m})^3 = 10^{-6}\,\text{m}^3$.
- Writing units – for example, you must leave a small space between each unit when writing a speed such as $3\,\text{m\,s}^{-1}$, because if you write it as $3\,\text{ms}^{-1}$ it would mean 3 millisecond^{-1}.

QUESTIONS

7 Find the area of one page of this book in cm^2 and then convert your value to m^2.

8 Write down in powers of ten the values of the following quantities:

a 60 pA

b 500 MW

c 20 000 mm

WORKED EXAMPLE

4 The density of water is $1.0\,\text{g\,cm}^{-3}$. Calculate this value in $\text{kg\,m}^{-3}$.

Step 1 Find the conversions for the units:

$1\,\text{g} = 1 \times 10^{-3}\,\text{kg}$

$1\,\text{cm}^3 = 1 \times 10^{-6}\,\text{m}^3$

Step 2 Use these in the value for the density of water:

$$1.0\,\text{g\,cm}^{-3} = \frac{1.0 \times 1 \times 10^{-3}}{1 \times 10^{-6}}$$

$$= 1.0 \times 10^3\,\text{kg\,m}^{-3}$$

The pull of gravity

Now we need to consider some specific forces – such as weight and friction.

When Isaac Newton was confined to his rural home to avoid the plague which was rampant in other parts of England, he is said to have noticed an apple fall to the ground. From this, he developed his theory of gravity which relates the motion of falling objects here on Earth to the motion of the Moon around the Earth, and the planets around the Sun.

The force which caused the apple to accelerate was the pull of the Earth's gravity. Another name for this force is the **weight** of the apple. The force is shown as an arrow, pulling vertically downwards on the apple (Figure 3.4). It is usual to show the arrow coming from the centre of the apple – its **centre of gravity**. The centre of gravity of an object is defined as the point where its entire weight appears to act.

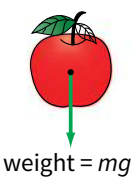


Figure 3.4 The weight of an object is a force caused by the Earth's gravity. It acts vertically down on the object.

Large and small

A large rock has a greater weight than a small rock, but if you push both rocks over a cliff at the same time, they will fall at the same rate. In other words, they have the **same** acceleration, regardless of their mass. This is a surprising result. Common sense may suggest that a heavier object will fall faster than a lighter one. It is said that Galileo dropped a large cannon ball and a small cannon ball from the top of the Leaning Tower of Pisa in Italy, and showed that they landed simultaneously. He may never actually have done this, but the story illustrates that the result is not intuitively obvious – if everyone thought that the two cannon balls would accelerate at the same rate, there would not have been any experiment or story.

In fact, we are used to lighter objects falling more slowly than heavy ones. A feather drifts down to the floor, while a stone falls quickly. However, we are being misled by the presence of **air resistance**. The force of air resistance has a large effect on the falling feather, and almost no effect on the falling stone. When astronauts visited the Moon (where there is virtually no atmosphere and so no air resistance), they were able to show that a feather and a stone fell side-by-side to the ground.

As we saw in Chapter 2, an object falling freely close to the Earth's surface has an acceleration of roughly 9.81 m s^{-2}, the acceleration of free fall g.

We can find the force causing this acceleration using $F = ma$. This force is the object's **weight**. Hence the weight W of an object is given by:

weight = mass × acceleration of free fall

or

$W = mg$

Gravitational field strength

Here is another way to think about the significance of g. This quantity indicates how strong gravity is at a particular place. The Earth's gravitational field is stronger than the Moon's. On the Earth's surface, gravity gives an acceleration of free fall of about 9.81 m s^{-2}. On the Moon, gravity is weaker; it only gives an acceleration of free fall of about 1.6 m s^{-2}. So g indicates the strength of the gravitational field at a particular place:

g = gravitational field strength

and

weight = mass × gravitational field strength

(Gravitational field strength has units of N kg^{-1}. This unit is equivalent to m s^{-2}.)

QUESTION

9 Estimate the mass and weight of each of the following at the surface of the Earth:

a a kilogram of potatoes

b this book

c an average student

d a mouse

e a 40-tonne truck.

(For estimates, use $g = 10$ m s^{-2}; 1 tonne = 1000 kg.)

On the Moon

The Moon is smaller and has less mass than the Earth, and so its gravity is weaker. If you were to drop a stone on the Moon, it would have a smaller acceleration. Your hand is about 1 m above ground level; a stone takes about 0.45 s to fall through this distance on the Earth, but about 1.1 s on the surface of the Moon. The acceleration of free fall on the Moon is about one-sixth of that on the Earth:

$g_{Moon} = 1.6$ m s^{-2}

It follows that objects weigh less on the Moon than on the Earth. They are not completely weightless, because the Moon's gravity is not zero.

Mass and weight

We have now considered two related quantities, mass and weight. It is important to distinguish carefully between these (Table 3.4).

If your moon-buggy breaks down (Figure 3.5), it will be no easier to get it moving on the Moon than on the Earth. This is because its mass does not change, because it is made from just the same atoms and molecules wherever it is. From $F = ma$, it follows that if m doesn't change, you will need the same force F to start it moving.

However, your moon-buggy will be easier to lift on the Moon, because its weight will be less. From $W = mg$, since g is less on the Moon, it has a smaller weight than when on the Earth.

Quantity	Symbol	Unit	Comment
mass	m	kg	this does not vary from place to place
weight	mg	N	this a force – it depends on the strength of gravity

Table 3.4 Distinguishing between mass and weight.

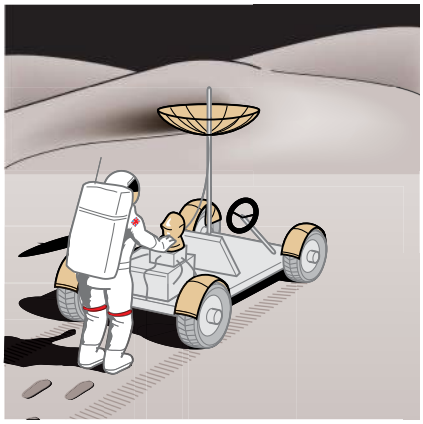


Figure 3.5 The mass of a moon-buggy is the same on the Moon as on the Earth, but its weight is smaller.

Mass and inertia

It took a long time for scientists to develop correct ideas about forces and motion. We will start by thinking about some wrong ideas, and then consider why Galileo, Newton and others decided new ideas were needed.

Observations and ideas

Here are some observations to think about:

- The large tree trunk shown in Figure 3.6 is being dragged from a forest. The elephant provides the force needed to pull it along. If the elephant stops pulling, the tree trunk will stop moving.
- A horse is pulling a cart. If the horse stops pulling, the cart soon stops.
- You are riding a bicycle. If you stop pedalling, the bicycle will come to a halt.
- You are driving along the road. You must keep your foot on the accelerator pedal, otherwise the car will not keep moving.
- You kick a football. The ball rolls along the ground and gradually stops.

In each of these cases, there is a force which makes something move – the pull of the elephant or the horse, your push on the bicycle pedals, the force of the car engine, the push of your foot. Without the force, the moving object comes to a halt. So what conclusion might we draw?

A moving object needs a force to keep it moving.

This might seem a sensible conclusion to draw, but it is wrong. We have not thought about all the forces involved. The missing force is friction.

In each example above, friction (or air resistance) makes the object slow down and stop when there is no force pushing or pulling it forwards. For example, if you stop pedalling your cycle, air resistance will slow you down. There is also friction at the axles of the wheels, and this too will slow you down. If you could lubricate your axles and cycle in a vacuum, you could travel along at a steady speed for ever, without pedalling!

In the 17th century, astronomers began to use telescopes to observe the night sky. They saw that objects such as the planets could move freely through space. They simply kept on moving, without anything providing a force to push them. Galileo came to the conclusion that this was the natural motion of objects.

- An object at rest will stay at rest, unless a force causes it to start moving.
- A moving object will continue to move at a steady speed in a straight line, unless a force acts on it.

So objects move with a constant velocity, unless a force acts on them. (Being stationary is simply a particular case of this, where the velocity is zero.) Nowadays it is much easier to appreciate this law of motion, because we have more experience of objects moving with little or no friction – roller-skates with low-friction bearings, ice skates, and spacecraft in empty space. In Galileo's day, people's everyday experience was of dragging things along the ground, or pulling things on carts with high-friction axles. Before Galileo, the orthodox scientific idea was that a force must act all the time to keep an object moving – this had been handed down from the time of the ancient Greek philosopher Aristotle. So it was a great achievement when scientists were able to develop a picture of a world without friction.



Figure 3.6 An elephant provides the force needed to drag this tree from the forest.

The idea of inertia

The tendency of a moving object to carry on moving is sometimes known as **inertia**.

- An object with a large mass is difficult to stop moving – think about catching a cricket ball, compared with a tennis ball.
- Similarly, a stationary object with a large mass is difficult to start moving – think about pushing a car to get it started.
- It is difficult to make a massive object change direction – think about the way a fully laden supermarket trolley tries to keep moving in a straight line.

All of these examples suggest another way to think of an object's mass; it is a measure of its inertia – how difficult it is to change the object's motion. Uniform motion is the natural state of motion of an object. Here, **uniform motion** means 'moving with constant velocity' or 'moving at a steady speed in a straight line'. Now we can summarise these findings as **Newton's first law of motion**.

> An object will remain at rest or in a state of uniform motion unless it is acted on by a resultant force.

In fact, this is already contained in the simple equation we have been using to calculate acceleration, $F = ma$. If no resultant force acts on an object ($F = 0$), it will not accelerate ($a = 0$). The object will either remain stationary or it will continue to travel at a constant velocity. If we rewrite the equation as $a = \frac{F}{m}$, we can see that the greater the mass m, the smaller the acceleration a produced by a force F.

QUESTIONS

10 Use the idea of inertia to explain why some large cars have power-assisted brakes.

11 A car crashes head-on into a brick wall. Use the idea of inertia to explain why the driver is more likely to come out through the windscreen if he or she is not wearing a seat belt.

Top speed

The vehicle shown in Figure 3.7 is capable of speeds as high as 760 mph, greater than the speed of sound. Its streamlined shape is designed to cut down air resistance and its jet engines provide a strong forward force to accelerate it up to top speed. All vehicles have a top speed.



Figure 3.7 The Thrust SSC rocket car broke the world land-speed record in 1997. It achieved a top speed of 763 mph (just over 340 m s^{-1}) over a distance of 1 mile (1.6 km).

But why can't they go any faster? Why can't a car driver keep pressing on the accelerator pedal, and simply go faster and faster?

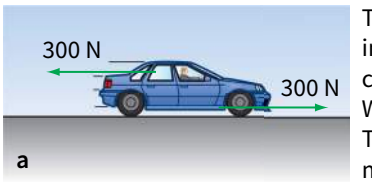
To answer this, we have to think about the two forces mentioned above: air resistance and the forward thrust (force) of the engine. The vehicle will accelerate so long as the thrust is greater than the air resistance. When the two forces are equal, the resultant force on the vehicle is zero, and the vehicle moves at a steady velocity.

Balanced and unbalanced forces

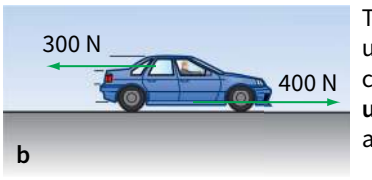
If an object has two or more forces acting on it, we have to consider whether or not they are 'balanced' (Figure 3.8). Forces on an object are balanced when the resultant force on the object is zero. The object will either remain at rest or have a constant velocity.

We can calculate the **resultant force** by adding up two (or more) forces which act in the same straight line. We must take account of the direction of each force. In the examples in Figure 3.8, forces to the right are positive and forces to the left are negative.

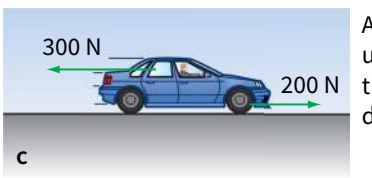
When a car travels slowly, it encounters little air resistance. However, the faster it goes, the more air it has to push out of the way each second, and so the greater



Two equal forces acting in opposite directions cancel each other out. We say they are **balance** The car will continue to move at a steady velocit in a straight line.
resultant force = 0 N



These two forces are unequal, so they do not cancel out. They are **unbalanced**. The car will accelerate.
resultant force
= 400 N – 300
= 100 N to the **right**



Again the forces are unbalanced. This time, the car will slow down o decelerate.
resultant force
= 400 N – 300 N
= 100 N to the **left**

Figure 3.8 Balanced and unbalanced forces.

the air resistance. Eventually the backward force of air resistance equals the forward force provided between the tyres and the road, and the forces on the car are balanced. It can go no faster – it has reached its top speed.

Free fall

Skydivers (Figure 3.9) are rather like cars – at first, they accelerate freely. At the start of the fall, the only force acting on the diver is his or her weight. The acceleration of the diver at the start must therefore be *g*. Then increasing air resistance opposes their fall and their acceleration decreases. Eventually they reach a maximum velocity, known as the **terminal velocity**. At the terminal velocity the air resistance is equal to the weight. The terminal velocity is approximately 120 miles per hour (about 50 m s^{-1}), but it depends on the diver's weight and orientation. Head-first is fastest.



Figure 3.9 A skydiver falling freely.

The idea of a parachute is to greatly increase the air resistance. Then terminal velocity is reduced, and the parachutist can land safely. Figure 3.10 shows how a parachutist's velocity might change during descent.

Terminal velocity depends on the weight and surface area of the object. For insects, air resistance is much greater relative to their weight than for a human being and so their terminal velocity is quite low. Insects can be swept up several kilometres into the atmosphere by rising air streams. Later, they fall back to Earth uninjured. It is said that mice can survive a fall from a high building for the same reason.

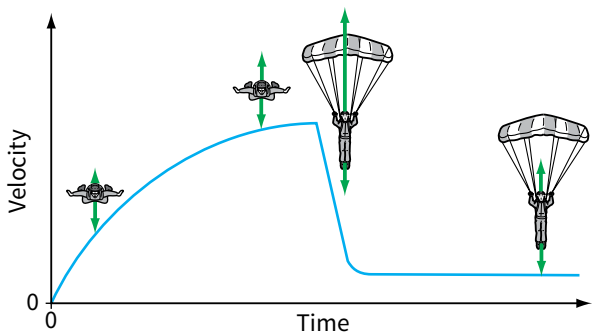


Figure 3.10 The velocity of a parachutist varies during a descent. The force arrows show weight (downwards) and air resistance (upwards).

Moving through fluids

Air resistance is just one example of the resistive or **viscous forces** which objects experience when they move through a fluid – a liquid or a gas. If you have ever run down the beach and into the sea, or tried to wade quickly through the water of a swimming pool, you will have experienced the force of **drag**. The deeper the water gets, the more it resists your movement and the harder you have to work to make progress through it. In deep water, it is easier to swim than to wade.

You can observe the effect of drag on a falling object if you drop a key or a coin into the deep end of a swimming pool. For the first few centimetres, it speeds up, but for the remainder of its fall, it has a steady speed. (If it fell through the same distance in air, it would accelerate all the way.) The drag of water means that the falling object reaches its terminal velocity very soon after it is released. Compare this with a skydiver, who has to fall hundreds of metres before reaching terminal velocity.

Moving through air

We rarely experience drag in air. This is because air is much less dense than water; its density is roughly $\frac{1}{800}$ that of water. At typical walking speed, we do not notice the effects of drag. However, if you want to move faster, they can be important. Racing cyclists, like the one shown in Figure 3.11, wear tight-fitting clothing and streamlined



Figure 3.11 A racing cyclist adopts a posture which helps to reduce drag. Clothing, helmet and even the cycle itself are designed to allow them to go as fast as possible.

helmets. Other athletes may take advantage of the drag of air. The runner in Figure 3.12 is undergoing resistance training. The parachute provides a backward force against which his muscles must work. This should help to develop his muscles.

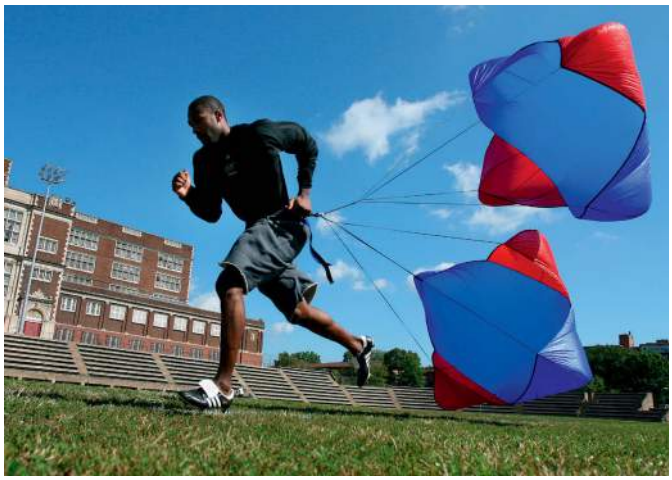


Figure 3.12 A runner making use of air resistance to build up his muscles.

QUESTIONS

12 If you drop a large stone and a small stone from the top of a tall building, which one will reach the ground first? Explain your answer.

13 In a race, downhill skiers want to travel as quickly as possible. They are always looking for ways to increase their top speed. Explain how they might do this. Think about:

- **a** their skis
- **b** their clothing
- **c** their muscles
- **d** the slope.

14 Skydivers jump from a plane at intervals of a few seconds. If two divers wish to join up as they fall, the second must catch up with the first.

- **a** If one diver is more massive than the other, which should jump first? Use the idea of forces and terminal velocity to explain your answer.
- **b** If both divers are equally massive, suggest what the second might do to catch up with the first.

WORKED EXAMPLES

5 A car of mass 500 kg is travelling along a flat road. The forward force provided between the car tyres and the road is 300 N and the air resistance is 200 N. Calculate the acceleration of the car.

Step 1 Start by drawing a diagram of the car, showing the forces mentioned in the question (Figure 3.13). Calculate the resultant force on the car; the force to the right is taken as positive:

resultant force = 300 − 200 = 100 N

Step 2 Now use $F = ma$ to calculate the car's acceleration:

$$a = \frac{F}{m} = \frac{100}{500} = 0.20\ \text{m s}^{-2}$$

So the car's acceleration is $0.20\ \text{m s}^{-2}$.

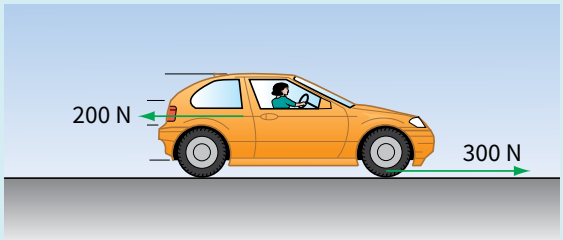


Figure 3.13 The forces on an accelerating car.

6 The maximum forward force a car can provide is 500 N. The air resistance F which the car experiences depends on its speed according to $F = 0.2v^2$, where v is the speed in m s^{-1}. Determine the top speed of the car.

Step 1 From the equation $F = 0.2v^2$, you can see that the air resistance increases as the car goes faster. Top speed is reached when the forward force equals the air resistance. So, at top speed:

$500 = 0.2v^2$

Step 2 Rearranging gives:

$$v^2 = \frac{500}{0.2} = 2500$$

$$v = 50\ \text{m s}^{-1}$$

So the car's top speed is $50\ \text{m s}^{-1}$ (this is about $180\ \text{km h}^{-1}$).

Identifying forces

It is important to be able to identify the forces which act on an object. When we know what forces are acting, we can predict how it will move. Figure 3.14 shows some important forces, how they arise, and how we represent them in diagrams.

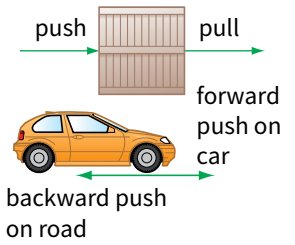
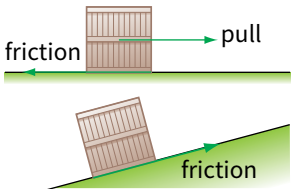
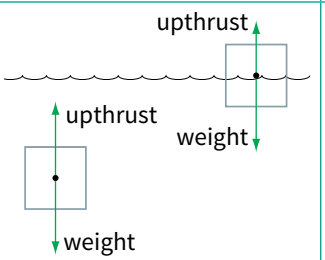
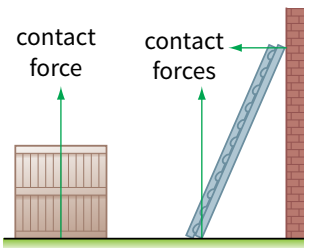
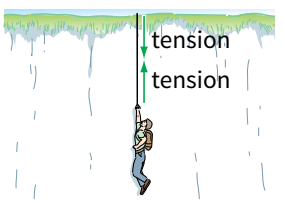
Diagram	Force	Important situations
	**Pushes and pulls.** You can make an object accelerate by pushing and pulling it. Your force is shown by an arrow pushing (or pulling) the object. The engine of a car provides a force to push backwards on the road. Frictional forces from the road on the tyre push the car forwards.	■ pushing and pulling ■ lifting ■ force of car engine ■ attraction and repulsion by magnets and by electric charges
	**Weight.** This is the force of gravity acting on the object. It is usually shown by an arrow pointing vertically downwards from the object's centre of gravity.	■ any object in a gravitational field ■ less on the Moon
	**Friction.** This is the force which arises when two surfaces rub over one another. If an object is sliding along the ground, friction acts in the opposite direction to its motion. If an object is stationary, but tending to slide – perhaps because it is on a slope – the force of friction acts up the slope to stop it from sliding down. Friction always acts along a surface, never at an angle to it.	■ pulling an object along the ground ■ vehicles cornering or skidding ■ sliding down a slope
	**Drag.** This force is similar to friction. When an object moves through air, there is friction between it and the air. Also, the object has to push aside the air as it moves along. Together, these effects make up drag. Similarly, when an object moves through a liquid, it experiences a drag force. Drag acts to oppose the motion of an object; it acts in the opposite direction to the object's velocity. It can be reduced by giving the object a streamlined shape.	■ vehicles moving ■ aircraft flying ■ parachuting ■ objects falling through air or water ■ ships sailing
	**Upthrust.** Any object placed in a fluid such as water or air experiences an upwards force. This is what makes it possible for something to float in water. Upthrust arises from the pressure which a fluid exerts on an object. The deeper you go, the greater the pressure. So there is more pressure on the lower surface of an object than on the upper surface, and this tends to push it upwards. If upthrust is greater than the object's weight, it will float up to the surface.	■ boats and icebergs floating ■ people swimming ■ divers surfacing ■ a hot air balloon rising
	**Contact force.** When you stand on the floor or sit on a chair, there is usually a force which pushes up against your weight, and which supports you so that you do not fall down. The contact force is sometimes known as the normal reaction of the floor or chair. (In this context, normal means 'perpendicular'.) The contact force always acts at right angles to the surface which produces it. The floor pushes straight upwards; if you lean against a wall, it pushes back against you horizontally.	■ standing on the ground ■ one object sitting on top of another ■ leaning against a wall ■ one object bouncing off another
	**Tension.** This is the force in a rope or string when it is stretched. If you pull on the ends of a string, it tends to stretch. The tension in the string pulls back against you. It tries to shorten the string. Tension can also act in springs. If you stretch a spring, the tension pulls back to try to shorten the spring. If you squash (compress) the spring, the tension acts to expand the spring.	■ pulling with a rope ■ squashing or stretching a spring

Figure 3.14 Some important forces.

Contact forces and upthrust

We will now think about the forces which act when two objects are in contact with each other. When two objects touch each other, each exerts a force on the other. These are called **contact forces**. For example, when you stand on the floor (Figure 3.15), your feet push downwards on the floor and the floor pushes back upwards on your feet. This is a vital force – the upward push of the floor prevents you from falling downwards under the pull of your weight.

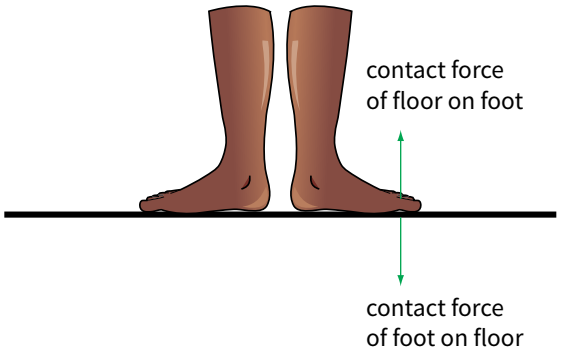


Figure 3.15 Equal and opposite contact forces act when you stand on the floor.

Where do these contact forces come from? When you stand on the floor, the floor becomes slightly compressed. Its atoms are pushed slightly closer together, and the interatomic forces push back against the compressing force. At the same time, the atoms in your feet are also pushed together so that they push back in the opposite direction. (It is hard to see the compression of the floor when you stand on it, but if you stand on a soft material such as foam rubber or a mattress you will be able to see the compression clearly.)

You can see from Figure 3.15 that the two contact forces act in opposite directions. They are also equal in magnitude. As we will see shortly, this is a consequence of Newton's third law of motion.

When an object is immersed in a fluid (a liquid or a gas), it experiences an upward force called **upthrust**. It is the upthrust of water which keeps a boat floating (Figure 3.16) and the upthrust of air which lifts a hot air balloon upwards.

The upthrust of water on a boat can be thought of as the contact force of the water on the boat. It is caused by the pressure of the water pushing upwards on the boat. Pressure arises from the motion of the water molecules colliding with the boat and the net effect of all these collisions is an upward force.

An object in air, such as a ball, has a very small upthrust acting on it, because the density of the air around it is low. Molecules hit the top surface of the ball pushing down, but only a few more molecules push upwards on the bottom of the ball, so the resultant force upwards, or the upthrust is low. If the ball is falling, air resistance is greater than this small upthrust but both these forces are acting upwards on the ball.

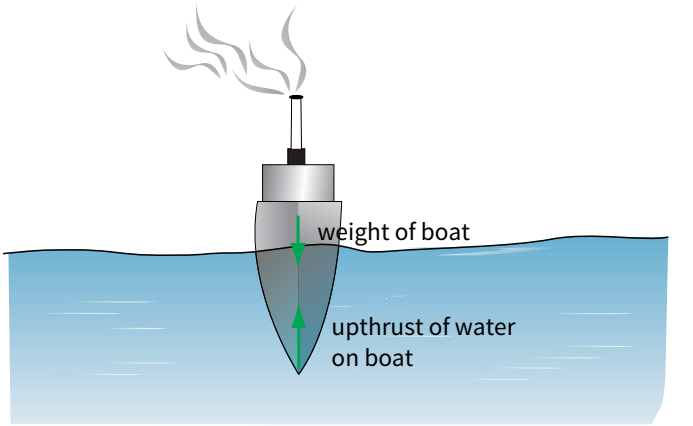


Figure 3.16 Without sufficient upthrust from the water, the boat would sink.

QUESTIONS

15 Name these forces:

- **a** the upward push of water on a submerged object
- **b** the force which wears away two surfaces as they move over one another
- **c** the force which pulled the apple off Isaac Newton's tree
- **d** the force which stops you falling through the floor
- **e** the force in a string which is holding up an apple
- **f** the force which makes it difficult to run through shallow water.

16 Draw a diagram to show the forces which act on a car as it travels along a level road at its top speed.

17 Imagine throwing a shuttlecock straight up in the air. Air resistance is more important for shuttlecocks than for a tennis ball. Air resistance always acts in the opposite direction to the velocity of an object.

Draw diagrams to show the two forces, weight and air resistance, acting on the shuttlecock:

- **a** as it moves upwards
- **b** as it falls back downwards.

Newton's third law of motion

For completeness, we should now consider **Newton's third law of motion**. (There is more about this in Chapter 6.)

When two objects interact, each exerts a force on the other. Newton's third law says that these forces are equal and opposite to each other:

> When two bodies interact, the forces they exert on each other are equal in magnitude and opposite in direction.

(These two forces are sometimes described as **action** and **reaction**, but this is misleading as it sounds as though one force arises as a consequence of the other. In fact, the two forces appear at the same time and we can't say that one caused the other.)

The two forces which make up a 'Newton's third law pair' have the following characteristics:

- They act on **different** objects.
- They are equal in magnitude.
- They are opposite in direction.
- They are forces **of the same type**.

What does it mean to say that the forces are 'of the same type'? We need to think about the type of interaction which causes the forces to appear.

- Two objects may attract each other because of the gravity of their masses – these are gravitational forces.
- Two objects may attract or repel because of their electrical charges – electrical forces.
- Two objects may touch – contact forces.
- Two objects may be attached by a string and pull on each other – tension forces.
- Two objects may attract or repel because of their magnetic fields – magnetic forces.

Figure 3.17 shows a person standing on the Earth's surface. The two gravitational forces are a Newton's third law pair, as are the two contact forces. Don't be misled into thinking that the person's weight and the contact force of the floor are a Newton's third law pair. Although they are 'equal and opposite', they do not act on different objects and they are not of the same type.

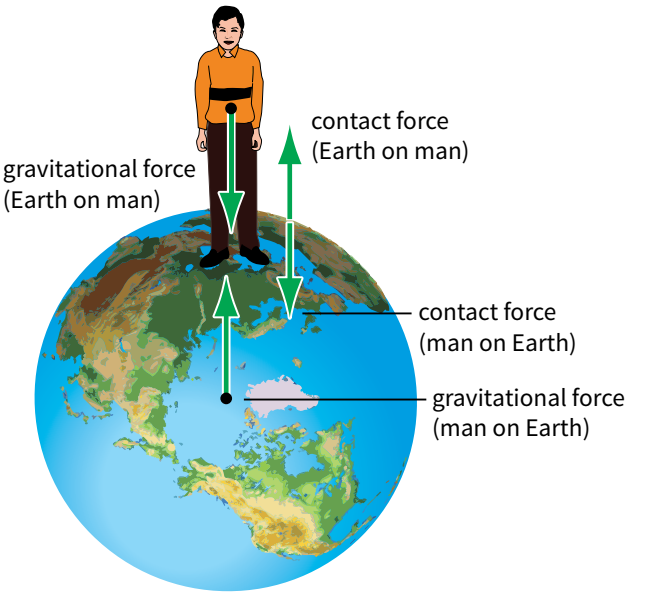


Figure 3.17 For each of the forces that the Earth exerts on you, an equal and opposite force acts on the Earth.

QUESTION

18 Describe one 'Newton's third law pair' of forces involved in the following situations. In each case, state the object that each force acts on, the type of force and the direction of the force.

a You step on someone's toe.

b A car hits a brick wall and comes to rest.

c A car slows down by applying the brakes.

d You throw a ball upwards into the air.

Summary

- An object will remain at rest or in a state of uniform motion unless it is acted on by an external force. This is Newton's first law of motion.
- For a body of constant mass, the acceleration is directly proportional to the resultant force applied to it. Resultant force F, mass m and acceleration a are related by the equation:

  resultant force = mass × acceleration

  $F = ma$

  This is a form of Newton's second law of motion.
- When two bodies interact, the forces they exert on each other are equal in magnitude and opposite in direction.
  This is Newton's third law of motion.
- The acceleration produced by a force is in the same direction as the force. Where there are two or more forces, we must determine the resultant force.
- A newton (N) is the force required to give a mass of 1 kg an acceleration of 1 m s^{-2} in the direction of the force.
- The greater the mass of an object, the more it resists changes in its motion. Mass is a measure of the object's inertia.
- The weight of an object is a result of the pull of gravity on it:

  weight = mass × acceleration of free fall ($W = mg$)

  weight = mass × gravitational field strength
- An object falling freely under gravity has a constant acceleration provided the gravitational field strength is constant. However, fluid resistance (such as air resistance) reduces its acceleration. Terminal velocity is reached when the fluid resistance is equal to the weight of the object.

End-of-chapter questions

1 When a golfer hits a ball his club is in contact with the ball for about 0.0005 s and the ball leaves the club with a speed of 70 m s^{-1}. The mass of the ball is 46 g.

a Determine the mean accelerating force. [4]

b What mass, resting on the ball, would exert the same force as in **a**? [2]

2 The mass of a spacecraft is 70 kg. As the spacecraft takes off from the Moon, the upwards force on the spacecraft caused by the engines is 500 N. The gravitational field strength on the Moon is 1.6 N kg^{-1}.

Determine:

a the weight of the spacecraft on the Moon [2]

b the resultant force on the spacecraft [2]

c the acceleration of the spacecraft. [2]

3 A metal ball is dropped into a tall cylinder of oil. The ball initially accelerates but soon reaches a terminal velocity.

a By considering the forces on the metal ball bearing, explain why it first accelerates but then reaches terminal velocity. [3]

b Describe how you would show that the metal ball reaches terminal velocity. [3]

4 Determine the speed in m s^{-1} of an object that travels:

a 3 µm in 5 ms [2]

b 6 km in 3 Ms [2]

c 8 pm in 4 ns. [2]

5 Figure 3.18 shows a man who is just supporting the weight of a box. Two of the forces acting are shown in the diagram. According to Newton's third law, each of these forces is paired with **another** force.

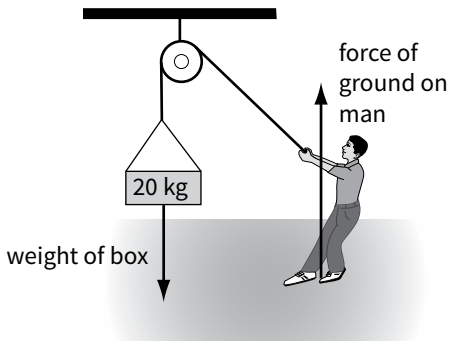


Figure 3.18 For End-of-chapter Question 5.

For **a** the weight of the box and **b** the force of the ground on the man, state:

i the body that the other force acts upon [2]

ii the direction of the other force [2]

iii the type of force involved. [2]

6 A car starts to move along a straight, level road. For the first 10 s, the driver maintains a constant acceleration of 1.5 m s^{-2}. The mass of the car is 1.1 × 10^3 kg.

a Calculate the driving force provided by the wheels, when:

i the force opposing motion is negligible [1]

ii the total force opposing the motion of the car is 600 N. [1]

b Calculate the distance travelled by the car in the first 10 s. [2]

7 Figure 3.19 shows the speed–time graphs for two falling balls.

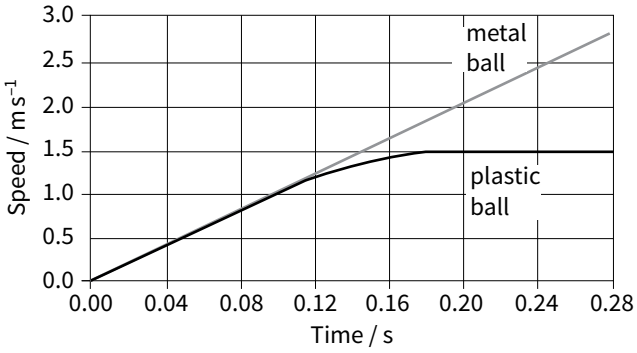


Figure 3.19 For End-of-chapter Question 7.

a Determine the terminal velocity of the plastic ball. [1]

b Both balls are of the same size and shape but the metal ball has a greater mass. Explain, in terms of Newton's laws of motion and the forces involved, why the plastic ball reaches a constant velocity but the metal ball does not. [3]

c Explain why both balls have the same initial acceleration. [2]

8 A car of mass 1200 kg accelerates from rest to a speed of 8.0 m s^{-1} in a time of 2.0 s.

a Calculate the forward driving force acting on the car while it is accelerating. Assume that, at low speeds, all frictional forces are negligible. [2]

b At high speeds the resistive frictional force F produced by air on a body moving with velocity v is given by the equation $F = bv^2$, where b is a constant.

i Derive the base units of force in the SI system. [1]

ii Determine the base units of b in the SI system. [1]

iii The car continues with the same forward driving force and accelerates until it reaches a top speed of 50 m s^{-1}. At this speed the resistive force is given by the equation $F = bv^2$. Determine the value of b for the car. [2]

iv Sketch a graph showing how the value of F varies with v over the range 0 to 50 m s^{-1}. Use your graph to describe what happens to the acceleration of the car during this time. [2]

9 **a** Explain what is meant by the **mass** of a body and the **weight** of a body. [3]

b State and explain **one** situation in which the weight of a body changes while its mass remains constant. [2]

c State the difference between the base units of mass and weight in the SI system. [2]

10 **a** State Newton's second law of motion. [2]

b When you jump from a wall on to the ground, it is advisable to bend your knees on landing.

i State how bending your knees affects the time it takes to stop when hitting the ground. [1]

ii Using Newton's second law of motion, explain why it is sensible to bend your knees. [2]

Chapter 4: Forces – vectors and moments

Learning outcomes

You should be able to:

- add two or more coplanar forces
- resolve a force into perpendicular components
- define and apply the moment of a force and the torque of a couple
- apply the principle of moments
- state the conditions for a body to be in equilibrium

Sailing ahead

Force is a vector quantity. Sailors know a lot about the vector nature of forces. For example, they can sail 'into the wind'. The sails of a yacht can be angled to provide a component of force in the forward direction and the boat can then sail at almost 45° to the wind. The boat tends to 'heel over' and the crew sit on the side of the boat to provide a turning effect in the opposite direction (Figure 4.1).



Figure 4.1 Sailing into the wind.

Combining forces

You should recall that a vector quantity has both magnitude and direction. An object may have two or more forces acting on it and, since these are vectors, we must use vector addition (Chapter 1) to find their combined effect (their resultant).

There are several forces acting on the car (Figure 4.2) as it struggles up the steep hill. They are:

- its weight W (= mg)
- the contact force N of the road (its normal reaction)
- air resistance D
- the forward force F caused by friction between the car tyres and the road.

If we knew the magnitude and direction of each of these forces, we could work out their combined effect on the car. Will it accelerate up the hill? Or will it slide backwards down the hill?

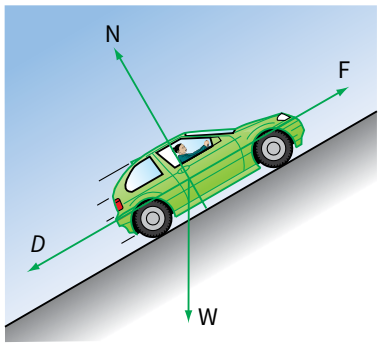


Figure 4.2 Four forces act on this car as it moves uphill.

The combined effect of several forces is known as the **resultant force**. To see how to work out the resultant of two or more forces, we will start with a relatively simple example.

Two forces in a straight line

We saw some examples in Chapter 3 of two forces acting in a straight line. For example, a falling tennis ball may be acted on by two forces: its weight mg, downwards, and air resistance D, upwards (Figure 4.3). The resultant force is then:

$$\text{resultant force} = mg - D = 1.0 - 0.2 = 0.8\,\text{N}$$

When adding two or more forces which act in a straight line, we have to take account of their directions. A force may be positive or negative; we adopt a **sign convention** to help us decide which is which.

If you apply a sign convention correctly, the sign of your final answer will tell you the direction of the resultant force (and hence acceleration).

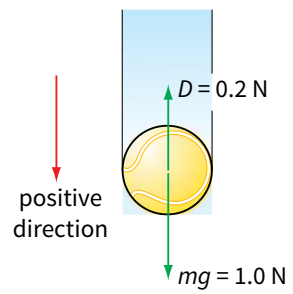


Figure 4.3 Two forces on a falling tennis ball.

Two forces at right angles

Figure 4.4 shows a shuttlecock falling on a windy day. There are two forces acting on the shuttlecock: its weight vertically downwards, and the horizontal push of the wind. (It helps if you draw the force arrows of different lengths, to show which force is greater.) We must add these two forces together to find the resultant force acting on the shuttlecock.

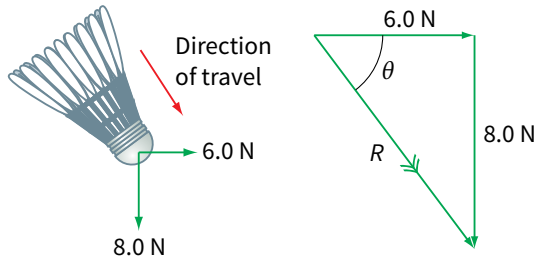


Figure 4.4 Two forces act on this shuttlecock as it travels through the air; the vector triangle shows how to find the resultant force.

We add the forces by drawing two arrows, end-to-end, as shown on the right of Figure 4.4.

- First, a horizontal arrow is drawn to represent the 6.0 N push of the wind.
- Next, starting from the end of this arrow, we draw a second arrow, downwards, representing the weight of 8.0 N.
- Now we draw a line from the start of the first arrow to the end of the second arrow. This arrow represents the resultant force *R*, in both magnitude and direction.

The arrows are added by drawing them end-to-end; the end of the first arrow is the start of the second arrow. Now we can find the resultant force either by scale drawing or by calculation. In this case, we have a 3–4–5 right-angled triangle, so calculation is simple:

$$R^2 = 6.0^2 + 8.0^2 = 36 + 64 = 100$$

$$R = 10\,\text{N}$$

$$\tan\theta = \frac{\text{opp}}{\text{adj}} = \frac{8.0}{6.0} = \frac{4}{3}$$

$$\theta = \tan^{-1}\frac{4}{3} \approx 53°$$

So the resultant force is 10 N, at an angle of 53° below the horizontal. This is a reasonable answer; the weight is pulling the shuttlecock downwards and the wind is pushing it to the right. The angle is greater than 45° because the downward force is greater than the horizontal force.

If you draw a scale drawing be careful to:

- state the scale used
- draw a large diagram to reduce the uncertainty.

Three or more forces

The spider shown in Figure 4.5 is hanging by a thread. It is blown sideways by the wind. The diagram shows the three forces acting on it:

- weight acting downwards
- the tension in the thread
- the push of the wind.

The diagram also shows how these can be added together. In this case, we arrive at an interesting result. Arrows are drawn to represent each of the three forces, end-to-end. The end of the third arrow coincides with the start of the first arrow, so the three arrows form a closed triangle. This tells us that the resultant force *R* on the spider is zero, that is, $R = 0$. The closed triangle in Figure 4.5 is known as a **triangle of forces**.

So there is no resultant force. The forces on the spider balance each other out, and we say that the spider is in **equilibrium**. If the wind blew a little harder, there would be an unbalanced force on the spider, and it would move off to the right.

We can use this idea in two ways:

- If we work out the resultant force on an object and find that it is zero, this tells us that the object is in equilibrium.
- If we know that an object is in equilibrium, we know that the forces on it must add up to zero. We can use this to work out the values of one or more unknown forces.

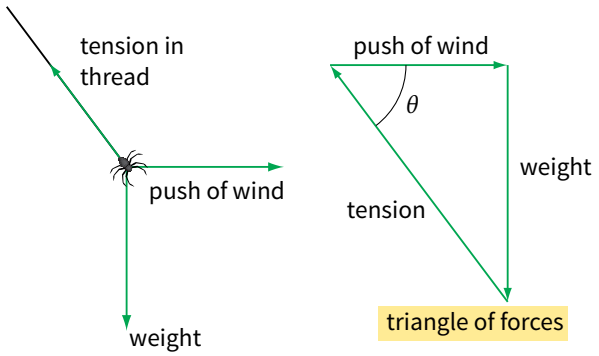


Figure 4.5 Blowing in the wind – this spider is hanging in equilibrium.

QUESTIONS

1 A parachutist weighs 1000 N. When she opens her parachute, it pulls upwards on her with a force of 2000 N.
   - **a** Draw a diagram to show the forces acting on the parachutist.
   - **b** Calculate the resultant force acting on her.
   - **c** What effect will this force have on her?

2 The ship shown in Figure 4.6 is travelling at a constant velocity.
   - **a** Is the ship in equilibrium (in other words, is the resultant force on the ship equal to zero)? How do you know?
   - **b** What is the upthrust U of the water?
   - **c** What is the drag D of the water?

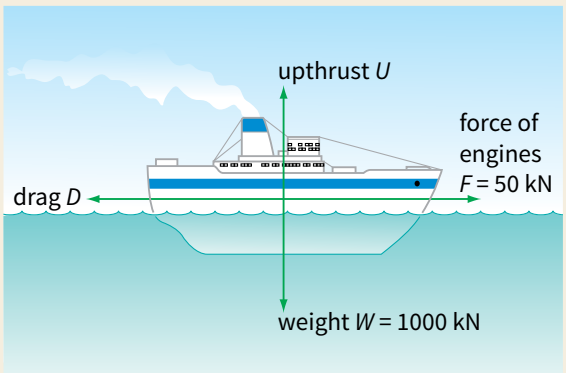


Figure 4.6 For Question 2. The force D is the frictional drag of the water on the boat. Like air resistance, drag is always in the opposite direction to the object's motion.

3 A stone is dropped into a fast-flowing stream. It does not fall vertically, because of the sideways push of the water (Figure 4.7).
   - **a** Calculate the resultant force on the stone.
   - **b** Is the stone in equilibrium?

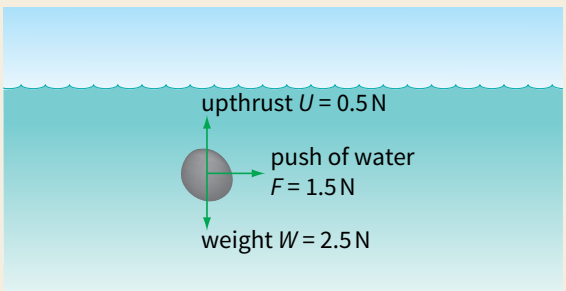


Figure 4.7 For Question 3.

Components of vectors

Look back to Figure 4.5. The spider is in equilibrium, even though three forces are acting on it. We can think of the tension in the thread as having two effects:

- it is pulling upwards, to counteract the downward effect of gravity
- it is pulling to the left, to counteract the effect of the wind.

We can say that this force has two effects or **components**: an upwards (vertical) component and a sideways (horizontal) component. It is often useful to split up a vector quantity into components like this, just as we did with velocity in Chapter 2. The components are in two directions at right angles to each other, often horizontal and vertical. The process is called **resolving** the vector. Then we can think about the effects of each component separately; we say that the perpendicular components are **independent** of one another. Because the two components are at 90° to each other, a change in one will have no effect on the other. Figure 4.8 shows how to resolve a force F into its horizontal and vertical components. These are:

horizontal component of F, $F_x = F\cos\theta$

vertical component of F, $F_y = F\sin\theta$

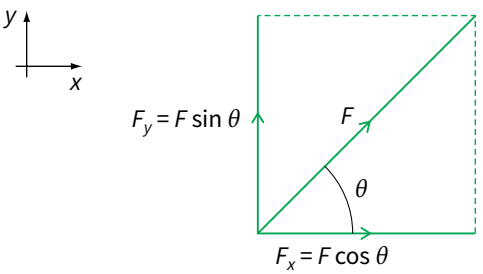


Figure 4.8 Resolving a vector into two components at right angles.

Making use of components

When the trolley shown in Figure 4.9 is released, it accelerates down the ramp. This happens because of the weight of the trolley. The weight acts vertically downwards, although this by itself does not determine the resulting motion. However, the weight has a component which acts down the slope. By calculating the component of the trolley's weight down the slope, we can determine its acceleration.

Figure 4.10 shows the forces acting on the trolley. To simplify the situation, we will assume there is no friction. The forces are:

- W, the weight of the trolley, which acts vertically downwards
- N, the contact force of the ramp, which acts at right angles to the ramp.



Figure 4.9 These students are investigating the acceleration of a trolley down a sloping ramp.

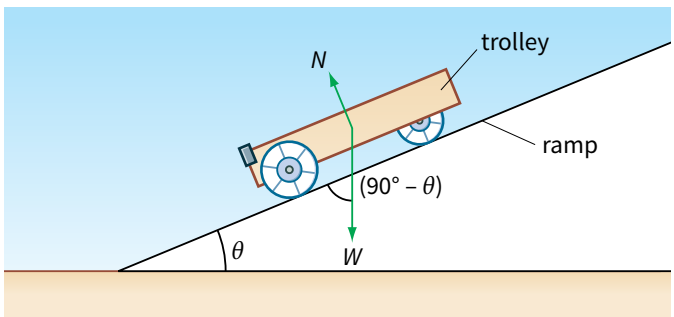


Figure 4.10 A force diagram for a trolley on a ramp.

You can see at once from the diagram that the forces cannot be balanced, since they do not act in the same straight line.

To find the component of W down the slope, we need to know the angle between W and the slope. The slope makes an angle θ with the horizontal, and from the diagram we can see that the angle between the weight and the ramp is $(90° - \theta)$. Using the rule for calculating the component of a vector given above, we have:

$$\text{component of } W \text{ down the slope} = W\cos(90° - \theta)$$
$$= W\sin\theta$$

(It is helpful to recall that $\cos(90° - \theta) = \sin\theta$; you can see this from Figure 4.10.)

Does the contact force N help to accelerate the trolley down the ramp? To answer this, we must calculate its component down the slope. The angle between N and the slope is 90°. So:

$$\text{component of } N \text{ down the slope} = N\cos 90° = 0$$

The cosine of 90° is zero, and so N has no component down the slope. This shows why it is useful to think in terms of the components of forces; we don't know the value of N, but, since it has no effect down the slope, we can ignore it.

(There's no surprise about this result. The trolley runs down the slope because of the influence of its weight, not because it is pushed by the contact force N.)

Changing the slope

If the students in Figure 4.9 increase the slope of their ramp, the trolley will move down the ramp with greater acceleration. They have increased θ, and so the component of W down the slope will have increased.

Now we can work out the trolley's acceleration. If the trolley's mass is m, its weight is mg. So the force F making it accelerate down the slope is:

$$F = mg\sin\theta$$

Since from Newton's second law for constant mass we have $a = \frac{F}{m}$, the trolley's acceleration a is given by:

$$a = \frac{mg\sin\theta}{m} = g\sin\theta$$

We could have arrived at this result simply by saying that the trolley's acceleration would be the component of g down the slope (Figure 4.11). The steeper the slope, the greater the value of $\sin\theta$, and hence the greater the trolley's acceleration.

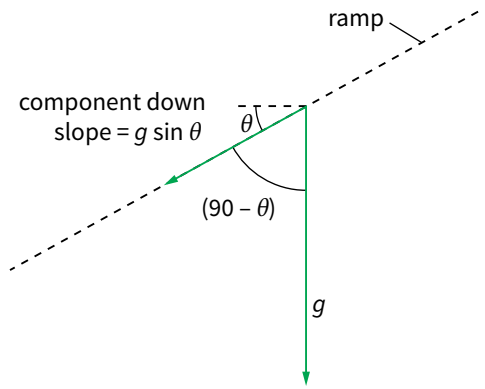


Figure 4.11 Resolving g down the ramp.

QUESTIONS

4 The person in Figure 4.12 is pulling a large box using a rope. Use the idea of components of a force to explain why they are more likely to get the box to move if the rope is horizontal (as in **a**) than if it is sloping upwards (as in **b**).

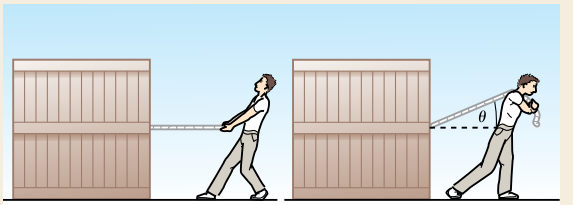


Figure 4.12 Why is it easier to move the box with the rope horizontal? See Question 4.

5 A crate is sliding down a slope. The weight of the crate is 500 N. The slope makes an angle of 30° with the horizontal.

- **a** Draw a diagram to show the situation. Include arrows to represent the weight of the crate and the contact force of the slope acting on the crate.
- **b** Calculate the component of the weight down the slope.
- **c** Explain why the contact force of the slope has no component down the slope.
- **d** What third force might act to oppose the motion? In which direction would it act?

Solving problems by resolving forces

A force can be resolved into two components at right angles to each other; these can then be treated independently of one another. This idea can be used to solve problems, as illustrated in Worked example 1.

QUESTION

6 A child of mass 40 kg is on a water slide. The slide slopes down at 25° to the horizontal. The acceleration of free fall is 9.81 m s^{-2}. Calculate the child's acceleration down the slope:

- **a** when there is no friction and the only force acting on the child is his weight
- **b** if a frictional force of 80 N acts up the slope.

WORKED EXAMPLE

1 A boy of mass 40 kg is on a waterslide which slopes at 30° to the horizontal. The frictional force up the slope is 120 N. Calculate the boy's acceleration down the slope. Take the acceleration of free fall g to be 9.81 m s^{-2}.

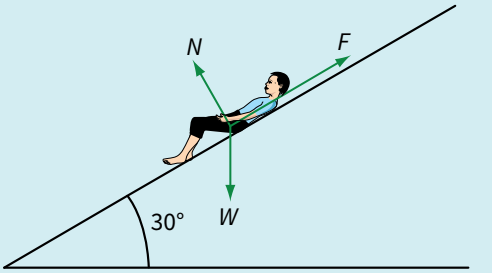


Figure 4.13 For Worked example 1.

Step 1 Draw a labelled diagram showing all the forces acting on the object of interest (Figure 4.13). This is known as a **free-body force diagram**. The forces are:

the boy's weight $W = 40 \times 9.81 = 392$ N

the frictional force up the slope $F = 120$ N

the contact force N at 90° to the slope.

Step 2 We are trying to find the resultant force on the boy which makes him accelerate down the slope. We resolve the forces down the slope, i.e. we find their components in that direction.

component of W down the slope $= 392 \times \cos 60°$
$= 196$ N

component of F down the slope $= -120$ N
(negative because F is directed up the slope)

component of N down the slope $= 0$
(because it is at 90° to the slope)

It is convenient that N has no component down the slope, since we do not know the value of N.

Step 3 Calculate the resultant force on the boy:

resultant force $= 196 - 120 = 76$ N

Step 4 Calculate his acceleration:

$$\text{acceleration} = \frac{\text{resultant force}}{\text{mass}} = \frac{76}{40} = 1.9 \text{ m s}^{-2}$$

So the boy's acceleration down the slope is 1.9 m s^{-2}. We could have arrived at the same result by resolving vertically and horizontally, but that would have led to two simultaneous equations from which we would have had to eliminate the unknown force N. It often helps to resolve forces at 90° to an unknown force.

Centre of gravity

We have weight because of the force of gravity of the Earth on us. Each part of our body – arms, legs, head, for example – experiences a force, caused by the force of gravity. However, it is much simpler to picture the overall effect of gravity as acting at a single point. This is our **centre of gravity**.

The centre of gravity of an object is defined as the point where all the weight of the object may be considered to act.

For a person standing upright, the centre of gravity is roughly in the middle of the body, behind the navel. For a sphere, it is at the centre. It is much easier to solve problems if we simply indicate an object's weight by a single force acting at the centre of gravity, rather than a large number of forces acting on each part of the object. Figure 4.14 illustrates this point. The athlete performs a complicated manoeuvre. However, we can see that his centre of gravity follows a smooth, parabolic path through the air, just like the paths of projectiles we discussed in Chapter 2.



Figure 4.14 The dots indicate the athlete's centre of gravity, which follows a smooth trajectory through the air. With his body curved like this, the athlete's centre of gravity is actually outside his body, just below the small of his back. At no time is the whole of his body above the bar.

BOX 4.1: Finding the centre of gravity

The centre of gravity of a thin sheet, or lamina, of cardboard or metal can be found by suspending it freely from two or three points (Figure 4.15).

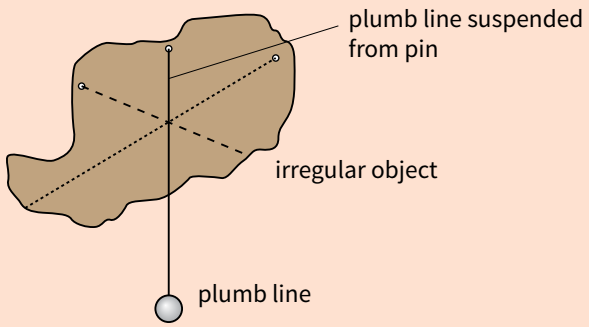


Figure 4.15 The centre of gravity is located at the intersection of the lines.

Small holes are made round the edge of the irregularly shaped object. A pin is put through one of the holes and held firmly in a clamp and stand so the object can swing freely. A length of string is attached to the pin. The other end of the string has a heavy mass attached to it. This arrangement is called a **plumb line**.

The object will stop swinging when its centre of gravity is vertically below the point of suspension. A line is drawn on the object along the vertical string of the plumb line. The centre of gravity must lie on this line. To find the position of the centre of gravity, the process is repeated with the object suspended from different holes. The centre of gravity will be at the point of intersection of the lines drawn on the object.

The turning effect of a force

Forces can make things accelerate. They can do something else as well: they can make an object turn round. We say that they can have a **turning effect**. Figure 4.16 shows how to use a spanner to turn a nut.

To maximise the turning effect of his force, the operator pulls close to the end of the spanner, as far as possible from the pivot (the centre of the nut) and at 90° to the spanner.



Figure 4.16 A mechanic turns a nut.

Moment of a force

The quantity which tells us about the turning effect of a force is its **moment**. The moment of a force depends on two quantities:

- the magnitude of the force (the bigger the force, the greater its moment)
- the perpendicular distance of the force from the pivot (the further the force acts from the pivot, the greater its moment).

The moment of a force is defined as follows:

> The moment of a force = force × perpendicular distance of the pivot from the line of action of the force.

Figure 4.17a shows these quantities. The force F_1 is pushing down on the lever, at a perpendicular distance x_1 from the pivot. The moment of the force F_1 about the pivot is then given by:

$$\text{moment} = \text{force} \times \text{distance from pivot}$$

$$= F_1 \times x_1$$

The unit of moment is the newton metre (N m). This is a unit which does not have a special name. You can also determine the moment of a force in N cm.

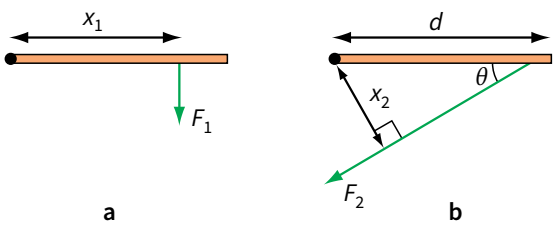


Figure 4.17 The quantities involved in calculating the moment of a force.

Figure 4.17b shows a slightly more complicated situation. F_2 is pushing at an angle θ to the lever, rather than at 90°. This makes it have less turning effect. There are two ways to calculate the moment of the force.

Method 1

Draw a perpendicular line from the pivot to the line of the force. Find the distance x_2. Calculate the moment of the force, $F_2 \times x_2$. From the right-angled triangle, we can see that:

$$x_2 = d \sin\theta$$

Hence:

$$\text{moment of force} = F_2 \times d \sin\theta = F_2 d \sin\theta$$

Method 2

Calculate the component of F_2 which is at 90° to the lever. This is $F_2 \sin\theta$. Multiply this by d.

$$\text{moment} = F_2 \sin\theta \times d$$

We get the same result as Method 1:

$$\text{moment of force} = F_2 d \sin\theta$$

Note that any force (such as the component $F_2 \cos\theta$) which passes through the pivot has no turning effect, because the distance from the pivot to the line of the force is zero.

Note also that we can calculate the moment of a force about any point, not just the pivot. However, in solving problems, it is often most convenient to take moments about the pivot as there is often an unknown force acting through the pivot (its contact force on the object).

Balanced or unbalanced?

We can use the idea of the moment of a force to solve two sorts of problem:

- We can check whether an object will remain balanced or start to rotate.
- We can calculate an unknown force or distance if we know that an object is balanced.

We can use the **principle of moments** to solve problems. The principle of moments states that:

> For any object that is in **equilibrium**, the sum of the clockwise moments about any point provided by the forces acting on the object equals the sum of the anticlockwise moments about that same point.

Note that, for an object to be in equilibrium, we also require that no resultant force acts on it. The Worked examples that follow illustrate how we can use these ideas to determine unknown forces.

WORKED EXAMPLES

2 Is the see-saw shown in Figure 4.18 in equilibrium (balanced), or will it start to rotate?

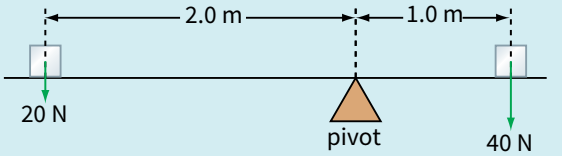


Figure 4.18 Will these forces make the see-saw rotate, or are their moments balanced?

The see-saw will remain balanced, because the 20 N force is twice as far from the pivot as the 40 N force.

To prove this, we need to think about each force individually. Which direction is each force trying to turn the see-saw, clockwise or anticlockwise? The 20 N force is tending to turn the see-saw anticlockwise, while the 40 N force is tending to turn it clockwise.

Step 1 Determine the anticlockwise moment:

moment of anticlockwise force = 20 × 2.0 = 40 N m

Step 2 Determine the clockwise moment:

moment of clockwise force = 40 × 1.0 = 40 N m

Step 3 We can see that:

clockwise moment = anticlockwise moment

So the see-saw is balanced and therefore does not rotate. The see-saw is in equilibrium.

3 The beam shown in Figure 4.19 is in equilibrium. Determine the force *X*.

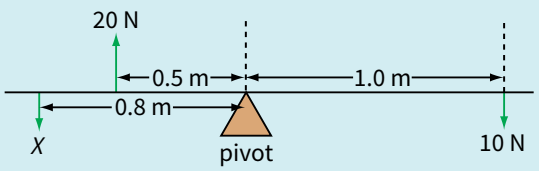


Figure 4.19 For Worked example 3.

The unknown force *X* is tending to turn the beam anticlockwise. The other two forces (10 N and 20 N) are tending to turn the beam clockwise. We will start by calculating their moments and adding them together.

Step 1 Determine the clockwise moments:

sum of moments of clockwise forces

$= (10 \times 1.0) + (20 \times 0.5)$

$= 10 + 10 = 20$ N m

Step 2 Determine the anticlockwise moment:

moment of anticlockwise force = $X \times 0.8$

Step 3 Since we know that the beam must be balanced, we can write:

sum of clockwise moments

= sum of anticlockwise moments

$$20 = X \times 0.8$$

$$X = \frac{20}{0.8} = 25 \text{ N}$$

So a force of 25 N at a distance of 0.8 m from the pivot will keep the beam still and prevent it from rotating (keep it balanced).

4 Figure 4.20 shows the internal structure of a human arm holding an object. The biceps are muscles attached to one of the bones of the forearm. These muscles provide an upward force.

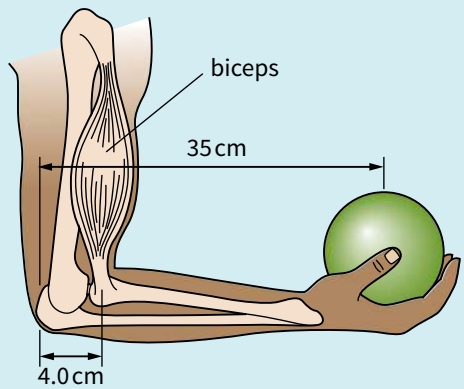


Figure 4.20 The human arm. For Worked example 4.

An object of weight 50 N is held in the hand with the forearm at right angles to the upper arm. Use the principle of moments to determine the muscular force *F* provided by the biceps, given the following data:

weight of forearm = 15 N

distance of biceps from elbow = 4.0 cm

distance of centre of gravity of forearm from elbow = 16 cm

distance of object in the hand from elbow = 35 cm

Step 1 There is a lot of information in this question. It is best to draw a simplified diagram of the forearm that shows all the forces and the relevant distances (Figure 4.21). All distances must be from the pivot, which in this case is the elbow.

WORKED EXAMPLES (continued)

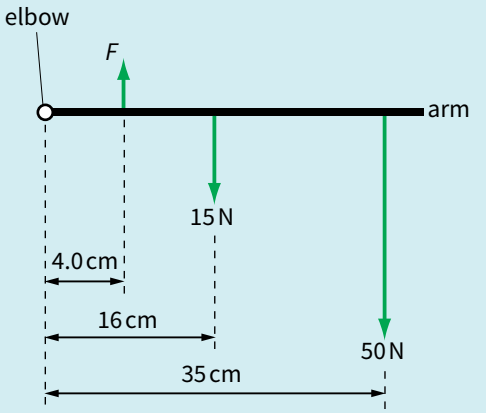


Figure 4.21 Simplified diagram showing forces on the forearm. For Worked example 4. Note that another force acts on the arm at the elbow; we do not know the size or direction of this force but we can ignore it by taking moments about the elbow.

Step 2 Determine the clockwise moments:

sum of moments of clockwise forces
$= (15 \times 0.16) + (50 \times 0.35)$
$= 19.9\ \text{N m}$

Step 3 Determine the anticlockwise moment:

moment of anticlockwise force $= F \times 0.04$

Step 4 Since the arm is in balance, according to the principle of moments we have:

sum of clockwise moments
= sum of anticlockwise moments

$19.9 = 0.04\ F$

$$F = \frac{19.9}{0.04} = 497.5\ \text{N} \approx 500\ \text{N}$$

The biceps provide a force of 500 N – a force large enough to lift 500 apples!

QUESTIONS

7 A wheelbarrow is loaded as shown in Figure 4.22.

a Calculate the force that the gardener needs to exert to hold the wheelbarrow's legs off the ground.

b Calculate the force exerted by the ground on the legs of the wheelbarrow (taken both together) when the gardener is not holding the handles.

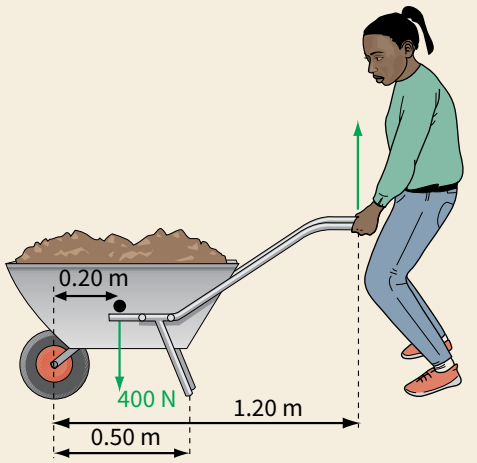


Figure 4.22 For Question 7.

8 A traditional pair of scales uses sliding masses of 10 g and 100 g to achieve a balance. A diagram of the arrangement is shown in Figure 4.23. The bar itself is supported with its centre of gravity at the pivot.

a Calculate the value of the mass *M*, attached at X.

b State **one** advantage of this method of measuring mass.

c Determine the upward force of the pivot on the bar.

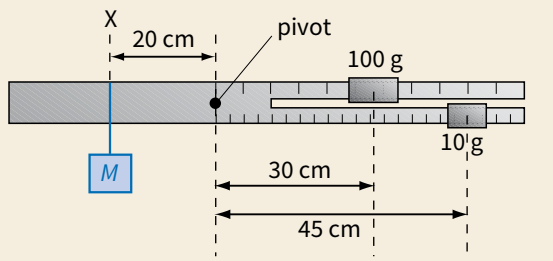


Figure 4.23 For Question 8.

9 Figure 4.24 shows a beam with four forces acting on it

a For each force, calculate the moment of the force about point P.

b State whether each moment is clockwise or anticlockwise.

c State whether or not the moments of the forces are balanced.

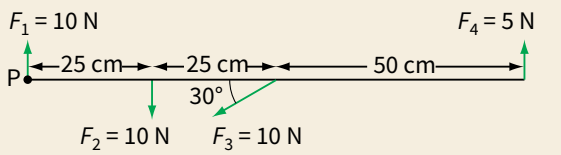


Figure 4.24 For Question 9.

The torque of a couple

Figure 4.25 shows the forces needed to turn a car's steering wheel. The two forces balance up and down (15 N up and 15 N down), so the wheel will not move up, down or sideways. However, the wheel is not in equilibrium. The pair of forces will cause it to rotate.

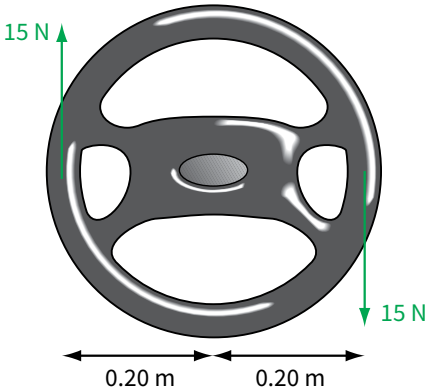


Figure 4.25 Two forces act on this steering wheel to make it turn.

A pair of forces like that in Figure 4.25 is known as a **couple**. A couple has a turning effect, but does not cause an object to accelerate. To form a couple, the two forces must be:

- equal in magnitude
- parallel, but opposite in direction
- separated by a distance *d*.

The turning effect or moment of a couple is known as its **torque**. We can calculate the torque of the couple in Figure 4.25 by adding the moments of each force about the centre of the wheel:

$$\text{torque of couple} = (15 \times 0.20) + (15 \times 0.20)$$

$$= 6.0\ \text{N m}$$

We could have found the same result by multiplying one of the forces by the perpendicular distance between them:

$$\text{torque of a couple} = 15 \times 0.4 = 6.0\ \text{N m}$$

The torque of a couple is defined as follows:

> torque of a couple = one of the forces × perpendicular distance between the forces

QUESTION

10 The driving wheel of a car travelling at a constant velocity has a torque of 137 N m applied to it by the axle that drives the car (Figure 4.26). The radius of the tyre is 0.18 m. Calculate the driving force provided by this wheel.

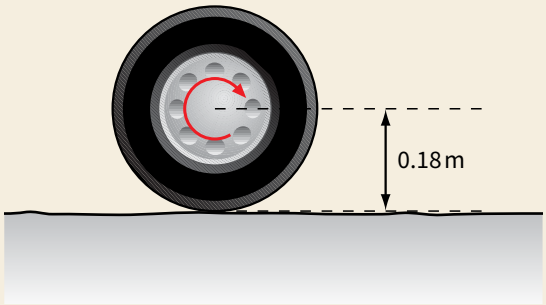


Figure 4.26 For Question 10.

Pure turning effect

When we calculate the moment of a single force, the result depends on the point or pivot about which the moment acts. The further the force is from the pivot, the greater the moment. A couple is different; the moment of a couple does not depend on the point about which it acts, only on the perpendicular distance between the two forces. A single force acting on an object will tend to make the object accelerate (unless there is another force to balance it). A couple, however, is a pair of equal and opposite forces, so it will not make the object accelerate. This means we can think of a couple as a pure 'turning effect', the size of which is given by its torque.

For an object to be in equilibrium, two conditions must be met at the same time:

- The resultant force acting on the object is zero.
- The resultant moment is zero.

Summary

- Forces are vector quantities that can be added by means of a vector triangle. Their resultant can be determined using trigonometry or by scale drawing.
- Vectors such as forces can be resolved into components. Components at right angles to one another can be treated independently of one another. For a force F at an angle θ to the x-direction, the components are:

  x-direction: $F\cos\theta$

  y-direction: $F\sin\theta$
- The moment of a force = force × perpendicular distance of the pivot from the line of action of the force.
- The principle of moments states that, for any object that is in equilibrium, the sum of the clockwise moments about any point provided by the forces acting on the object equals the sum of the anticlockwise moments about that same point.
- A couple is a pair of equal, parallel but opposite forces whose effect is to produce a turning effect on a body without giving it linear acceleration.

  torque of a couple = one of the forces × perpendicular distance between the forces
- For an object to be in equilibrium, the resultant force acting on the object must be zero and the resultant moment must be zero.

End-of-chapter questions

1 A ship is pulled at a constant speed by two small boats, A and B, as shown in Figure 4.27. The engine of the ship does not produce any force.

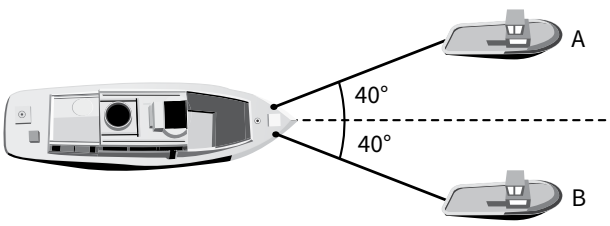


Figure 4.27 For End-of-chapter Question 1.

The tension in each cable between A and B and the ship is 4000 N.

a Draw a free-body diagram showing the three horizontal forces acting on the ship. [2]

b Draw a vector diagram to scale showing these three forces and use your diagram to find the value of the drag force on the ship. [2]

2 A block of mass 1.5 kg is at rest on a rough surface which is inclined at 20° to the horizontal as shown in Figure 4.28.

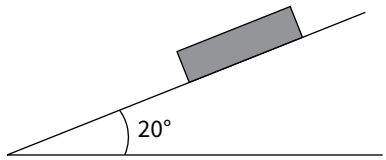


Figure 4.28 For End-of-chapter Question 2.

a Draw a free-body diagram showing the three forces acting on the block. [2]

b Calculate the component of the weight that acts down the slope. [2]

c Use your answer to **b** to determine the force of friction that acts on the block. [2]

d Determine the normal contact force between the block and the surface. [3]

3 The free-body diagram (Figure 4.29) shows three forces that act on a stone hanging at rest from two strings.

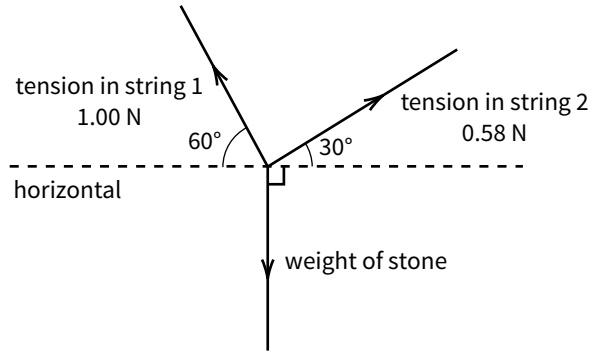


Figure 4.29 For End-of-chapter Question 3.

a Calculate the horizontal component of the tension in each string. Why should these two components be equal in magnitude? [5]

b Calculate the vertical component of the tension in each string. [4]

c Use your answer to **b** to calculate the weight of the stone. [2]

d Draw a vector diagram of the forces on the stone. This should be a triangle of forces. [1]

e Use your diagram in **d** to calculate the weight of the stone. [2]

4 The force *F* shown in Figure 4.30 has a moment of 40 N m about the pivot. Calculate the magnitude of the force *F*. [4]

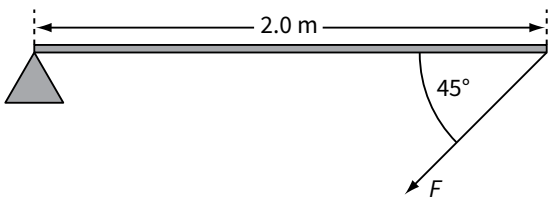


Figure 4.30 For End-of-chapter Question 4.

5 The asymmetric bar shown in Figure 4.31 has a weight of 7.6 N and a centre of gravity that is 0.040 m from the wider end, on which there is a load of 3.3 N. It is pivoted a distance of 0.060 m from its centre of gravity. Calculate the force *P* that is needed at the far end of the bar in order to maintain equilibrium. [4]

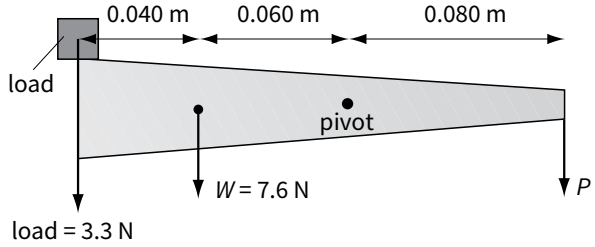


Figure 4.31 For End-of-chapter Question 5.

6 **a** Explain what is meant by:

i a couple [1]

ii torque. [2]

b The engine of a car produces a torque of 200 N m on the axle of the wheel in contact with the road. The car travels at a constant velocity towards the right (Figure 4.32).

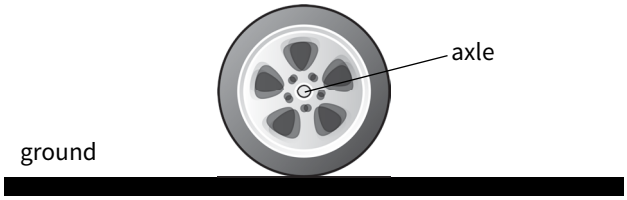


Figure 4.32 For End-of-chapter Question 6.

i Copy Figure 4.32 and show the direction of rotation of the wheel, and the horizontal component of the force that the road exerts on the wheel. [2]

ii State the resultant torque on the wheel. Explain your answer. [2]

iii The diameter of the car wheel is 0.58 m. Determine the value of the horizontal component of the force of the road on the wheel. [1]

7 **a** Explain what is meant by the **centre of gravity** of an object. [2]

b A flagpole of mass 25 kg is held in a horizontal position by a cable as shown in Figure 4.33. The centre of gravity of the flagpole is at a distance of 1.5 m from the fixed end.

i Write an equation to represent taking moments about the left-hand end of the flagpole. Use your equation to find the tension *T* in the cable. [4]

ii Determine the vertical component of the force at the left-hand end of the flagpole. [2]

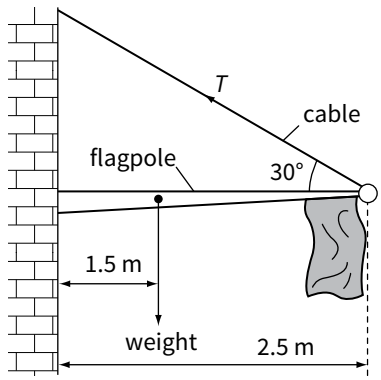


Figure 4.33 For End-of-chapter Question 7.

8 **a** State the **two** conditions necessary for an object to be in equilibrium. [2]

b A metal rod of length 90 cm has a disc of radius 24 cm fixed rigidly at its centre, as shown in Figure 4.34 The assembly is pivoted at its centre.

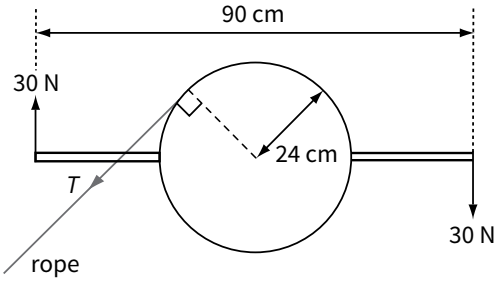


Figure 4.34 For End-of-chapter Question 8.

Two forces, each of magnitude 30 N, are applied normal to the rod at each end so as to produce a turning effect on the rod. A rope is attached to the edge of the disc to prevent rotation.

Calculate:

i the torque of the couple produced by the 30 N forces [1]

ii the tension *T* in the rope. [3]

9 **a** Explain what is meant by the **torque of a couple**. [2]

b Three strings, A, B and C, are attached to a circular ring, as shown in Figure 4.35.

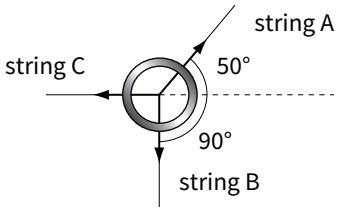


Figure 4.35 For End-of-chapter Question 9.

The strings and the ring all lie on a smooth horizontal surface and are at rest. The tension in string A is 8.0 N. Calculate the tension in strings B and C. [4]

10 Figure 4.36 shows a picture hanging symmetrically by two cords from a nail fixed to a wall. The picture is in equilibrium.

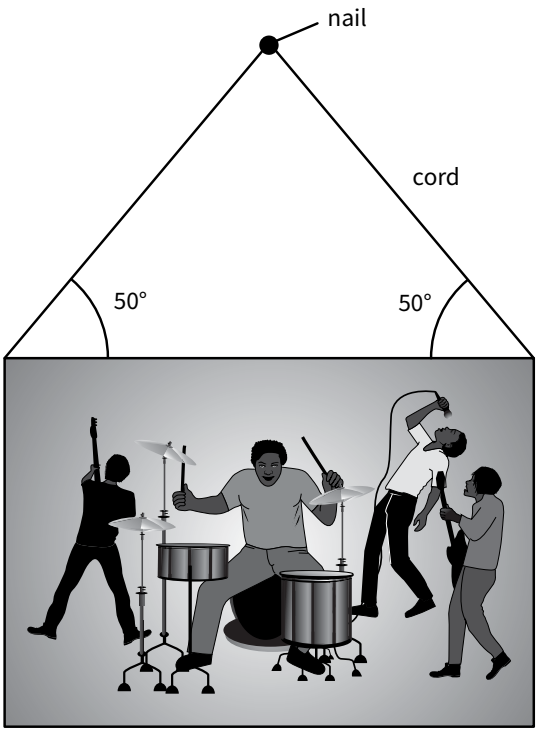


Figure 4.36 For End-of-chapter Question 10.

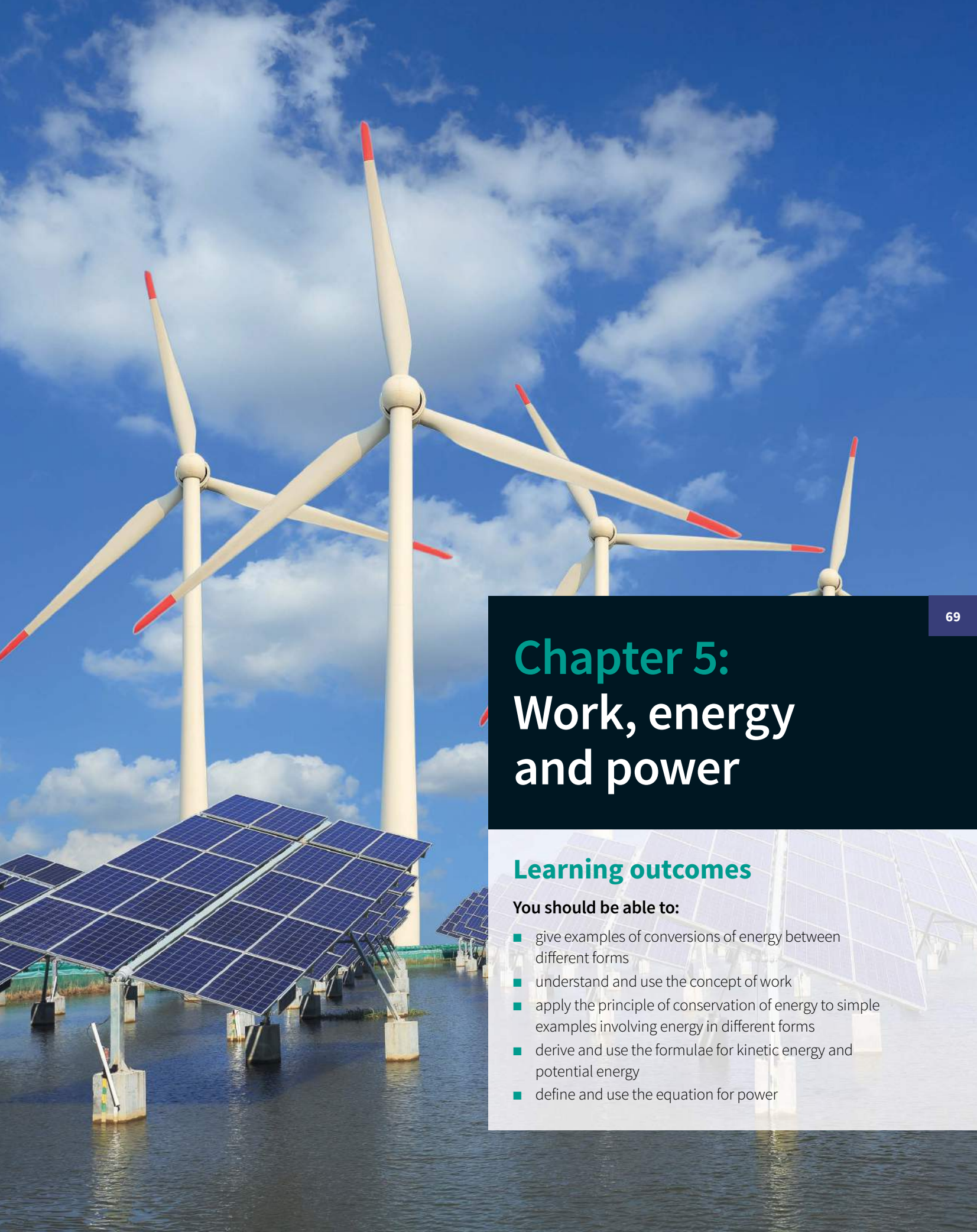
a Explain what is meant by **equilibrium**. [2]

b Draw a vector diagram to represent the three forces acting on the picture in the vertical plane. Label each force clearly with its name and show the direction of each force with an arrow. [2]

c The tension in the cord is 45 N and the angle that each end of the cord makes with the horizontal is 50°. Calculate:

i the vertical component of the tension in the cord [1]

ii the weight of the picture. [1]



Chapter 5: Work, energy and power

Learning outcomes

You should be able to:

- give examples of conversions of energy between different forms
- understand and use the concept of work
- apply the principle of conservation of energy to simple examples involving energy in different forms
- derive and use the formulae for kinetic energy and potential energy
- define and use the equation for power

The idea of energy

The Industrial Revolution started in the late 18th century in the British Isles. Today, many other countries are undergoing the process of industrialisation (Figure 5.1). Industrialisation began as engineers developed new machines which were capable of doing the work of hundreds of craftsmen and labourers. At first, they made use of the traditional techniques of water power and wind power. Water stored behind a dam was used to turn a wheel, which turned many machines. By developing new mechanisms, the designers tried to extract as much as possible of the energy stored in the water. Steam engines were developed, initially for pumping water out of mines. Steam engines use a fuel such as coal; there is much more energy stored in 1 kg of coal than in 1 kg of water held behind a dam. Steam engines soon powered the looms of the textile mills, and the British industry came to dominate world trade in textiles.

Nowadays, most factories and mills rely on electrical power, generated by burning coal or gas at a power station. The fuel is burnt to release its store of energy. High-pressure steam is generated, and this turns a turbine which turns a generator. Even in the most efficient coal-fired power station, only about 40% of the energy from the fuel is transferred to the electrical energy that the station supplies to the grid.

Engineers strove to develop machines which made the most efficient use of the energy supplied to them. At the same time, scientists were working out the basic ideas of energy transfer and energy transformations. The idea of energy itself had to be developed; it was not obvious at first that heat, light, electrical energy and so on could all be thought of as being, in some way, forms of the same thing. In fact, steam engines had been in use for 150 years before it was realised that their energy came from the heat supplied to them from their fuel.

The earliest steam engines had very low efficiencies – many converted less than 1% of the energy supplied to them into useful work. The understanding of the relationship between work and energy led to many ingenious ways of making the most of the energy supplied by fuel.

This improvement in energy efficiency has led to the design of modern engines such as the jet engines which have made long-distance air travel a commercial possibility (Figure 5.2).



Figure 5.2 The jet engines of this aircraft are designed to make efficient use of their fuel. If they were less efficient, their thrust might only be sufficient to lift the empty aircraft, and the passengers would have to be left behind.



Figure 5.1 Anshan steel works, China.

Doing work, transferring energy

The weight-lifter shown in Figure 5.3 has powerful muscles. They can provide the force needed to lift a large weight above her head – about 2 m above the ground. The force exerted by the weight-lifter transfers energy from her to the weights. We know that the weights have gained energy because, when the athlete releases them, they come crashing down to the ground.



Figure 5.3 It is hard work being a weight-lifter.

As the athlete lifts the weights and transfers energy to them, we say that her lifting force is doing work. 'Doing work' is a way of transferring energy from one object to another. In fact, if you want to know the scientific meaning of the word 'energy', we have to say it is 'that which is transferred when a force moves through a distance'. So work and energy are two closely linked concepts.

In physics, we often use an everyday word but with a special meaning. **Work** is an example of this. Table 5.1 describes some situations which illustrate the meaning of **doing work** in physics.

It is important to appreciate that our bodies sometimes mislead us. If you hold a heavy weight above your head for some time, your muscles will get tired. However, you are not doing any work **on the weights**, because you are not transferring energy to the weights once they are above your head. Your muscles get tired because they are constantly relaxing and contracting, and this uses energy, but none of the energy is being transferred to the weights.

Doing work	Not doing work
Pushing a car to start it moving: your force transfers energy to the car. The car's kinetic energy (i.e. 'movement energy') increases.	Pushing a car but it does not budge: no energy is transferred, because your force does not move it. The car's kinetic energy does not change.
Lifting weights: you are doing work as the weights move upwards. The gravitational potential energy of the weights increases.	Holding weights above your head: you are not doing work on the weights (even though you may find it tiring) because the force you apply is not moving them. The gravitational potential energy of the weights is not changing.
A falling stone: the force of gravity is doing work. The stone's kinetic energy is increasing.	The Moon orbiting the Earth: the force of gravity is not doing work. The Moon's kinetic energy is not changing.
Writing an essay: you are doing work because you need a force to move your pen across the page, or to press the keys on the keyboard.	Reading an essay: this may seem like 'hard work', but no force is involved, so you are not doing any work.

Table 5.1 The meaning of 'doing work' in physics.

Calculating work done

Because **doing work** defines what we mean by **energy**, we start this chapter by considering how to calculate **work done**. There is no doubt that you do work if you push a car along the road. A force transfers energy from you to the car. But how much work do you do? Figure 5.4 shows the two factors involved:

- the size of the force F – the bigger the force, the greater the amount of work you do
- the distance s you push the car – the further you push it, the greater the amount of work done.

So, the bigger the force, and the further it moves, the greater the amount of work done.

The **work done** by a force is defined as the product of the force and the distance moved in the direction of the force:

$$W = F \times s$$

where s is the distance moved in the direction of the force.

In the example shown in Figure 5.4, $F = 300\,\text{N}$ and $s = 5.0\,\text{m}$, so:

work done $W = F \times s = 300 \times 5.0 = 1500\,\text{J}$

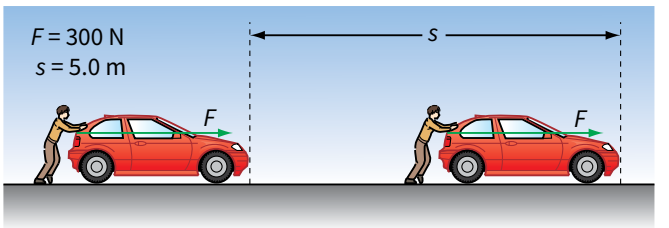


Figure 5.4 You have to do work to start the car moving.

Energy transferred

Doing work is a way of transferring energy. For both energy and work the correct SI unit is the joule (J). The amount of work done, calculated using $W = F \times s$, shows the amount of energy transferred:

work done = energy transferred

Newtons, metres and joules

From the equation $W = F \times s$ we can see how the unit of force (the newton), the unit of distance (the metre) and the unit of work or energy (the joule) are related.

1 joule = 1 newton × 1 metre

1 J = 1 N m

The joule is defined as the amount of work done when a force of 1 newton moves a distance of 1 metre in the direction of the force. Since **work done = energy transferred**, it follows that a joule is also the amount of energy transferred when a force of 1 newton moves a distance of 1 metre in the direction of the force.

QUESTIONS

1 In each of the following examples, explain whether or not any work is done by the force mentioned.

- **a** You pull a heavy sack along rough ground.
- **b** The force of gravity pulls you downwards when you fall off a wall.
- **c** The tension in a string pulls on a stone when you whirl it around in a circle at a steady speed.
- **d** The contact force of the bedroom floor stops you from falling into the room below.

2 A man of mass 70 kg climbs stairs of vertical height 2.5 m. Calculate the work done against the force of gravity. (Take $g = 9.81\,\text{m s}^{-2}$.)

3 A stone of weight 10 N falls from the top of a 250 m high cliff.

- **a** Calculate how much work is done by the force of gravity in pulling the stone to the foot of the cliff.
- **b** How much energy is transferred to the stone?

Force, distance and direction

It is important to appreciate that, for a force to do work, there must be movement in the direction of the force. Both the force F and the distance s moved **in the direction of the force** are vector quantities, so you should know that their directions are likely to be important. To illustrate this, we will consider three examples involving gravity (Figure 5.5). In the equation for work done, $W = F \times s$, the distance moved s is thus the displacement in the direction of the force.

Suppose that the force F moves through a distance s which is at an angle θ to F, as shown in Figure 5.6. To determine the work done by the force, it is simplest to determine the component of F in the direction of s. This component is $F\cos\theta$, and so we have:

work done = $(F\cos\theta) \times s$

or simply:

work done = $Fs\cos\theta$

Worked example 1 shows how to use this.

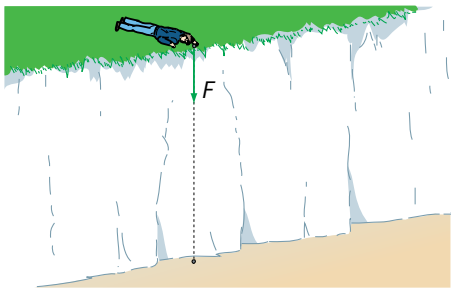
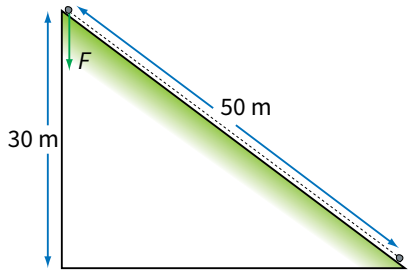
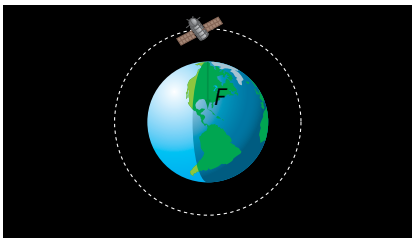
Doing work		Not doing work
		
1 You drop a stone weighing 5.0 N from the top of a 50 m high cliff. What is the work done by the force of gravity? force on stone F = pull of gravity = weight of stone = 5.0 N vertically downwards	**2** A stone weighing 5.0 N rolls 50 m down a slope. What is the work done by the force of gravity? force on stone F = pull of gravity = weight of stone = 5.0 N vertically downwards	**3** A satellite orbits the Earth at a constant height and at a constant speed. The weight of the satellite at this height is 500 N. What is the work done by the force of gravity? force on satellite F = pull of gravity = weight of satellite = 500 N towards centre of Earth
Distance moved by stone is s = 50 m vertically downwards.	Distance moved by stone down slope is 50 m, but distance moved in direction of force is 30 m.	Distance moved by satellite towards centre of Earth (i.e. in the direction of force) is s = 0.
Since F and s are in the same direction, there is no problem: work done = $F \times s$ = 5.0 × 50 = 250 J	The work done by the force of gravity is: work done = 5.0 × 30 = 150 J	The satellite remains at a constant distance from the Earth. It does not move in the direction of F. The work done by the Earth's pull on the satellite is zero because F = 500 N but s = 0: work done = 500 × 0 = 0 J

Figure 5.5 Three examples involving gravity.

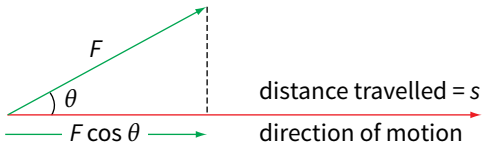


Figure 5.6 The work done by a force depends on the angle between the force and the distance it moves.

WORKED EXAMPLE

1 A man pulls a box along horizontal ground using a rope (Figure 5.7). The force provided by the rope is 200 N, at an angle of 30° to the horizontal. Calculate the work done if the box moves 5.0 m along the ground.

Step 1 Calculate the component of the force in the direction in which the box moves. This is the horizontal component of the force:

horizontal component of force = 200 cos 30° ≈ 173 N

Hint: *$F \cos\theta$ is the component of the force at an angle θ to the direction of motion.*

Step 2 Now calculate the work done:

work done = force × distance moved = 173 × 5.0 = 865 J

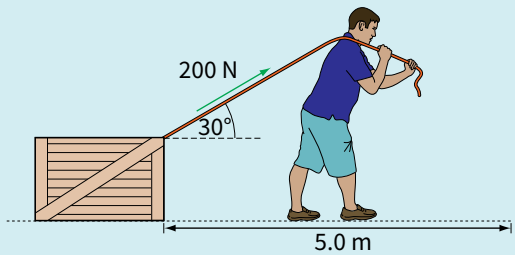


Figure 5.7 For Worked example 1.

Hint: *Note that we could have used the equation work done = $Fs \cos\theta$ to combine the two steps into one.*

A gas doing work

Gases exert pressure on the walls of their container. If a gas expands, the walls are pushed outwards – the gas has done work on its surroundings. In a steam engine, expanding steam pushes a piston to turn the engine, and in a car engine, the exploding mixture of fuel and air does the same thing, so this is an important situation.

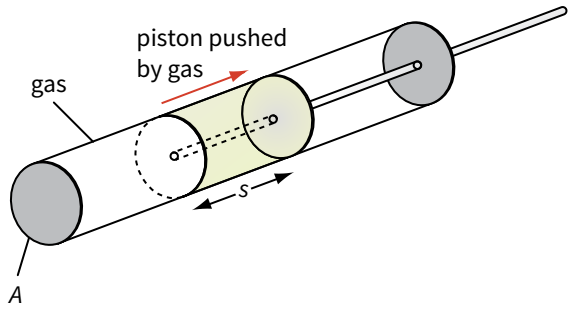


Figure 5.8 When a gas expands, it does work on its surroundings.

Figure 5.8 shows a gas at pressure p inside a cylinder of cross-sectional area A. The cylinder is closed by a moveable piston. The gas pushes the piston a distance s. If we know the force F exerted by the gas on the piston, we can deduce an expression for the amount of work done by the gas.

From the definition of pressure (pressure = $\frac{\text{force}}{\text{area}}$), the force exerted by the gas on the piston is given by:

$$\text{force} = \text{pressure} \times \text{area}$$

$$F = p \times A$$

and the work done is force × displacement:

$$W = p \times A \times s$$

But the quantity $A \times s$ is the **increase** in volume of the gas; that is, the shaded volume in Figure 5.8. We call this ΔV, where the Δ indicates that it is a **change** in V. Hence the work done by the gas in expanding is:

$$W = p\Delta V$$

Notice that we are assuming that the pressure p does not change as the gas expands. This will be true if the gas is expanding against the pressure of the atmosphere, which changes only very slowly.

QUESTIONS

4 The crane shown in Figure 5.9 lifts its 500 N load to the top of the building from A to B. Distances are as shown on the diagram. Calculate how much work is done by the crane.

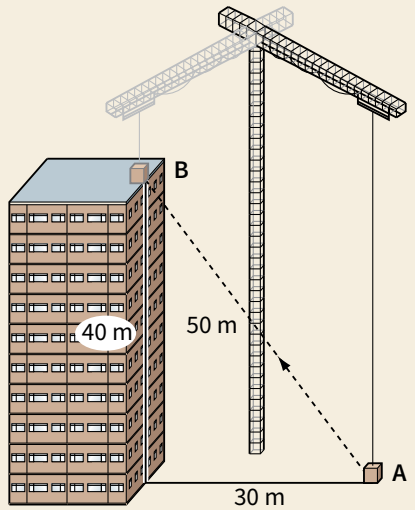


Figure 5.9 For Question 4. The dotted line shows the track of the load as it is lifted by the crane.

5 Figure 5.10 shows the forces acting on a box which is being pushed up a slope. Calculate the work done by each force if the box moves 0.50 m up the slope.

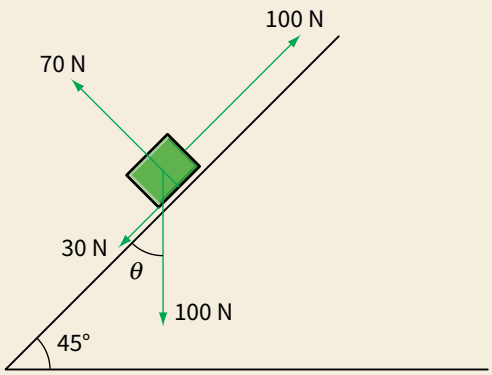


Figure 5.10 For Question 5.

6 When you blow up a balloon, the expanding balloon pushes aside the atmosphere. How much work is done against the atmosphere in blowing up a balloon to a volume of 2 litres (0.002 m^3)? (Atmospheric pressure = 1.0×10^5 N m^{-2}.)

Gravitational potential energy

If you lift a heavy object, you do work. You are providing an upward force to overcome the downward force of gravity on the object. The force moves the object upwards, so the force is doing work.

In this way, energy is transferred from you to the object. You lose energy, and the object gains energy. We say that the **gravitational potential energy** E_p of the object has increased. Worked example 2 shows how to calculate a change in gravitational potential energy – or g.p.e. for short.

WORKED EXAMPLE

2 A weight-lifter raises weights with a mass of 200 kg from the ground to a height of 1.5 m. Calculate how much work he does. By how much does the g.p.e. of the weights increase?

Step 1 As shown in Figure 5.11, the downward force on the weights is their weight, $W = mg$. An equal, upward force F is required to lift them.

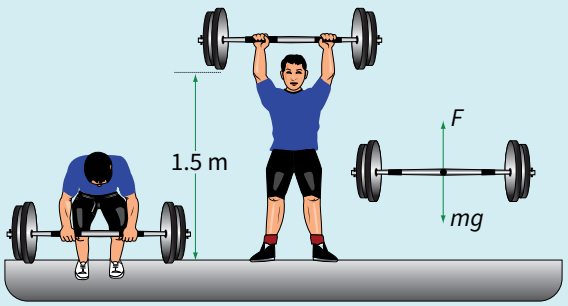


Figure 5.11 For Worked example 2.

$W = F = mg = 200 \times 9.81 = 1962$ N

Hint: *It helps to draw a diagram of the situation.*

Step 2 Now we can calculate the work done by the force F:

work done = force × distance moved

$= 1962 \times 1.5 \approx 2940$ J

Note that the distance moved is in the same direction as the force. So the work done on the weights is about 2940 J. This is also the value of the increase in their g.p.e.

An equation for gravitational potential energy

The change in the gravitational potential energy (g.p.e.) of an object, E_p, depends on the change in its height, h. We can calculate E_p using this equation:

change in g.p.e. = weight × change in height

$E_p = (mg) \times h$

or simply

$E_p = mgh$

It should be clear where this equation comes from. The force needed to lift an object is equal to its weight mg, where m is the mass of the object and g is the acceleration of free fall or the gravitational field strength on the Earth's surface. The work done by this force is given by force × distance moved, or weight × change in height. You might feel that it takes a force greater than the weight of the object being raised to lift it upwards, but this is not so. Provided the force is equal to the weight, the object will move upwards at a steady speed.

Note that h stands for the vertical height through which the object moves. Note also that we can only use the equation $E_p = mgh$ for relatively small changes in height. It would not work, for example, in the case of a satellite orbiting the Earth. Satellites orbit at a height of at least 200 km and g has a smaller value at this height.

Other forms of potential energy

Potential energy is the energy an object has because of its position or shape. So, for example, an object's gravitational potential energy changes when it moves through a gravitational field. (There is much more about gravitational fields in Chapter 18.)

We can identify other forms of potential energy. An electrically charged object has electric potential energy when it is placed in an electric field (see Chapter 8). An object may have elastic potential energy when it is stretched, squashed or twisted – if it is released it goes back to its original shape (see Chapter 7).

QUESTIONS

7 Calculate how much gravitational potential energy is gained if you climb a flight of stairs. Assume that you have a mass of 52 kg and that the height you lift yourself is 2.5 m.

8 A climber of mass 100 kg (including the equipment she is carrying) ascends from sea level to the top of a mountain 5500 m high. Calculate the change in her gravitational potential energy.

9 **a** A toy car works by means of a stretched rubber band. What form of potential energy does the car store when the band is stretched?

b A bar magnet is lying with its north pole next to the south pole of another bar magnet. A student pulls them apart. Why do we say that the magnets' potential energy has increased? Where has this energy come from?

Kinetic energy

As well as lifting an object, a force can make it accelerate. Again, work is done by the force and energy is transferred to the object. In this case, we say that it has gained kinetic energy, E_k. The faster an object is moving, the greater its kinetic energy (k.e.).

For an object of mass m travelling at a speed v, we have:

kinetic energy = $\frac{1}{2}$ × mass × speed^2

$$E_k = \frac{1}{2}mv^2$$

Deriving the formula for kinetic energy

The equation for k.e., $E_k = \frac{1}{2}mv^2$, is related to one of the equations of motion. We imagine a car being accelerated from rest ($u = 0$) to velocity v. To give it acceleration a, it is pushed by a force F for a distance s. Since $u = 0$, we can write the equation $v^2 = u^2 + 2as$ as:

$$v^2 = 2as$$

Multiplying both sides by $\frac{1}{2}m$ gives:

$$\frac{1}{2}mv^2 = mas$$

Now, ma is the force F accelerating the car, and mas is the force × the distance it moves, that is, the work done by the force. So we have:

$\frac{1}{2}mv^2$ = work done by force F

This is the energy transferred to the car, and hence its kinetic energy.

WORKED EXAMPLE

3 Calculate the increase in kinetic energy of a car of mass 800 kg when it accelerates from 20 m s^{-1} to 30 m s^{-1}.

Step 1 Calculate the initial k.e. of the car:

$E_k = \frac{1}{2}mv^2 = \frac{1}{2} \times 800 \times (20)^2 = 160\,000$ J

= 160 kJ

Step 2 Calculate the final k.e. of the car:

$E_k = \frac{1}{2}mv^2 = \frac{1}{2} \times 800 \times (30)^2 = 360\,000$ J

= 360 kJ

Step 3 Calculate the change in the car's k.e.:

change in k.e. = 360 − 160 = 200 kJ

Hint: *Take care! You can't calculate the change in k.e. by squaring the change in speed. In this example, the change in speed is 10 m s^{-1}, and this would give an incorrect value for the change in k.e.*

QUESTIONS

10 Which has more k.e., a car of mass 500 kg travelling at 15 m s^{-1} or a motorcycle of mass 250 kg travelling at 30 m s^{-1}?

11 Calculate the change in kinetic energy of a ball of mass 200 g when it bounces. Assume that it hits the ground with a speed of 15.8 m s^{-1} and leaves it at 12.2 m s^{-1}.

g.p.e.–k.e. transformations

A motor drags the roller-coaster car to the top of the first hill. The car runs down the other side, picking up speed as it goes (see Figure 5.12). It is moving just fast enough to reach the top of the second hill, slightly lower than the first. It accelerates downhill again. Everybody screams!

The motor provides a force to pull the roller-coaster car to the top of the hill. It transfers energy to the car. But where is this energy when the car is waiting at the top of the hill? The car now has gravitational potential energy; as soon as it is given a small push to set it moving, it accelerates. It gains kinetic energy and at the same time it loses g.p.e.



Figure 5.12 The roller-coaster car accelerates as it comes downhill. It's even more exciting if it runs through water.

As the car runs along the roller-coaster track (Figure 5.13), its energy changes.

1. At the top of the first hill, it has the most g.p.e.
2. As it runs downhill, its g.p.e. decreases and its k.e. increases.
3. At the bottom of the hill, all of its g.p.e. has been changed to k.e. and heat and sound energy.
4. As it runs back uphill, the force of gravity slows it down. k.e. is being changed to g.p.e.

Inevitably, some energy is lost by the car. There is friction with the track, and air resistance. So the car cannot return to its original height. That is why the second hill must be slightly lower than the first. It is fun if the car runs through a trough of water, but that takes even more energy, and the car cannot rise so high. There are many situations where an object's energy changes between gravitational potential energy and kinetic energy. For example:

- a high diver falling towards the water – g.p.e. changes to k.e.
- a ball is thrown upwards – k.e. changes to g.p.e.
- a child on a swing – energy changes back and forth between g.p.e. and k.e.

Down, up, down – energy changes

When an object falls, it speeds up. Its g.p.e. decreases and its k.e. increases. Energy is being transformed from gravitational potential energy to kinetic energy. Some energy is likely to be lost, usually as heat because of air resistance. However, if no energy is lost in the process, we have:

decrease in g.p.e. = gain in k.e.

We can use this idea to solve a variety of problems, as illustrated by Worked example 4.

WORKED EXAMPLE

4 A pendulum consists of a brass sphere of mass 5.0 kg hanging from a long string (see Figure 5.14). The sphere is pulled to the side so that it is 0.15 m above its lowest position. It is then released. How fast will it be moving when it passes through the lowest point along its path?

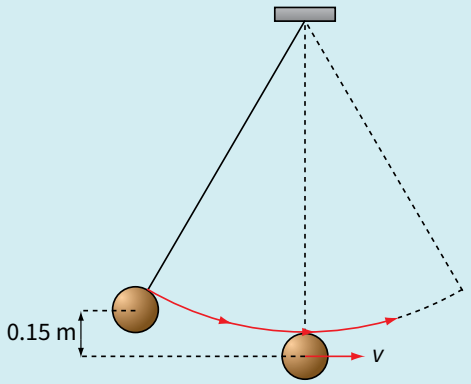


Figure 5.14 For Worked example 4.

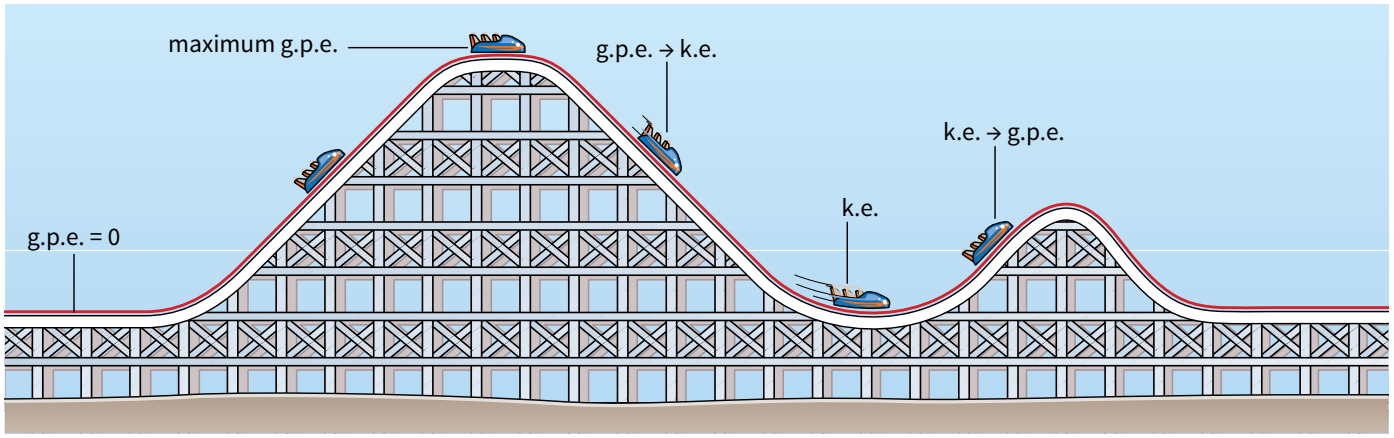


Figure 5.13 Energy changes along a roller-coaster.

WORKED EXAMPLE (continued)

Step 1 Calculate the loss in g.p.e. as the sphere falls from its highest position:

$E_p = mgh = 5.0 \times 9.81 \times 0.15 = 7.36\,\text{J}$

Step 2 The gain in the sphere's k.e. is 7.36 J. We can use this to calculate the sphere's speed. First calculate v^2, then v:

$\frac{1}{2}mv^2 = 7.36$

$\frac{1}{2} \times 5.0 \times v^2 = 7.36$

$$v^2 = 2 \times \frac{7.36}{5.0} = 2.944$$

$$v = \sqrt{2.944} \approx 1.72\,\text{ms}^{-1} \approx 1.7\,\text{ms}^{-1}$$

Note that we would obtain the same result in Worked example 4 no matter what the mass of the sphere. This is because both k.e. and g.p.e. depend on mass m. If we write:

change in g.p.e. = change in k.e.

$$mgh = \frac{1}{2}mv^2$$

we can cancel m from both sides. Hence:

$$gh = \frac{v^2}{2}$$

$v^2 = 2gh$

Therefore:

$v = \sqrt{2gh}$

The final speed v only depends on g and h. The mass m of the object is irrelevant. This is not surprising; we could use the same equation to calculate the speed of an object falling from height h. An object of small mass gains the same speed as an object of large mass, provided air resistance has no effect.

Energy transfers

Climbing bars

If you are going to climb a mountain, you will need a supply of energy. This is because your gravitational potential energy is greater at the top of the mountain than at the base. A good supply of energy would be some bars of chocolate. Each bar supplies 1200 kJ. Suppose your weight is 600 N and you climb a 2000 m high mountain. The work done by your muscles is:

work done = $Fs = 600 \times 2000 = 1200\,\text{kJ}$

So one bar of chocolate will do the trick. Of course, in reality, it would not. Your body is inefficient. It cannot convert 100% of the energy from food into gravitational potential energy. A lot of energy is wasted as your muscles warm up, you perspire, and your body rises and falls as you walk along the path. Your body is perhaps only 5% efficient as far as climbing is concerned, and you will need to eat 20 chocolate bars to get you to the top of the mountain. And you will need to eat more to get you back down again.

Many energy transfers are inefficient. That is, only part of the energy is transferred to where it is wanted. The rest is wasted, and appears in some form that is not wanted (such as waste heat), or in the wrong place. You can determine the efficiency of any device or system using the following equation:

$$\text{efficiency} = \frac{\text{useful output energy}}{\text{total input energy}} \times 100\%$$

A car engine is more efficient than a human body, but not much more. Figure 5.16 shows how this can be represented by a Sankey diagram. The width of the arrow represents the fraction of the energy which is transformed to each new form. In the case of a car engine, we want it to provide

QUESTIONS

12 Re-work Worked example 4 for a brass sphere of mass 10 kg, and show that you get the same result. Repeat with any other value of mass.

13 Calculate how much gravitational potential energy is lost by an aircraft of mass 80 000 kg if it descends from an altitude of 10 000 m to an altitude of 1000 m. What happens to this energy if the pilot keeps the aircraft's speed constant?

14 A high diver (see Figure 5.15) reaches the highest point in her jump with her centre of gravity 10 m above the water. Assuming that all her gravitational potential energy becomes kinetic energy during the dive, calculate her speed just before she enters the water.



Figure 5.15 A high dive is an example of converting (transforming) gravitational potential energy to kinetic energy.



Figure 5.16 We want a car engine to supply kinetic energy. This Sankey diagram shows that only 20% of the energy supplied to the engine ends up as kinetic energy – it is 20% efficient.

kinetic energy to turn the wheels. In practice, 80% of the energy is transformed into heat: the engine gets hot, and heat escapes into the surroundings. So the car engine is only 20% efficient.

We have previously considered situations where an object is falling, and all of its gravitational potential energy changes to kinetic energy. In Worked example 5, we will look at a similar situation, but in this case the energy change is not 100% efficient.

Conservation of energy

Where does the lost energy from the water in the reservoir go? Most of it ends up warming the water, or warming the pipes that the water flows through. The outflow of water is probably noisy, so some sound is produced.

Here, we are assuming that all of the energy ends up somewhere. None of it disappears. We assume the same thing when we draw a Sankey diagram. The total thickness of the arrow remains constant. We could not have an arrow which got thinner (energy disappearing) or thicker (energy appearing out of nowhere).

We are assuming that energy is conserved. This is a principle, known as the **principle of conservation of energy**, which we expect to apply in all situations.

> Energy cannot be created or destroyed. It can only be converted from one form to another.

We should always be able to add up the total amount of energy at the beginning, and be able to account for it all at the end. We cannot be sure that this is always the case, but we expect it to hold true.

We have to think about energy changes within a closed system; that is, we have to draw an imaginary boundary around all of the interacting objects which are involved in an energy transfer.

WORKED EXAMPLE

5 Figure 5.17 shows a dam which stores water. The outlet of the dam is 20 m below the surface of the water in the reservoir. Water leaving the dam is moving at 16 m s^{-1}. Calculate the percentage of the gravitational potential energy that is lost when converted into kinetic energy.

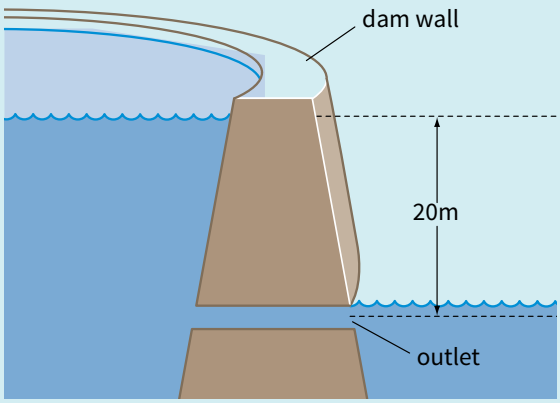


Figure 5.17 Water stored behind the dam has gravitational potential energy; the fast-flowing water leaving the foot of the dam has kinetic energy.

Step 1 We will picture 1 kg of water, starting at the surface of the lake (where it has g.p.e., but no k.e.) and flowing downwards and out at the foot (where it has k.e., but less g.p.e.). Then:

change in g.p.e. of water between surface and outflow
$= mgh = 1 \times 9.81 \times 20 = 196$ J

Step 2 Calculate the k.e. of 1 kg of water as it leaves the dam:

k.e. of water leaving dam $= \frac{1}{2}mv^2$
$= \frac{1}{2} \times 1 \times (16)^2$
$= 128$ J

Step 3 For each kilogram of water flowing out of the dam, the loss of energy is:

loss $= 196 - 128 = 68$ J

$$\text{percentage loss} = \frac{68}{196} \times 100\% \approx 35\%$$

If you wanted to use this moving water to generate electricity, you would have already lost more than a third of the energy which it stores when it is behind the dam.

Sometimes, applying the principle of conservation of energy can seem like a scientific fiddle. When physicists were investigating radioactive decay involving beta particles, they found that the particles after the decay had less energy in total than the particles before. They guessed that there was another, invisible particle which was carrying away the missing energy. This particle, named the neutrino, was proposed by the theoretical physicist Wolfgang Pauli in 1931. The neutrino was not detected by experimenters until 25 years later.

Although we cannot prove that energy is always conserved, this example shows that the principle of conservation of energy can be a powerful tool in helping us to understand what is going on in nature, and that it can help us to make fruitful predictions about future experiments.

QUESTION

15 A stone falls from the top of a cliff, 80 m high. When it reaches the foot of the cliff, its speed is 38 m s^{-1}.

a Calculate the proportion of the stone's initial g.p.e. that is converted to k.e.

b What happens to the rest of the stone's initial energy?

Power

The word **power** has several different meanings – political power, powers of ten, electrical power from power stations. In physics, it has a specific meaning which is related to these other meanings. Figure 5.18 illustrates what we mean by power in physics.

The lift shown in Figure 5.18 can lift a heavy load of people. The motor at the top of the building provides a force to raise the lift car, and this force does work against the force of gravity. The motor transfers energy to the lift car. The **power** P of the motor is the rate at which it does work. Power is defined as the rate of work done. As a word equation, power is given by:

$$\text{power} = \frac{\text{work done}}{\text{time taken}}$$

or

$$P = \frac{W}{t}$$

where W is the work done in a time t.

Units of power: the watt

Power is measured in watts, named after James Watt, the Scottish engineer famous for his development of the steam engine in the second half of the 18th century. The watt is defined as a rate of working of 1 joule per second. Hence:

1 watt = 1 joule per second

or

1 W = 1 J s^{-1}

In practice we also use kilowatts (kW) and megawatts (MW).

1000 watts = 1 kilowatt (1 kW)

1 000 000 watts = 1 megawatt (1 MW)

You are probably familiar with the labels on light bulbs which indicate their power in watts, for example 60 W or 10 W. The values of power on the labels tell you about the energy transferred by an electrical current, rather than by a force doing work.



Figure 5.18 A lift needs a powerful motor to raise the car when it has a full load of people. The motor does many thousands of joules of work each second.

QUESTIONS

16 Calculate how much work is done by a 50 kW car engine in a time of 1.0 minute.

17 A car engine does 4200 kJ of work in one minute. Calculate its output power, in kilowatts.

18 A particular car engine provides a force of 700 N when the car is moving at its top speed of 40 m s^{-1}.

a Calculate how much work is done by the car's engine in one second.

b State the output power of the engine.

WORKED EXAMPLE

6 The motor of the lift shown in Figure 5.18 provides a force of 20 kN; this force is enough to raise the lift by 18 m in 10 s. Calculate the output power of the motor.

Step 1 First, we must calculate the work done:

work done = force × distance moved

= 20 × 18 = 360 kJ

Step 2 Now we can calculate the motor's output power:

$$\text{power} = \frac{\text{work done}}{\text{time taken}} = \frac{360 \times 10^3}{10} = 36\,\text{kW}$$

Hint: *Take care not to confuse the two uses of the letter 'W':*
W = *watt (a unit)*
W = work done (a quantity)

So the lift motor's power is 36 kW. Note that this is its mechanical power output. The motor cannot be 100% efficient since some energy is bound to be wasted as heat due to friction, so the electrical power input must be more than 36 kW.

Moving power

An aircraft is kept moving forwards by the force of its engines pushing air backwards. The greater the force and the faster the aircraft is moving, the greater the power supplied by its engines.

Suppose that an aircraft is moving with velocity v. Its engines provide the force F needed to overcome the drag of the air. In time t, the aircraft moves a distance s equal to $v \times t$. So the work done by the engines is:

work done = force × distance

$$W = F \times v \times t$$

and the power P $\left(= \frac{\text{work done}}{\text{time taken}}\right)$ is given by:

$$P = \frac{W}{t} = \frac{F \times v \times t}{t}$$

and we have:

$$P = F \times v$$

power = force × velocity

It may help to think of this equation in terms of units. The right-hand side is in $\text{N} \times \text{m s}^{-1}$, and N m is the same as J. So the right-hand side has units of J s^{-1}, or W, the unit of power. If you look back to Question 18 above, you will see that, to find the power of the car engine, rather than considering the work done in 1 s, we could simply have multiplied the engine's force by the car's speed.

Human power

Our energy supply comes from our food. A typical diet supplies 2000–3000 kcal (kilocalories) per day. This is equivalent (in SI units) to about 10 MJ of energy. We need this energy for our daily requirements – keeping warm, moving about, brainwork and so on. We can determine the average power of all the activities of our body:

average power = 10 MJ per day

$$= 10 \times \frac{10^6}{86\,400} = 116\,\text{W}$$

So we dissipate energy at the rate of about 100 W. We supply roughly as much energy to our surroundings as a 100 W light bulb. Twenty people will keep a room as warm as a 2 kW electric heater.

Note that this is our average power. If you are doing some demanding physical task, your power will be greater. This is illustrated in Worked example 7.

Note also that the human body is not a perfectly efficient system; a lot of energy is wasted when, for example, we lift a heavy load. We might increase an object's g.p.e. by 1000 J when we lift it, but this might require five or ten times this amount of energy to be expended by our bodies.

QUESTION

19 In an experiment to measure a student's power, she times herself running up a flight of steps. Use the data below to work out her useful power.

number of steps = 28
height of each step = 20 cm
acceleration of free fall = 9.81 m s^{-2}
mass of student = 55 kg
time taken = 5.4 s

WORKED EXAMPLE

7 A person who weighs 500 N runs up a flight of stairs in 5.0 s (Figure 5.19). Their gain in height is 3.0 m. Calculate the rate at which work is done against the force of gravity.

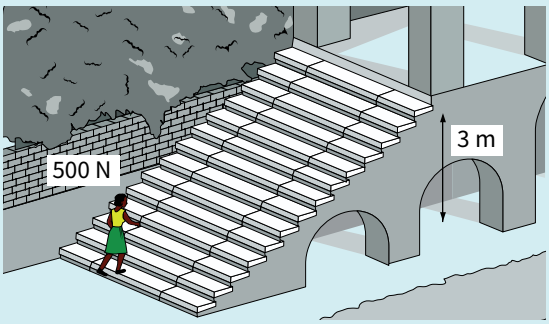


Figure 5.19 Running up stairs can require a high rate of doing work. You may have investigated your own power in this way.

Step 1 Calculate the work done against gravity:

work done $W = F \times s = 500 \times 3.0 = 1500\,\text{J}$

Step 2 Now calculate the power:

$$\text{power } P = \frac{W}{t} = \frac{1500}{5.0} = 300\,\text{W}$$

So, while the person is running up the stairs, they are doing work against gravity at a greater rate than their average power – perhaps three times as great. And, since our muscles are not very efficient, they need to be supplied with energy even faster, perhaps at a rate of 1 kW. This is why we cannot run up stairs all day long without greatly increasing the amount we eat. The inefficiency of our muscles also explains why we get hot when we exert ourselves.

Summary

- The work done W when a force F moves through a displacement s in the direction of the force:

  $W = Fs$ or $W = Fs\cos\theta$

  where θ is the angle between the force and the displacement.
- A joule is defined as the work done (or energy transferred) when a force of 1 N moves a distance of 1 m in the direction of the force.
- The work done W by a gas at pressure p when it expands:

  $W = p\Delta V$

  where ΔV is the increase in its volume.
- When an object of mass m rises through a height h, its gravitational potential energy E_p increases by an amount:

  $E_p = mgh$
- The kinetic energy E_k of a body of mass m moving at speed v is:

  $E_k = \frac{1}{2}mv^2$
- The principle of conservation of energy states that, for a closed system, energy can be transformed to other forms but the total amount of energy remains constant.
- The efficiency of a device or system is determined using the equation:

  $$\text{efficiency} = \frac{\text{useful output energy}}{\text{total input energy}} \times 100\%$$
- Power is the rate at which work is done (or energy is transferred):

  $P = \frac{W}{t}$ and $P = Fv$
- A watt is defined as a rate of transfer of energy of one joule per second.

End-of-chapter questions

1 In each case below, discuss the energy changes taking place:

a An apple falling towards the ground [1]

b A car decelerating when the brakes are applied [1]

c A space probe falling towards the surface of a planet. [1]

2 A 120 kg crate is dragged along the horizontal ground by a 200 N force acting at an angle of 30° to the horizontal, as shown in Figure 5.20. The crate moves along the surface with a constant velocity of 0.5 m s^{-1}. The 200 N force is applied for a time of 16 s.

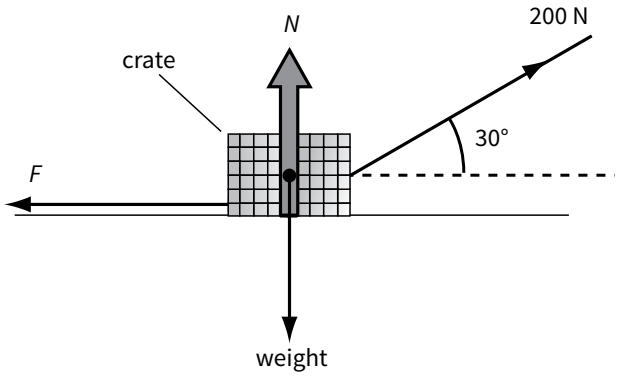


Figure 5.20 For End-of-chapter Question 2.

a Calculate the work done on the crate by:

i the 200 N force [3]

ii the weight of the crate [2]

iii the normal contact force *N*. [2]

b Calculate the rate of work done against the frictional force *F*. [2]

3 Which of the following has greater kinetic energy?

- A 20-tonne truck travelling at a speed of 30 m s^{-1}
- A 1.2 g dust particle travelling at 150 km s^{-1} through space. [3]

4 A 950 kg sack of cement is lifted to the top of a building 50 m high by an electric motor.

a Calculate the increase in the gravitational potential energy of the sack of cement. [2]

b The output power of the motor is 4.0 kW. Calculate how long it took to raise the sack to the top of the building. [2]

c The electrical power transferred by the motor is 6.9 kW. In raising the sack to the top of the building, how much energy is wasted in the motor as heat? [3]

5 **a** Define **power** and state its unit. [2]

b Write a word equation for the kinetic energy of a moving object. [1]

c A car of mass 1100 kg starting from rest reaches a speed of 18 m s^{-1} in 25 s. Calculate the average power developed by the engine of the car. [2]

6 A cyclist pedals a long slope which is at 5.0° to the horizontal (Figure 5.21). The cyclist starts from rest at the top of the slope and reaches a speed of 12 m s^{-1} after a time of 67 s, having travelled 40 m down the slope. The total mass of the cyclist and bicycle is 90 kg.

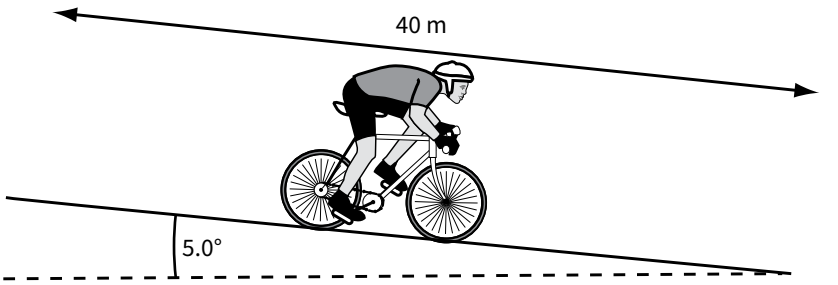


Figure 5.21 For End-of-chapter Question 6.

a Calculate:

i the loss in gravitational potential energy as he travels down the slope [3]

ii the increase in kinetic energy as he travels down the slope. [2]

b i Use your answers to **a** to determine the useful power output of the cyclist. [3]

ii Suggest **one** reason why the actual power output of the cyclist is larger than your value in **i**. [2]

7 a Explain what is meant by **work**. [2]

b i Explain how the principle of conservation of energy applies to a man sliding from rest down a vertical pole, if there is a constant force of friction acting on him. [2]

ii The man slides down the pole and reaches the ground after falling a distance h = 15 m. His potential energy at the top of the pole is 1000 J. Sketch a graph to show how his gravitational potential energy E_p varies with h. Add to your graph a line to show the variation of his kinetic energy E_k with h. [3]

8 a Use the equations of motion to show that the kinetic energy of an object of mass m moving with velocity v is $\frac{1}{2}mv^2$. [2]

b A car of mass 800 kg accelerates from rest to a speed of 20 m s^{-1} in a time of 6.0 s.

i Calculate the average power used to accelerate the car in the first 6.0 s. [2]

ii The power passed by the engine of the car to the wheels is constant. Explain why the acceleration of the car decreases as the car accelerates. [2]

9 a i Define **potential energy**. [1]

ii Distinguish between gravitational potential energy and elastic potential energy. [2]

b Seawater is trapped behind a dam at high tide and then released through turbines. The level of the water trapped by the dam falls 10.0 m until it is all at the same height as the sea.

i Calculate the mass of seawater covering an area of 1.4 × 10^6 m^2 and with a depth of 10.0 m. (Density of seawater = 1030 kg m^{-3}) [1]

ii Calculate the maximum loss of potential energy of the seawater in **i** when passed through the turbines. [2]

iii The potential energy of the seawater, calculated in **ii**, is lost over a period of 6.0 hours. Estimate the average power output of the power station over this time period, given that the efficiency of the power station is 50%. [3]



Chapter 6: Momentum

Learning outcomes

You should be able to:

- define linear momentum
- state and apply the principle of conservation of momentum to collisions in one and two dimensions
- relate force to the rate of change of momentum
- discuss energy changes in perfectly elastic and inelastic collisions

Understanding collisions

To improve the safety of cars the motion of a car during a crash must be understood and the forces on the driver minimised (Figure 6.1). In this way safer cars have been developed and many lives have been saved.

In this chapter, we will explore how the idea of momentum can allow us to predict how objects move after colliding (interacting) with each other. We will also see how Newton's laws of motion can be expressed in terms of momentum.



Figure 6.1 A high-speed photograph of a crash test. The cars collide head-on at 15 m s^{-1} with dummies as drivers.

The idea of momentum

Snooker players can perform some amazing moves on the table, without necessarily knowing Newton's laws of motion – see Figure 6.2. However, the laws of physics can help us to understand what happens when two snooker balls collide or when one bounces off the side cushion of the table.

Here are some examples of situations involving collisions:

- Two cars collide head-on.
- A fast-moving car runs into the back of a slower car in front.
- A footballer runs into an opponent.
- A hockey stick strikes a ball.
- A comet or an asteroid collides with a planet as it orbits the Sun.
- The atoms of the air collide constantly with each other, and with the walls of their surroundings.



Figure 6.2 If you play pool often enough, you will be able to predict how the balls will move on the table. Alternatively, you can use the laws of physics to predict their motion.

- Electrons that form an electric current collide with the vibrating ions that make up a metal wire.
- Two distant galaxies collide over millions of years.

From these examples, we can see that collisions are happening all around us, all the time. They happen on the microscopic scale of atoms and electrons, they happen in our everyday world, and they also happen on the cosmic scale of our Universe.

Modelling collisions

Springy collisions

Figure 6.3a shows what happens when one snooker ball collides head-on with a second, stationary ball. The result can seem surprising. The moving ball stops dead. The ball initially at rest moves off with the same velocity as that of the original ball. To achieve this, a snooker player must observe two conditions:

- The collision must be head-on. (If one ball strikes a glancing blow on the side of the other, they will both move off at different angles.)
- The moving ball must not be given any spin. (Spin is an added complication which we will ignore in our present study, although it plays a vital part in the games of pool and snooker.)

You can mimic the collision of two snooker balls in the laboratory using two identical trolleys, as shown in Figure 6.3b. The moving trolley has its spring-load released, so that the collision is springy. As one trolley runs into the

other, the spring is at first compressed, and then it pushes out again to set the second trolley moving. The first trolley comes to a complete halt. The 'motion' of one trolley has been transferred to the other.

You can see another interesting result if two moving identical trolleys collide head-on. If the collision is springy, both trolleys bounce backwards. If a fast-moving trolley collides with a slower one, the fast trolley bounces back at the speed of the slow one, and the slow one bounces back at the speed of the fast one. In this collision, it is as if the velocities of the trolleys have been swapped.

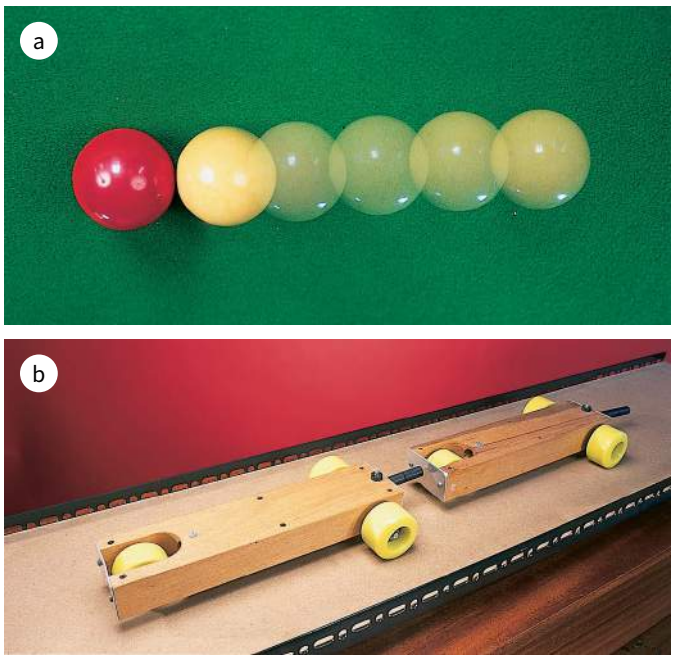


Figure 6.3 **a** The red snooker ball, coming from the left, has hit the yellow ball head-on. **b** You can do the same thing with two trolleys in the laboratory.

Sticky collisions

Figure 6.4 shows another type of collision. In this case, the trolleys have adhesive pads so that they stick together when they collide. A sticky collision like this is the opposite of a springy collision like the ones described above.

If a single moving trolley collides with an identical stationary one, they both move off together. After the collision, the speed of the combined trolleys is half that of the original trolley. It is as if the 'motion' of the original trolley has been shared between the two. If a single moving trolley collides with a stationary double trolley (twice the mass), they move off with one-third of the original velocity.

From these examples of sticky collisions, you can see that, when the mass of the trolley increases as a result of a collision, its velocity decreases. Doubling the mass halves the velocity, and so on.

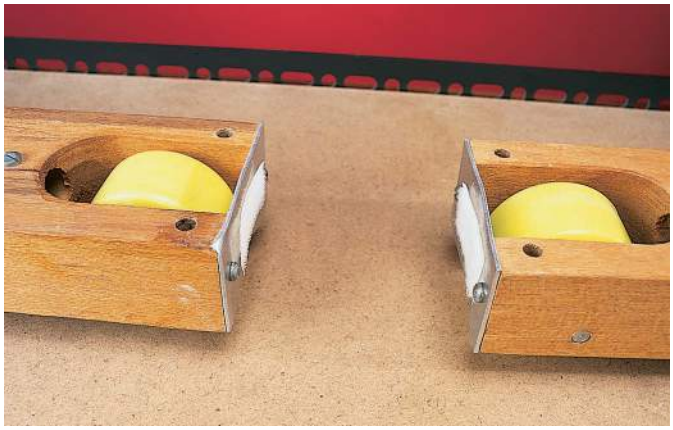


Figure 6.4 If a moving trolley sticks to a stationary trolley, they both move off together.

QUESTION

1 Here are two collisions to picture in your mind. Answer the question for each.

a Ball A, moving towards the right, collides with stationary ball B. Ball A bounces back; B moves off slowly to the right. Which has the greater mass, A or B?

b Trolley A, moving towards the right, collides with stationary trolley B. They stick together, and move off at less than half A's original speed. Which has the greater mass, A or B?

Defining linear momentum

From the examples discussed above, we can see that two quantities are important in understanding collisions:

- the mass m of the object
- the velocity v of the object.

These are combined to give a single quantity, called the **linear momentum** (or simply momentum) p of an object. The momentum of an object is defined as the product of the mass of the object and its velocity. Hence:

$$\text{momentum} = \text{mass} \times \text{velocity}$$
$$p = mv$$

The unit of momentum is kg m s^{-1}. There is no special name for this unit in the SI system.

Momentum is a vector quantity because it is a product of a vector quantity (velocity) and a scalar quantity (mass). Momentum has both magnitude and direction. Its direction is the same as the direction of the object's velocity.

In the earlier examples, we described how the 'motion' of one trolley appeared to be transferred to a second trolley, or shared with it. It is more correct to say that it is the trolley's momentum that is transferred or shared. (Strictly speaking, we should refer to linear momentum, because there is another quantity called **angular momentum** which is possessed by spinning objects.)

As with energy, we find that momentum is also conserved. We have to consider objects which form a **closed system** – that is, no external force acts on them. The principle of **conservation of momentum** states that:

> Within a closed system, the total momentum in any direction is constant.

The principle of conservation of momentum can also be expressed as follows:

> For a closed system, in any direction:
> total momentum of objects before collision
> = total momentum of objects after collision

A group of colliding objects always has as much momentum after the collision as it had before the collision. This principle is illustrated in Worked example 1.

QUESTIONS

2 Calculate the momentum of each of the following objects:

a a 0.50 kg stone travelling at a velocity of 20 m s^{-1}

b a 25 000 kg bus travelling at 20 m s^{-1} on a road

c an electron travelling at 2.0×10^{7} m s^{-1}.
(The mass of the electron is 9.1×10^{-31} kg.)

3 Two balls, each of mass 0.50 kg, collide as shown in Figure 6.6. Show that their total momentum before the collision is equal to their total momentum after the collision.

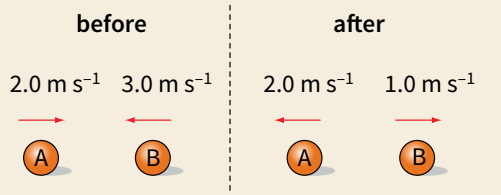


Figure 6.6 For Question 3.

WORKED EXAMPLE

1 In Figure 6.5, trolley A of mass 0.80 kg travelling at a velocity of 3.0 m s^{-1} collides head-on with a stationary trolley B. Trolley B has twice the mass of trolley A. The trolleys stick together and have a common velocity of 1.0 m s^{-1} after the collision. Show that momentum is conserved in this collision.

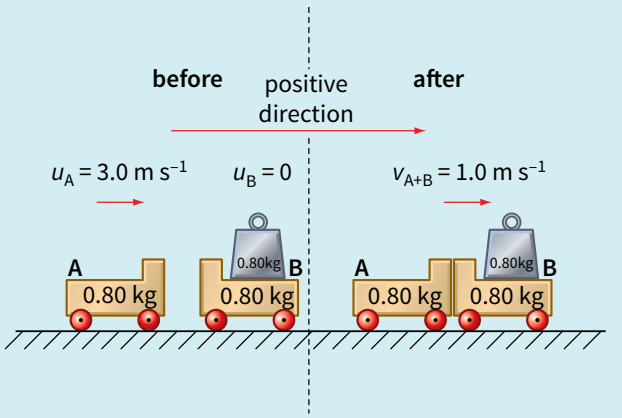


Figure 6.5 The state of trolleys A and B, before and after the collision.

Step 1 Make a sketch using the information given in the question. Notice that we need two diagrams to show the situations, one before and one after the collision. Similarly, we need two calculations – one for the momentum of the trolleys before the collision and one for their momentum after the collision.

Step 2 Calculate the momentum before the collision:

momentum of trolleys before collision

$= m_A \times u_A + m_B \times u_B$

$= (0.80 \times 3.0) + 0$

$= 2.4\ \text{kg m s}^{-1}$

Trolley B has no momentum before the collision, because it is not moving.

Step 3 Calculate the momentum after the collision:

momentum of trolleys after collision

$= (m_A + m_B) \times v_{A+B}$

$= (0.80 + 1.60) \times 1.0$

$= 2.4\ \text{kg m s}^{-1}$

So, both before and after the collision, the trolleys have a combined momentum of 2.4 kg m s^{-1}. Momentum has been conserved.

Understanding collisions

The cars in Figure 6.7 have been badly damaged by a collision. The front of a car is designed to absorb the impact of the crash. It has a 'crumple zone', which collapses on impact. This absorbs most of the kinetic energy that the car had before the collision. It is better that the car's kinetic energy should be transferred to the crumple zone than to the driver and passengers.

Motor manufacturers make use of test labs to investigate how their cars respond to impacts. When a car is designed, the manufacturers combine soft, compressible materials that absorb energy with rigid structures that protect the car's occupants. Old-fashioned cars had much more rigid structures. In a collision, they were more likely to bounce back and the violent forces involved were much more likely to prove fatal.



Figure 6.7 The front of each car has crumpled in, as a result of a head-on collision.

Two types of collision

When two objects collide, they may crumple and deform. Their kinetic energy may also disappear completely as they come to a halt. This is an example of an inelastic collision. Alternatively, they may spring apart, retaining all of their kinetic energy. This is a perfectly elastic collision. In practice, in most collisions, some kinetic energy is transformed into other forms (e.g. heat or sound) and the collision is inelastic. Previously we described the collisions as being 'springy' or 'sticky'. We should now use the correct scientific terms, **perfectly elastic** and **inelastic**.

We will look at examples of these two types of collision and consider what happens to linear momentum and kinetic energy in each.

A perfectly elastic collision

Two identical objects A and B, moving at the same speed but in opposite directions, have a head-on collision, as shown in Figure 6.8. Each object bounces back with its velocity reversed. This is a perfectly elastic collision.

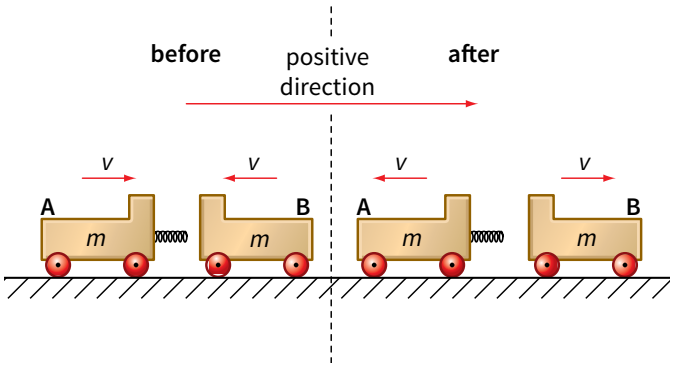


Figure 6.8 Two objects may collide in different ways: this is an elastic collision. An inelastic collision of the same two objects is shown in Figure 6.9.

You should be able to see that, in this collision, both momentum and kinetic energy are conserved. Before the collision, object A of mass m is moving to the right at speed v and object B of mass m is moving to the left at speed v. Afterwards, we still have two masses m moving with speed v, but now object A is moving to the left and object B is moving to the right. We can express this mathematically as follows:

Before the collision

object A: mass = m velocity = v momentum = mv
object B: mass = m velocity = $-v$ momentum = $-mv$

Object B has negative velocity and momentum because it is travelling in the opposite direction to object A. Therefore we have:

total momentum before collision

= momentum of A + momentum of B

= $mv + (-mv) = 0$

total kinetic energy before collision

= k.e. of A + k.e. of B

= $\frac{1}{2}mv^2 + \frac{1}{2}mv^2 = mv^2$

The magnitude of the momentum of each object is the same. Momentum is a vector quantity and we have to consider the directions in which the objects travel. The combined momentum is zero. On the other hand, kinetic energy is a scalar quantity and direction of travel is irrelevant. Both objects have the same kinetic energy and therefore the combined kinetic energy is twice the kinetic energy of a single object.

After the collision

Both objects have their velocities reversed, and we have:

total momentum after collision = $(-mv) + mv = 0$

total kinetic energy after collision = $\frac{1}{2}mv^2 + \frac{1}{2}mv^2 = mv^2$

So the total momentum and the total kinetic energy are unchanged. They are both conserved in a perfectly elastic collision such as this.

In this collision, the objects have a **relative speed** of $2v$ before the collision. After their collision, their velocities are reversed so their relative speed is $2v$ again. This is a feature of perfectly elastic collisions.

The relative speed of approach is the speed of one object measured relative to another. If two objects are travelling directly towards each other with speed v, as measured by someone stationary on the ground, then each object 'sees' the other one approaching with a speed of $2v$. Thus if objects are travelling in opposite directions we add their speeds to find the relative speed. If the objects are travelling in the same direction then we subtract their speeds to find the relative speed.

In a perfectly elastic collision,
relative speed of approach = relative speed of separation.

An inelastic collision

In Figure 6.9, the same two objects collide, but this time they stick together after the collision and come to a halt. Clearly, the total momentum and the total kinetic energy are both zero after the collision, since neither mass is moving. We have:

	Before collision	After collision
momentum	0	0
kinetic energy	$\frac{1}{2}mv^2$	0

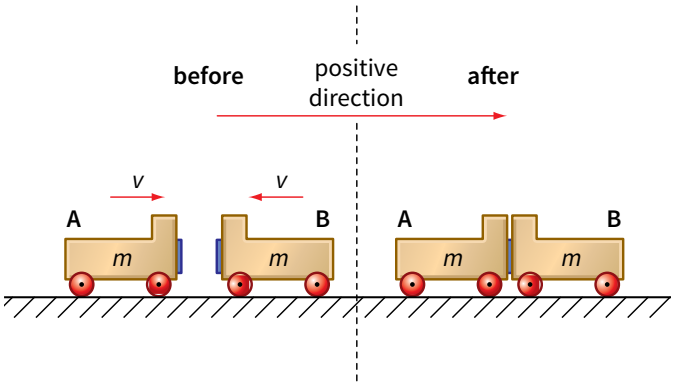


Figure 6.9 An inelastic collision between two identical objects. The trolleys are stationary after the collision.

Again we see that momentum is conserved. However, kinetic energy is not conserved. It is lost because work is done in deforming the two objects.

In fact, **momentum is always conserved** in all collisions. There is nothing else into which momentum can be converted. Kinetic energy is usually not conserved in a collision, because it can be transformed into other forms of energy – sound energy if the collision is noisy, and the energy involved in deforming the objects (which usually ends up as internal energy – they get warmer). Of course, the total amount of energy remains constant, as prescribed by the principle of conservation of energy.

QUESTION

4 Copy Table 6.1 below, choosing the correct words from each pair.

Type of collision	perfectly elastic	inelastic
Momentum	conserved / not conserved	conserved / not conserved
Kinetic energy	conserved / not conserved	conserved / not conserved
Total energy	conserved / not conserved	conserved / not conserved

Table 6.1 For Question 4.

Solving collision problems

We can use the idea of conservation of momentum to solve numerical problems, as illustrated by Worked example 2.

WORKED EXAMPLE

2 In the game of bowls, a player rolls a large ball towards a smaller, stationary ball. A large ball of mass 5.0 kg moving at 10.0 m s^{-1} strikes a stationary ball of mass 1.0 kg. The smaller ball flies off at 10.0 m s^{-1}.

a Determine the final velocity of the large ball after the impact.

b Calculate the kinetic energy 'lost' in the impact.

Step 1 Draw two diagrams, showing the situations before and after the collision. Figure 6.10 shows the values of masses and velocities; since we don't know the velocity of the large ball after the collision, this is shown as *v*. The direction from left to right has been assigned the 'positive' direction.

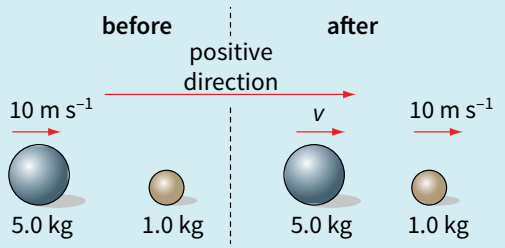


Figure 6.10 When solving problems involving collisions, it is useful to draw diagrams showing the situations before and after the collision. Include the values of all the quantities that you know.

Step 2 Using the principle of conservation of momentum, set up an equation and solve for the value of *v*:

total momentum before collision = total momentum after collision

$(5.0 \times 10) + (1.0 \times 0) = (5.0 \times v) + (1.0 \times 10)$

$50 + 0 = 5.0v + 10$

$$v = \frac{40}{5.0} = 8.0 \text{ m s}^{-1}$$

So the speed of the large ball decreases to 8.0 m s^{-1} after the collision. Its direction of motion is unchanged – the velocity remains positive.

Step 3 Knowing the large ball's final velocity, calculate the change in kinetic energy during the collision:

total k.e. before collision = $\frac{1}{2} \times 5.0 \times 10^2 + 0 = 250$ J

total k.e. after collision = $\frac{1}{2} \times 5.0 \times 8.0^2 + \frac{1}{2} \times 1.0 \times 10^2$

= 210 J

k.e. 'lost' in the collision = 250 J − 210 J = 40 J

This lost kinetic energy will appear as internal energy (the two balls get warmer) and as sound energy (we hear the collision between the balls).

QUESTIONS

5 Figure 6.11 shows two identical balls A and B about to make a head-on collision. After the collision, ball A rebounds at a speed of 1.5 m s^{-1} and ball B rebounds at a speed of 2.5 m s^{-1}. The mass of each ball is 4.0 kg.

a Calculate the momentum of each ball before the collision.

b Calculate the momentum of each ball after the collision.

c Is the momentum conserved in the collision?

d Show that the total kinetic energy of the two balls is conserved in the collision.

e Show that the relative speed of the balls is the same before and after the collision.

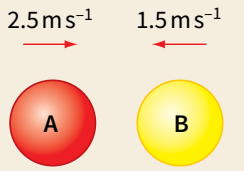


Figure 6.11 For Question 5.

6 A trolley of mass 1.0 kg is moving at 2.0 m s^{-1}. It collides with a stationary trolley of mass 2.0 kg. This second trolley moves off at 1.2 m s^{-1}.

a Draw 'before' and 'after' diagrams to show the situation.

b Use the principle of conservation of momentum to calculate the speed of the first trolley after the collision. In what direction does it move?

Explosions and crash-landings

There are situations where it may appear that momentum is being created out of nothing, or that it is disappearing without trace. Do these contradict the principle of conservation of momentum?

The rockets shown in Figure 6.12 rise high into the sky. As they start to fall, they send out showers of chemical packages, each of which explodes to produce a brilliant sphere of burning chemicals. Material flies out in all directions to create a spectacular effect.

Does an explosion create momentum out of nothing? The important point to note here is that the burning material spreads out equally in all directions. Each tiny spark has momentum, but for every spark, there is another moving in the opposite direction, i.e. with opposite momentum. Since momentum is a vector quantity, the total amount of momentum created is zero.



Figure 6.12 These exploding rockets produce a spectacular display of bright sparks in the night sky.

At the same time, kinetic energy is created in an explosion. Burning material flies outwards; its kinetic energy has come from the chemical potential energy stored in the chemical materials before they burn.

More fireworks

A roman candle fires a jet of burning material up into the sky. This is another type of explosion, but it doesn't send material in all directions. The firework tube directs the material upwards. Has momentum been created out of nothing here?

Again, the answer is no. The chemicals have momentum upwards, but at the same time, the roman candle pushes downwards on the Earth. An equal amount of downwards momentum is given to the Earth. Of course, the Earth is massive, and we don't notice the tiny change in its velocity which results.

Down to Earth

If you push a large rock over a cliff, its speed increases as it falls. Where does its momentum come from? And when it lands, where does its momentum disappear to?

The rock falls because of the pull of the Earth's gravity on it. This force is its weight and it makes the rock accelerate towards the Earth. Its weight does work and the rock gains kinetic energy. It gains momentum downwards. Something must be gaining an equal amount of momentum in the opposite (upward) direction. It is the Earth, which starts to move upwards as the rock falls downwards. The mass of the Earth is so great that its change in velocity is small – far too small to be noticeable.

When the rock hits the ground, its momentum becomes zero. At the same instant, the Earth also stops moving upwards. The rock's momentum cancels out the Earth's momentum. At all times during the rock's fall and crash-landing, momentum has been conserved.

If a rock of mass 60 kg is falling towards the Earth at a speed of 20 m s^{-1}, how fast is the Earth moving towards it? Figure 6.13 shows the situation. The mass of the Earth is 6.0×10^{24} kg. We have:

total momentum of Earth and rock = 0

Hence:

$$(60 \times 20) + (6.0 \times 10^{24} \times v) = 0$$

$$v = -2.0 \times 10^{-22}\ \text{m s}^{-1}$$

The minus sign shows that the Earth's velocity is in the opposite direction to that of the rock. The Earth moves very slowly indeed. In the time of the rock's fall, it will move much less than the diameter of the nucleus of an atom!

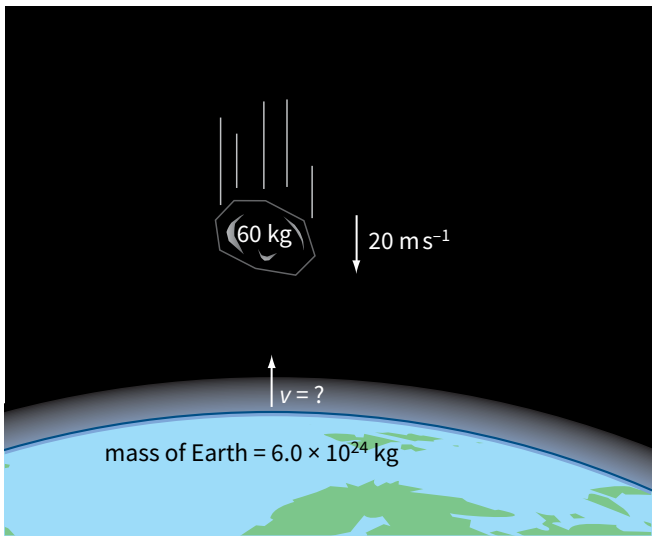


Figure 6.13 The rock and Earth gain momentum in opposite directions.

QUESTIONS

7 Discuss whether momentum is conserved in each of the following situations.
   a A star explodes in all directions – a supernova.
   b You jump up from a trampoline. As you go up, your speed decreases; as you come down again, your speed increases.

8 A ball of mass 0.40 kg is thrown at a wall. It strikes the wall with a speed of 1.5 m s^{-1} perpendicular to the wall and bounces off the wall with a speed of 1.2 m s^{-1}. Explain the changes in momentum and energy which happen in the collision between the ball and the wall. Give numerical values where possible.

Collisions in two dimensions

It is rare that collisions happen in a straight line – in one dimension. Figure 6.14 shows a two-dimensional collision between two snooker balls. From the multiple images, we can see how the velocities of the two balls change:

- At first, the white ball is moving straight forwards. When it hits the red ball, it moves off to the right. Its speed decreases; we can see this because the images get closer together.
- The red ball moves off to the left. It moves off at a bigger angle than the white ball, but more slowly – the images are even closer together.

How can we understand what happens in this collision, using the ideas of momentum and kinetic energy?

At first, only the white ball has momentum, and this is in the forward direction. During the collision, this momentum is shared between the two balls. We can see this because each has a component of velocity in the forward direction.

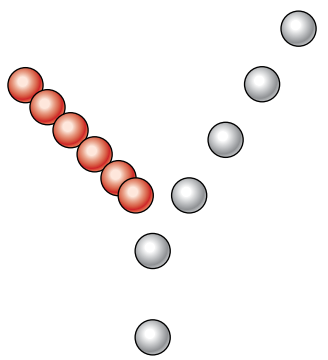


Figure 6.14 The white ball strikes the red ball a glancing blow. The two balls move off in different directions.

At the same time, each ball gains momentum in the sideways direction, because each has a sideways component of velocity – the white ball to the right, and the red ball to the left. These must be equal in magnitude and opposite in direction, otherwise we would conclude that momentum had been created out of nothing. The red ball moves at a greater angle, but its velocity is less than that of the white ball, so that the component of its velocity at right angles to the original track is the same as the white ball's.

Figure 6.15a shows the momentum of each ball before and after the collision. We can draw a vector triangle to represent the changes of momentum in this collision (Figure 6.15b). The two momentum vectors after the collision add up to equal the momentum of the white ball before the collision. The vectors form a closed triangle because momentum is conserved in this two-dimensional collision.

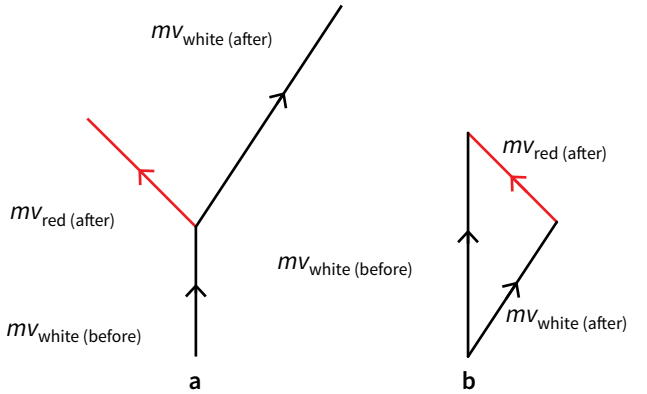


Figure 6.15 **a** These vectors represent the momenta of the colliding balls shown in Figure 6.14. **b** The closed vector triangle shows that momentum is conserved in the collision.

Components of momentum

Momentum is a vector quantity and so we can split it into components in order to solve problems.

Worked example 3 shows how to find an unknown velocity.

Worked example 4 shows how to demonstrate that momentum has been conserved in a two-dimensional collision.

WORKED EXAMPLES

3 A white ball of mass $m = 1.0\,\text{kg}$ and moving with initial speed $u = 0.5\,\text{m s}^{-1}$ collides with a stationary red ball of the same mass. They move off so that each has the same speed and the angle between their paths is 90°. What is their speed?

Step 1 Draw a diagram to show the velocity vectors of the two balls, before and after the collision (Figure 6.16). We will show the white ball initially travelling along the y-direction.

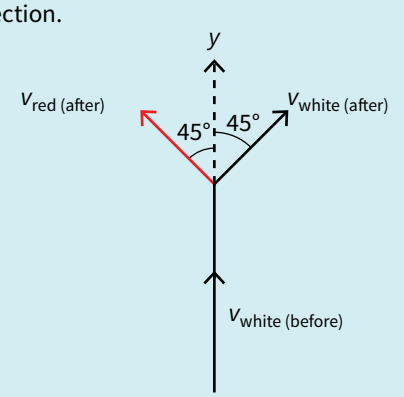


Figure 6.16 Velocity vectors for the white and red balls.

Because we know that the two balls have the same final speed v, their paths must be symmetrical about the y-direction. Since their paths are at 90° to one other, each must be at 45° to the y-direction.

Step 2 We know that momentum is conserved in the y-direction. Hence we can say:

initial momentum of white ball in y-direction
= final component of momentum of white ball in y-direction
+ final component of momentum of red ball in y-direction

This is easier to understand using symbols:

$mu = mv_y + mv_y$

where v_y is the component of v in the y-direction. The right-hand side of this equation has two identical terms, one for the white ball and one for the red. We can simplify the equation to give:

$mu = 2mv_y$

Step 3 The component of v in the y-direction is $v\cos 45°$. Substituting this, and including values of m and u, gives

$0.5 = 2v\cos 45°$

and hence

$$v = \frac{0.5}{2\cos 45°} \approx 0.354\,\text{m s}^{-1}$$

So each ball moves off at $0.354\,\text{m s}^{-1}$ at an angle of 45° to the initial direction of the white ball.

4 Figure 6.17 shows the momentum vectors for particles 1 and 2, before and after a collision. Show that momentum is conserved in this collision.

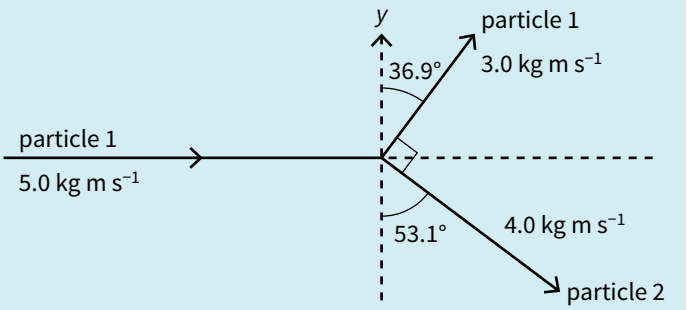


Figure 6.17 Momentum vectors: particle 1 has come from the left and collided with particle 2.

Step 1 Consider momentum changes in the y-direction.

Before collision:

momentum = 0

(because particle 1 is moving in the x-direction and particle 2 is stationary).

After collision:

component of momentum of particle 1
$= 3.0\cos 36.9° \approx 2.40\,\text{kg m s}^{-1}$ upwards

component of momentum of particle 2
$= 4.0\cos 53.1° \approx 2.40\,\text{kg m s}^{-1}$ downwards

These components are equal and opposite and hence their sum is zero. Hence momentum is conserved in the y-direction.

Step 2 Consider momentum changes in the x-direction.

Before collision: momentum $= 5.0\,\text{kg m s}^{-1}$ to the right

After collision:

component of momentum of particle 1
$= 3.0\cos 53.1° \approx 1.80\,\text{kg m s}^{-1}$ to the right

component of momentum of particle 2
$= 4.0\cos 36.9° \approx 3.20\,\text{kg m s}^{-1}$ to the right

total momentum to the right $= 5.0\,\text{kg m s}^{-1}$

Hence momentum is conserved in the x-direction.

Step 3 An alternative approach would be to draw a vector triangle similar to Figure 6.15b. In this case, the numbers have been chosen to make this easy; the vectors form a 3–4–5 right-angled triangle.

Because the vectors form a closed triangle, we can conclude that:

momentum before collision = momentum after collision
i.e. momentum is conserved.

QUESTIONS

9 A snooker ball strikes a stationary ball. The second ball moves off sideways, at 60° to the initial path of the first ball.

Use the idea of conservation of momentum to explain why the first ball cannot travel in its initial direction after the collision. Illustrate your answer with a diagram.

10 Look back to Worked example 4 above. Draw the vector triangle which shows that momentum is conserved in the collision described in the question. Show the value of each angle in the triangle.

11 Figure 6.18 shows the momentum vectors for two particles, 1 and 2, before and after a collision. Show that momentum is conserved in this collision.

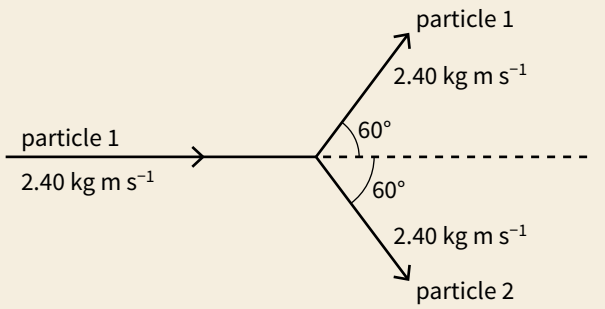


Figure 6.18 For Question 11.

12 A snooker ball collides with a second identical ball as shown in Figure 6.19.

a Determine the components of the velocity of the first ball in the *x*- and *y*-directions.

b Hence determine the components of the velocity of the second ball in the *x*- and *y*-directions.

c Hence determine the velocity (magnitude and direction) of the second ball.

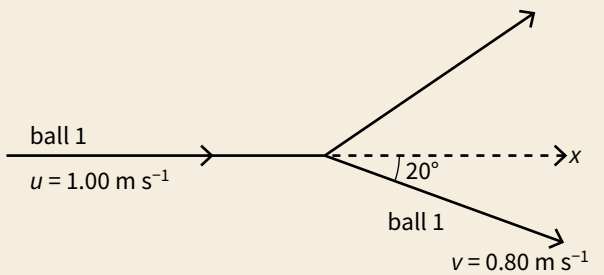


Figure 6.19 For Question 12.

Momentum and Newton's laws

The big ideas of physics are often very simple; that is to say, it takes only a few words to express them and they can be applied in many situations. However, 'simple' does not usually mean 'easy'. Concepts such as force, energy and voltage, for example, are not immediately obvious. They usually took someone to make a giant leap of imagination to first establish them. Then the community of physicists spent decades worrying away at them, refining them until they are the fundamental ideas which we use today.

Take Isaac Newton's work on motion. He published his ideas in a book commonly known as the *Principia* (see Figure 6.20); its full title translated from Latin is *Mathematical Principles of Natural Philosophy*.

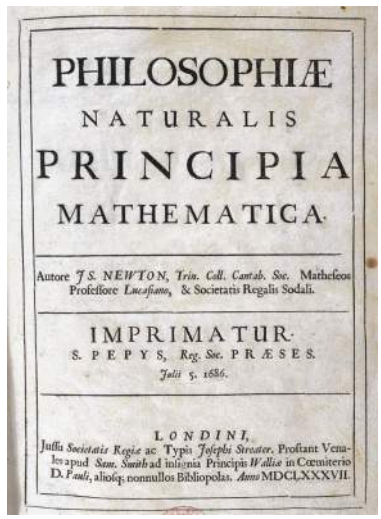

PHILOSOPHIÆ
NATURALIS
PRINCIPIA
MATHEMATICA
Autore *J.S. NEWTON, Trin. Coll. Cantab. Soc.* Matheseos Professore *Lucasiano*, & Societatis Regalis Sodali.
IMPRIMATUR.
S. PEPYS, *Reg. Soc.* PRÆSES.
Julii 5. 1686.
LONDINI,
Justu Societatis Regiæ ac Typis Josephi Streater. Prostant Venales apud *Sam. Smith* ad insignia Principis *Walliæ* in Cœmiterio D. *Pauli*, aliosq; nonnullos Bibliopolas. *Anno* MDCLXXXVII.

Figure 6.20 The title page of Newton's *Principia*, in which he outlined his theories of the laws that govern the motion of objects.

The *Principia* represents the results of 20 years of thinking. Newton was able to build on Galileo's ideas and he was in correspondence with many other scientists and mathematicians. Indeed, there was an ongoing feud with Robert Hooke as to who was the first to come up with certain ideas. Among scientists, this is known as 'priority', and publication is usually taken as proof of priority.

Newton wanted to develop an understanding of the idea of 'force'. You may have been told in your early studies of science that 'a force is a push or a pull'. That doesn't tell us very much. Newton's idea was that forces are interactions between bodies and that they change the motion of the body that they act on. Forces acting on an object can produce acceleration. For an object of constant mass, this acceleration is directly proportional to the resultant force acting on the object. That is much more like a scientific definition of force.

Understanding motion

In Chapter 3, we looked at Newton's laws of motion. We can get further insight into these laws by thinking about them in terms of momentum.

Newton's first law of motion

In everyday speech, we sometimes say that something has momentum when we mean that it has a tendency to keep on moving of its own free will. An oil tanker is difficult to stop at sea, because of its momentum. We use the same word in a figurative sense: 'The election campaign is gaining momentum.' This idea of keeping on moving is just what we discussed in connection with **Newton's first law of motion**:

> An object will remain at rest or keep travelling at constant velocity unless it is acted on by a resultant force.

An object travelling at constant velocity has constant momentum. Hence the first law is really saying that the momentum of an object remains the same unless the object experiences an external force.

Newton's second law of motion

Newton's second law of motion links the idea of the resultant force acting on an object and its momentum. A statement of Newton's second law is:

> The resultant force acting on an object is directly proportional to the rate of change of the linear momentum of that object. The resultant force and the change in momentum are in the same direction.

Hence:

resultant force ∝ rate of change of momentum

This can be written as:

$$F \propto \frac{\Delta p}{\Delta t}$$

where F is the resultant force and Δp is the change in momentum taking place in a time interval of Δt. (Remember that the Greek letter delta, Δ, is a shorthand for 'change in …', so Δp means 'change in momentum'.) The changes in momentum and force are both vector quantities, hence these two quantities must be in the same direction.

The unit of force (the newton, N) is defined to make the constant of proportionality equal to one, so we can write the second law of motion mathematically as:

$$F = \frac{\Delta p}{\Delta t}$$

If the forces acting on an object are balanced, there is no resultant force and the object's momentum will remain constant. If a resultant force acts on an object, its momentum (velocity and/or direction) will change. The equation above gives us another way of stating Newton's second law of motion:

> The resultant force acting on an object is equal to the rate of change of its momentum. The resultant force and the change in momentum are in the same direction.

This statement effectively defines what we mean by a force; it is an interaction that causes an object's momentum to change. So, if an object's momentum is changing, there must be a force acting on it. We can find the size and direction of the force by measuring the rate of change of the object's momentum:

> force = rate of change of momentum
>
> $$F = \frac{\Delta p}{\Delta t}$$

Worked example 5 shows how to use this equation.

WORKED EXAMPLE

5 Calculate the average force acting on a 900 kg car when its velocity changes from 5.0 m s^{-1} to 30 m s^{-1} in a time of 12 s.

Step 1 Write down the quantities given:

$m = 900$ kg

initial velocity $u = 5.0$ m s^{-1}

$\Delta t = 12$ s

Step 2 Calculate the initial momentum and the final momentum of the car:

momentum = mass × velocity

initial momentum = $mu = 900 \times 5.0 = 4500$ kg m s^{-1}

final momentum = $mv = 900 \times 30 = 27\,000$ kg m s^{-1}

Step 3 Use Newton's second law of motion to calculate the average force on the car:

$$F = \frac{\Delta p}{\Delta t}$$

$$= \frac{27\,500 - 4500}{12}$$

$$= 1875 \text{ N} \approx 1900 \text{ N}$$

The average force acting on the car is about 1.9 kN.

A special case of Newton's second law of motion

Imagine an object of constant mass m acted upon by a resultant force F. The force will change the momentum of the object. According to Newton's second law of motion, we have:

$$F = \frac{\Delta p}{\Delta t} = \frac{mv - mu}{t}$$

where u is the initial velocity of the object, v is the final velocity of the object and t is the time taken for the change in velocity. The mass m of the object is a constant; hence the above equation can be rewritten as:

$$F = \frac{m(v-u)}{t} = m\left(\frac{v-u}{t}\right)$$

The term in brackets on the right-hand side is the acceleration a of the object. Therefore a special case of Newton's second law is:

$$F = ma$$

We have already met this equation in Chapter 3. In Worked example 5, you could have determined the average force acting on the car using this simplified equation for Newton's second law of motion. Remember that the equation $F = ma$ is a special case of $F = \frac{\Delta p}{\Delta t}$ which only applies when the mass of the object is constant. There are situations where the mass of an object changes as it moves, for example a rocket, which burns a phenomenal amount of chemical fuel as it accelerates upwards.

Newton's third law of motion

Newton's third law of motion is about interacting objects. These could be two magnets attracting or repelling each other, two electrons repelling each other, etc. Newton's third law states:

> When two bodies interact, the forces they exert on each other are equal and opposite.

How can we relate this to the idea of momentum? Picture holding two magnets, one in each hand. You gradually bring them towards each other (Figure 6.21) so that they start to attract each other. Each feels a force pulling it towards the other. The two forces are the same size, even if one magnet is stronger than the other. Indeed, one magnet could be replaced by an unmagnetised piece of steel and they would still attract each other equally.

If you release the magnets, they will gain momentum as they are pulled towards each other. One gains momentum to the left while the other gains equal momentum to the right.

Each is acted on by the same force, and for the same time. Hence momentum is conserved.

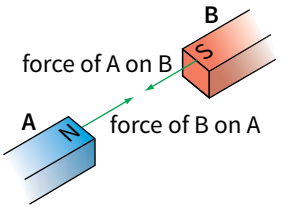


Figure 6.21 Newton's third law states that the forces these two magnets exert one each other must be equal and opposite.

QUESTIONS

13 A car of mass 1000 kg is travelling at a velocity of 10 m s^{-1}. It accelerates for 15 s, reaching a velocity of 24 m s^{-1}. Calculate:

- **a** the change in the momentum of the car in the 15 s period
- **b** the average force acting on the car as it accelerates.

14 A ball is kicked by a footballer. The average force on the ball is 240 N and the impact lasts for a time interval of 0.25 s.

- **a** Calculate the change in the ball's momentum.
- **b** State the direction of the change in momentum.

15 Water pouring from a broken pipe lands on a flat roof. The water is moving at 5.0 m s^{-1} when it strikes the roof. The water hits the roof at a rate of 10 kg s^{-1}. Calculate the force of the water hitting the roof. (Assume that the water does not bounce as it hits the roof. If it did bounce, would your answer be greater or smaller?)

16 A golf ball has a mass of 0.046 kg. The final velocity of the ball after being struck by a golf club is 50 m s^{-1}. The golf club is in contact with the ball for a time of 1.3 ms. Calculate the average force exerted by the golf club on the ball.

Summary

- Linear momentum is the product of mass and velocity:

  momentum = mass × velocity or $p = mv$
- The principle of conservation of momentum:

  For a closed system, in any direction the total momentum before an interaction (e.g. collision) is equal to the total momentum after the interaction.
- In all interactions or collisions, momentum and total energy are conserved.
- Kinetic energy is conserved in a perfectly elastic collision; relative speed is unchanged in a perfectly elastic collision.
- In an inelastic collision, kinetic energy is not conserved. It is transferred into other forms of energy (e.g. heat or sound). Most collisions are inelastic.
- Newton's first law of motion: An object will remain at rest or keep travelling at constant velocity unless it is acted on by a resultant force.
- Newton's second law of motion: The resultant force acting on a body is equal to the rate of change of its momentum:

  resultant force = rate of change of momentum

  or

  $F = \frac{\Delta p}{\Delta t}$
- Newton's third law of motion: When two bodies interact, the forces they exert on each other are equal and opposite.
- The equation $F = ma$ is a special case of Newton's second law of motion when mass m remains constant.

End-of-chapter questions

1 An object is dropped and its momentum increases as it falls toward the ground. Explain how the law of conservation of momentum and Newton's third law of motion can be applied to this situation. [2]

2 A ball of mass 2 kg, moving at 3.0 m s^{-1}, strikes a wall and rebounds with the same speed. State and explain whether there is a change in:

a the momentum of the ball [3]

b the kinetic energy of the ball. [1]

3 **a** Define **linear momentum**. [1]

b Determine the base units of linear momentum in the SI system. [1]

c A car of mass 900 kg starting from rest has a constant acceleration of 3.5 m s^{-2}. Calculate its momentum after it has travelled a distance of 40 m. [2]

d Figure 6.22 shows two identical objects about to make a head-on collision. The objects stick together during the collision. Determine the final speed of the objects. State the direction in which they move. [3]

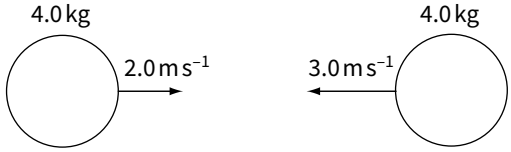


Figure 6.22 For End-of-chapter Question 3.

4 **a** Explain what is meant by an:

i elastic collision [1]

ii inelastic collision. [1]

b A snooker ball of mass 0.35 kg hits the side of a snooker table at right angles and bounces off also at right angles. Its speed before collision is 2.8 m s^{-1} and its speed after is 2.5 m s^{-1}. Calculate the change in the momentum of the ball. [2]

c Explain whether or not momentum is conserved in the situation described in **b**. [3]

5 A car of mass 1100 kg is travelling at 24 m s^{-1}. The driver applies the brakes and the car decelerates uniformly and comes to rest in 20 s.

a Calculate the change in momentum of the car. [2]

b Calculate the braking force on the car. [2]

c Determine the braking distance of the car. [2]

6 A marble of mass 100 g is moving at a speed of 0.40 m s^{-1} in the *x*-direction.

a Calculate the marble's momentum. [2]

The marble strikes a second, identical marble. Each moves off at an angle of 45° to the *x*-direction.

b Use the principle of conservation of momentum to determine the speed of each marble after the collision. [3]

c Show that kinetic energy is conserved in this collision. [2]

7 A cricket bat strikes a ball of mass 0.16 kg travelling towards it. The ball initially hits the bat at a speed of 25 m s^{-1} and returns along the same path with the same speed. The time of impact is 0.0030 s.

a Determine the change in momentum of the cricket ball. [2]

b Determine the force exerted by the bat on the ball. [2]

c Describe how the laws of conservation of energy and momentum apply to this impact and state whether the impact is elastic or inelastic. [4]

8 **a** State the principle of conservation of momentum and state the conditions under which it is valid. [2]

b An arrow of mass 0.25 kg is fired horizontally towards an apple of mass 0.10 kg which is hanging on a string (Figure 6.23).

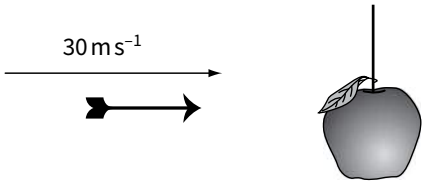


Figure 6.23 For End-of-chapter Question 8.

The horizontal velocity of the arrow as it enters the apple is 30 m s^{-1}. The apple was initially at rest and the arrow sticks in the apple.

i Calculate the horizontal velocity of the apple and arrow immediately after the impact. [2]

ii Calculate the change in momentum of the arrow during the impact. [2]

iii Calculate the change in total kinetic energy of the arrow and apple during the impact. [2]

iv An identical arrow is fired at the centre of a stationary ball of mass 0.25 kg. The collision is perfectly elastic. Describe what happens and state the relative speed of separation of the arrow and the ball. [2]

9 **a** State what is meant by:

i a perfectly elastic collision [1]

ii a completely inelastic collision. [1]

b A stationary uranium nucleus disintegrates, emitting an alpha-particle of mass 6.65×10^{-27} kg and another nucleus X of mass 3.89×10^{-25} kg (Figure 6.24).

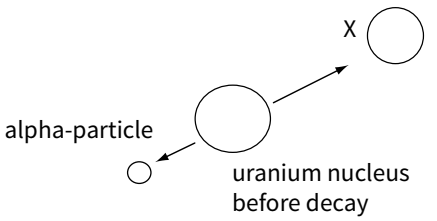


Figure 6.24 For End-of-chapter Question 9.

i Explain why the alpha-particle and nucleus X must be emitted in exactly opposite directions. [2]

ii Using the symbols v_α and v_X for velocities, write an equation for the conservation of momentum in this disintegration. [1]

iii Using your answer to **ii**, calculate the ratio v_α / v_X after the disintegration. [1]

10 **a** State **two** quantities that are conserved in an elastic collision. [1]

b A machine gun fires bullets of mass 0.014 kg at a speed of 640 m s^{-1}.

i Calculate the momentum of each bullet as it leaves the gun. [1]

ii Explain why a soldier holding the machine gun experiences a force when the gun is firing. [2]

iii The maximum steady horizontal force that a soldier can exert on the gun is 140 N. Calculate the maximum number of bullets that the gun can fire in one second. [2]

11 Two railway trucks are travelling in the same direction and collide. The mass of truck X is 2.0×10^4 kg and the mass of truck Y is 3.0×10^4 kg. Figure 6.25 shows how the velocity of each truck varies with time.

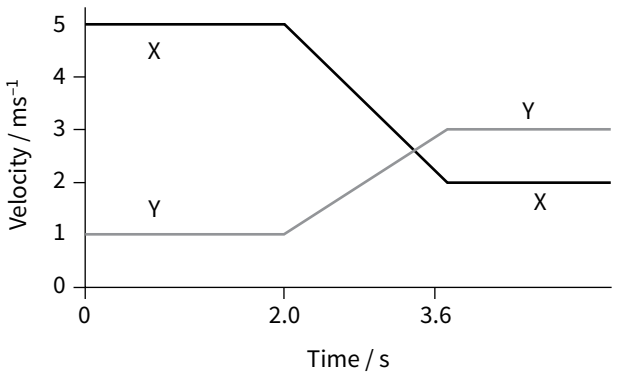


Figure 6.25 For End-of-chapter Question 11.

a Copy and complete the table. [6]

	Change in momentum / kg m s^{-1}	**Initial kinetic energy / J**	**Final kinetic energy / J**
truck X			
truck Y			

b State and explain whether the collision of the two trucks is an example of an elastic collision. [2]

c Determine the force that acts on each truck during the collision. [2]

Chapter 7: Matter and materials

Learning outcomes

You should be able to:

- define density
- define pressure and calculate the pressure in a fluid
- understand how tensile and compressive forces cause deformation
- describe the behaviour of springs and understand Hooke's law
- distinguish between elastic and plastic deformation
- define and use stress, strain and the Young modulus
- describe an experiment to measure the Young modulus

Springy stuff

In everyday life, we make great use of elastic materials. The term elastic means springy; that is, the material deforms when a force is applied and returns to its original shape when the force is removed. Rubber is an elastic material. This is obviously important for a bungee jumper (Figure 7.1). The bungee rope must have the correct degree of elasticity. The jumper must be brought gently to a halt. If the rope is too stiff, the jumper will be jerked violently so that the deceleration is greater than their body can withstand. On the other hand, if the rope is too stretchy, they may bounce up and down endlessly, or even strike the ground.

In this chapter we will look at how forces can change the shape of an object. Before that, we will look at two important quantities, density and pressure.



Figure 7.1 The stiffness and elasticity of rubber are crucial factors in bungee jumping.

Density

Density is a property of matter. It tells us about how concentrated the matter is in a particular material. Density is a constant for a given material under specific conditions.

Density is defined as the mass per unit volume of a substance:

$$\text{density} = \frac{\text{mass}}{\text{volume}}$$

$$\rho = \frac{m}{v}$$

The symbol used here for density, ρ, is the Greek letter rho.

The standard unit for density in the SI system is kg m^{-3}, but you may also find values quoted in g cm^{-3}. It is useful to remember that these units are related by:

$$1000\,\text{kg m}^{-3} = 1\,\text{g cm}^{-3}$$

and that the density of water is approximately $1000\,\text{kg m}^{-3}$.

QUESTIONS

1 A cube of copper has a mass of 240 g. Each side of the cube is 3.0 cm long. Calculate the density of copper in g cm^{-3} and in kg m^{-3}.

2 The density of steel is $7850\,\text{kg m}^{-3}$. Calculate the mass of a steel sphere of radius 0.15 m. (First calculate the volume of the sphere using the formula $V = \frac{4}{3}\pi r^3$ and then use the density equation.)

Pressure

A fluid (liquid or gas) exerts pressure on the walls of its container, or on any surface with which it is in contact. A big force on a small area produces a high pressure.

Pressure is defined as the normal force acting per unit cross-sectional area.

We can write this as a word equation:

$$\text{pressure} = \frac{\text{normal force}}{\text{cross-sectional area}}$$

$$p = \frac{F}{A}$$

Force is measured in newtons and area is measured in square metres. The units of pressure are thus newtons per square metre (N m^{-2}), which are given the special name of pascals (Pa).

$$1\,\text{Pa} = 1\,\text{N m}^{-2}$$

QUESTIONS

3 A chair stands on four feet, each of area $10\,\text{cm}^2$. The chair weighs 80 N. Calculate the pressure it exerts on the floor.

4 Estimate the pressure you exert on the floor when you stand on both feet. (You could draw a rough rectangle around both your feet placed together to find the area in contact with the floor. You will also need to calculate your weight from your mass.)

Pressure in a fluid

The pressure in a fluid (a liquid or gas) increases with depth. Divers know this: the further down they dive, the greater the water pressure acting on them. Pilots know this: the higher they fly, the lower is the pressure of the atmosphere. The atmospheric pressure we experience down here on the surface of the Earth is due to the weight of the atmosphere above us, pressing downwards. It is pulled downwards by gravity.

The pressure in a fluid depends on three factors:

- the depth h below the surface
- the density ρ of the fluid
- the acceleration due to gravity, g.

In fact, pressure p is proportional to each of these and we have:

$$\text{pressure} = \text{density} \times \text{acceleration due to gravity} \times \text{depth}$$

$$p = \rho g h$$

We can derive this relationship using Figure 7.2. The force acting on the shaded area A on the bottom of the tank is caused by the weight of water above it, pressing downwards. We can calculate this force and hence the pressure as follows:

$$\text{volume of water} = A \times h$$

$$\text{mass of water} = \text{density} \times \text{volume} = \rho \times A \times h$$

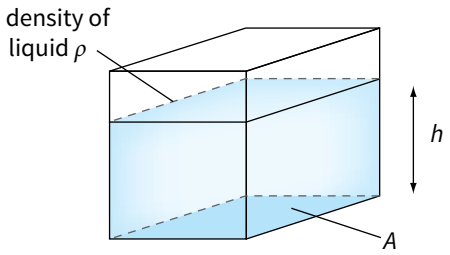


Figure 7.2 The weight of water in a tank exerts pressure on its base.

$$\text{weight of water} = \text{mass} \times g = \rho \times A \times h \times g$$

$$\text{pressure} = \frac{\text{force}}{\text{area}} = \rho \times A \times h \times \frac{g}{A}$$

$$= \rho \times g \times h$$

QUESTIONS

5 Calculate the pressure of water on the bottom of a swimming pool if the depth of water in the pool varies between 0.8 m and 2.4 m. (Density of water = 1000 kg m^{-3}.) If atmospheric pressure is 1.01×10^5 Pa, calculate the maximum total pressure at the bottom of the swimming pool.

6 Estimate the height of the atmosphere if atmospheric density at the Earth's surface is 1.29 kg m^{-3}. (Atmospheric pressure = 101 kPa.)

WORKED EXAMPLES

1 A cube of side 0.20 m floats in water with 0.15 m below the surface of the water. The density of water is 1000 kg m^{-3}. Calculate the pressure of the water acting on the bottom surface of the cube and the force upwards on the cube caused by this pressure. (This force is the upthrust on the cube.)

Step 1 Use the equation for pressure:

$p = \rho \times g \times h = 1000 \times 9.81 \times 0.15 = 1470$ Pa

Step 2 Calculate the area of the base of the cube, and use this area in the equation for force.

area of base of cube = $0.2 \times 0.2 = 0.04$ m^2

force = pressure × area = $1470 \times 0.04 = 58.8$ N

2 Figure 7.3 shows a manometer used to measure the pressure of a gas supply. Calculate the pressure difference between the gas inside the pipe and atmospheric pressure.

Step 1 Determine the difference in height h of the water on the two sides of the manometer.

$h = 60 - 30 = 30$ cm

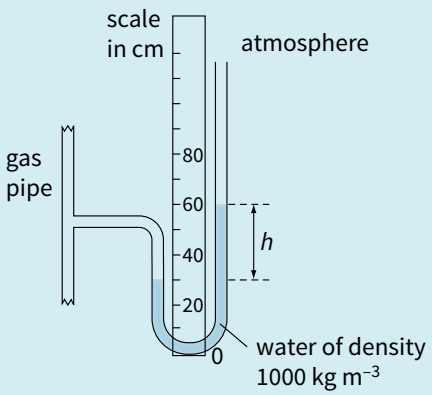


Figure 7.3 For Worked example 2.

Step 2 Because the level of water on the side of the tube next to the gas pipe is lower than on the side open to the atmosphere, the pressure in the gas pipe is above atmospheric pressure.

pressure difference = $\rho \times g \times h = 1000 \times 9.81 \times 0.30 = 2940$ Pa

Compressive and tensile forces

A pair of forces is needed to change the shape of a spring. If the spring is being squashed and shortened, we say that the forces are **compressive**. More usually, we are concerned with stretching a spring, in which case the forces are described as **tensile** (Figure 7.4).

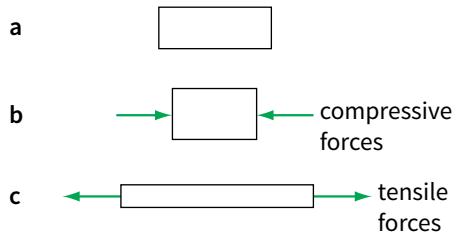


Figure 7.4 The effects of compressive and tensile forces.

When a wire is bent, some parts become longer and are in tension while other parts become shorter and are in compression. Figure 7.5 shows that the line AA becomes longer when the wire is bent and the line BB becomes shorter. The thicker the wire, the greater the compression and tension forces along its edges.

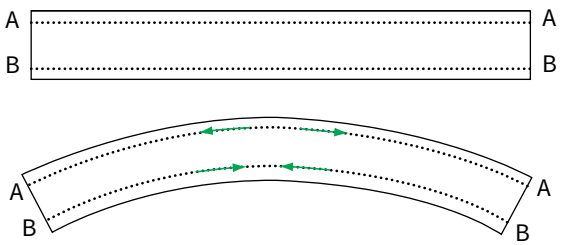


Figure 7.5 Bending a straight wire or beam results in tensile forces along the upper surface (the outside of the bend) and compressive forces on the inside of the bend.

It is simple to investigate how the length of a helical spring is affected by the applied force or load. The spring hangs freely with the top end clamped firmly (Figure 7.6). A load is added and gradually increased. For each value of the load, the extension of the spring is measured. Note that it is important to determine the increase in length of the spring, which we call the **extension**. We can plot a graph of force against extension to find the stiffness of the spring, as shown in Figure 7.7.

Hooke's law

The conventional way of plotting the results would be to have the force along the horizontal axis and the extension along the vertical axis. This is because we are changing the force (the independent variable) and this results in a change in the extension (the dependent variable). The graph shown in Figure 7.7 has extension on the horizontal

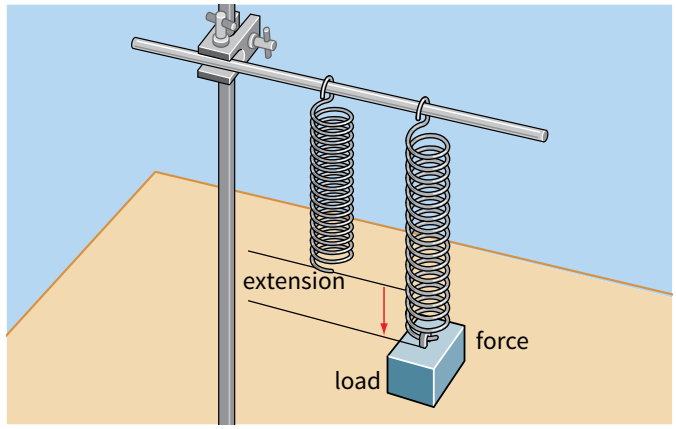


Figure 7.6 Stretching a spring.

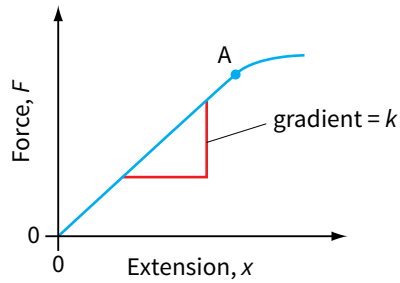


Figure 7.7 Force–extension graph for a spring.

axis and force on the vertical axis. This is a departure from the convention because the **gradient** of the straight section of this graph turns out to be an important quantity, known as the **force constant** of the spring. For a typical spring, the first section of this graph OA is a straight line passing through the origin. The extension x is directly proportional to the applied force (load) F. The behaviour of the spring in the linear region OA of the graph can be expressed by the following equation:

$$x \propto F$$

or $$F = kx$$

where k is the force constant of the spring (sometimes called either the stiffness or the spring constant of the spring). The force constant is the force per unit extension. The force constant k of the spring is given by the equation:

$$k = \frac{F}{x}$$

The SI unit for the force constant is newtons per metre or N m^{-1}. We can find the force constant k from the gradient of section OA of the graph:

k = gradient

A stiffer spring will have a larger value for the force constant k. Beyond point A, the graph is no longer a straight line; its gradient changes and we can no longer use the equation $F = kx$.

If a spring or anything else responds to a pair of tensile forces in the way shown in section OA of Figure 7.7, we say that it obeys **Hooke's law**:

> A material obeys Hooke's law if the extension produced in it is proportional to the applied force (load).

If you apply a small force to a spring and then release it, it will return to its original length. This behaviour is described as 'elastic'. However, if you apply a large force, the spring may not return to its original length. It has become permanently deformed. The force beyond which the spring becomes permanently deformed is known as the **elastic limit**.

QUESTION

7 Figure 7.8 shows the force–extension graphs for four springs, A, B, C and D.

- **a** State which spring has the greatest value of force constant.
- **b** State which is the least stiff.
- **c** State which of the four springs does not obey Hooke's law.

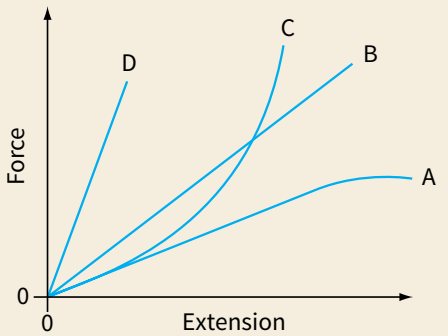


Figure 7.8 Force–extension graphs for four different springs.

Stretching materials

When we determine the force constant of a spring, we are only finding out about the stiffness of that particular spring. However, we can compare the stiffness of different **materials**. For example, steel is stiffer than copper, but copper is stiffer than lead.

Stress and strain

Figure 7.10 shows a simple way of assessing the stiffness of a wire in the laboratory. As the long wire is stretched, the position of the sticky tape pointer can be read from the scale on the bench.

BOX 7.1: Investigating springs

Springs can be combined in different ways (Figure 7.9): end-to-end (in series) and side-by-side (in parallel). Using identical springs, you can measure the force constant of a single spring, and of springs in series and in parallel. Before you do this, predict the outcome of such an experiment. If the force constant of a single spring is k, what will be the equivalent force constant of:

- two springs in series?
- two springs in parallel?

This approach can be applied to combinations of three or more springs.

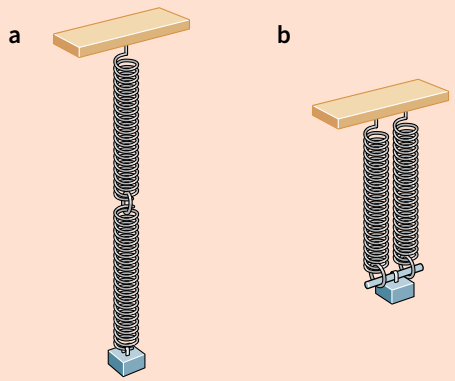


Figure 7.9 Two ways to combine a pair of springs: **a** in series; **b** in parallel.

Why do we use a long wire? Obviously, this is because a short wire would not stretch as much as a long one. We need to take account of this in our calculations, and we do this by calculating the strain produced by the load. The **strain** is defined as the fractional increase in the original length of the wire. That is:

$$\text{strain} = \frac{\text{extension}}{\text{original length}}$$

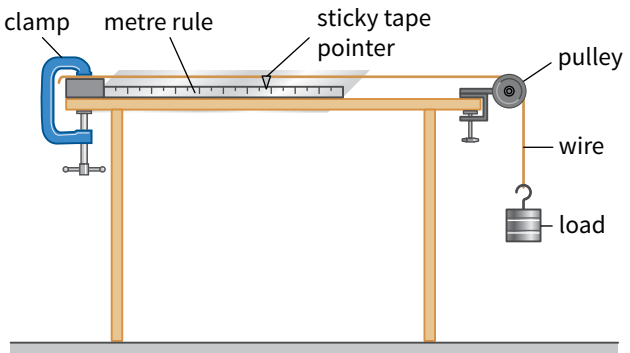


Figure 7.10 Stretching a wire in the laboratory. WEAR EYE PROTECTION and be careful not to overload the wire.

This may be written as:

$$\text{strain} = \frac{x}{L}$$

where x is the extension of the wire and L is its original length.

Note that both extension and original length must be in the same units, and so strain is a ratio, without units. Sometimes strain is given as a percentage. For example, a strain of 0.012 is equivalent to 1.2%.

Why do we use a thin wire? This is because a thick wire would not stretch as much for the same force. Again, we need to take account of this in our calculations, and we do this by calculating the stress produced by the load. The **stress** is defined as the force applied per unit cross-sectional area of the wire. That is:

$$\text{stress} = \frac{\text{force}}{\text{cross-sectional area}}$$

This may be written as:

$$\text{stress} = \frac{F}{A}$$

where F is the applied force on a wire of cross-sectional area A.

The units of stress are newtons per square metre (N m^{-2}) or pascals (Pa), the same as the units of pressure:

$$1\,\text{Pa} = 1\,\text{N m}^{-2}$$

The Young modulus

We can now find the **stiffness** of the **material** we are stretching. Rather than calculating the ratio of force to extension as we would for a spring or a wire, we calculate the ratio of stress to strain. This ratio is a constant for a particular material and does not depend on its shape or size. The ratio of stress to strain is called the **Young modulus** of the material. That is:

$$\text{Young modulus} = \frac{\text{stress}}{\text{strain}}$$

$$\text{or} \qquad = \frac{\sigma}{\varepsilon}$$

where E is the Young modulus of the material, σ (Greek letter sigma) is the stress and ε (epsilon) is the strain.

The unit of the Young modulus is the same as that for stress, N m^{-2} or Pa. In practice, values may be quoted in MPa or GPa. These units are related as:

$$1\,\text{MPa} = 10^6\,\text{Pa}$$

$$1\,\text{GPa} = 10^9\,\text{Pa}$$

Usually, we plot a graph with stress on the vertical axis and strain on the horizontal axis (Figure 7.11). It is drawn like this so that the gradient is the Young modulus of the material. It is important to consider only the first, linear section of the graph. In the linear section stress is proportional to strain and the wire under test obeys Hooke's law.

Table 7.1 gives some values of the Young modulus for different materials.

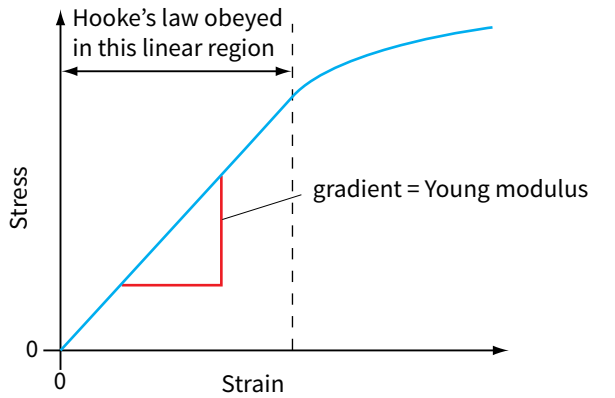


Figure 7.11 Stress–strain graph, and how to deduce the Young modulus. Note that we can only use the first, straight-line section of the graph.

Material	Young modulus / GPa
aluminium	70
brass	90–110
brick	7–20
concrete	40
copper	130
glass	70–80
iron (wrought)	200
lead	18
Perspex®	3
polystyrene	2.7–4.2
rubber	0.01
steel	210
tin	50
wood	10 approx.

Table 7.1 The Young modulus of various materials. Many of these values depend on the precise composition of the material concerned. (Remember, 1 GPa = 10^9 Pa.)

QUESTIONS

8 List the metals in Table 7.1 from stiffest to least stiff.

9 Which of the non-metals in Table 7.1 is the stiffest?

QUESTIONS

10 Figure 7.12 shows stress–strain graphs for two materials, A and B. Use the graphs to determine the Young modulus of each material.

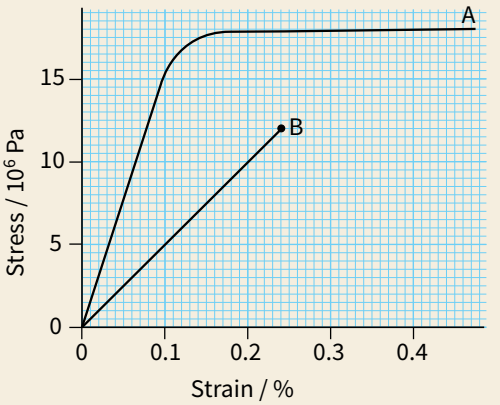


Figure 7.12 Stress–strain graphs for two different materials.

11 A piece of steel wire, 200.0 cm long and having cross-sectional area of 0.50 mm^2, is stretched by a force of 50 N. Its new length is found to be 200.1 cm. Calculate the stress and strain, and the Young modulus of steel.

12 Calculate the extension of a copper wire of length 1.00 m and diameter 1.00 mm when a tensile force of 10 N is applied to the end of the wire. (Young modulus *E* of copper = 130 GPa.)

13 In an experiment to measure the Young modulus of glass, a student draws out a glass rod to form a fibre 0.800 m in length. Using a travelling microscope, she estimates its diameter to be 0.40 mm. Unfortunately it proves impossible to obtain a series of readings for load and extension. The fibre snaps when a load of 1.00 N is hung on the end. The student judges that the fibre extended by no more than 1 mm before it snapped.

Use these values to obtain an estimate for the Young modulus of the glass used. Explain how the actual or accepted value for the Young modulus might differ from this estimate.

14 For each of the materials whose stress–strain graphs are shown in Figure 7.13, deduce the values of the Young modulus.

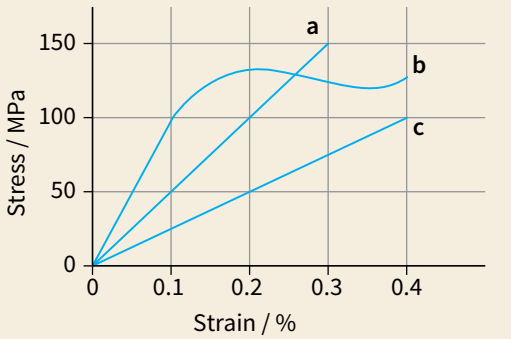


Figure 7.13 Stress–strain graphs for three materials.

BOX 7.2: Determining the Young modulus

Metals are not very elastic. In practice, they can only be stretched by about 0.1% of their original length. Beyond this, they become permanently deformed. As a result, some careful thought must be given to getting results that are good enough to give an accurate value of the Young modulus.

First, the wire used must be long. The increase in length is proportional to the original length, and so a longer wire gives larger and more measurable extensions. Typically, extensions up to 1 mm must be measured for a wire of length 1 m. To get suitable measurements of extension there are two possibilities: use a very long wire, or use a method that allows measurement of extensions that are a fraction of a millimetre.

The apparatus shown in Figure 7.10 can be used with a travelling microscope placed above the wire and focused on the sticky tape pointer. When the pointer moves, the microscope is adjusted to keep the pointer at the middle of the cross-hairs on the microscope. The distance that the pointer has moved can then be measured accurately from the scale on the microscope.

Second, the cross-sectional area of the wire must be known accurately. The diameter of the wire is measured using a micrometer screw gauge. This is reliable to within ±0.01 mm. Once the wire has been loaded in increasing steps, the load must be gradually decreased to ensure that there has been no permanent deformation of the wire.

Other materials such as glass and many plastics are also quite stiff, and so it is difficult to measure their Young modulus. Rubber is not as stiff, and strains of several hundred per cent can be achieved. However, the stress–strain graph for rubber is not a straight line. This means the value of the Young modulus found is not very precise, because it only has a very small linear region on a stress–strain graph.

Elastic potential energy

Whenever you stretch a material, you are doing work. This is because you have to apply a **force** and the material **extends** in the direction of the force. You will know this if you have ever used an exercise machine with springs intended to develop your muscles (Figure 7.14). Similarly, when you push down on the end of a springboard before diving, you are doing work. You transfer energy to the springboard, and you recover the energy when it pushes you up into the air.



Figure 7.14 Using an exercise machine is hard work.

We call the energy in a deformed solid the **elastic potential energy** or **strain energy**. If the material has been strained elastically (the elastic limit has not been exceeded), the energy can be recovered. If the material has been plastically deformed, some of the work done has gone into moving atoms past one another, and the energy cannot be recovered. The material warms up slightly.

We can determine how much elastic potential energy is involved from a force–extension graph: see Figure 7.15. We need to use the equation that defines the amount of work done by a force. That is:

work done
= force × distance moved in the direction of the force

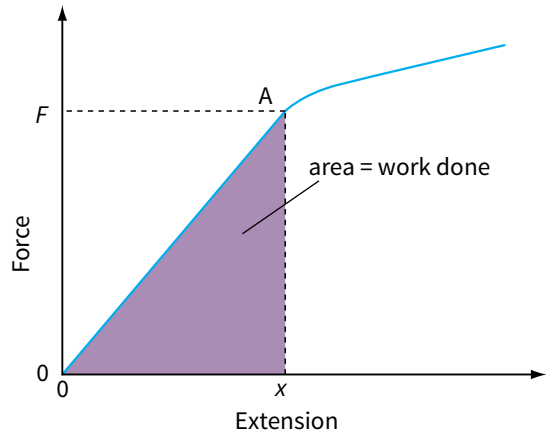


Figure 7.15 Elastic potential energy is equal to the area under the force–extension graph.

First, consider the linear region of the graph where Hooke's law is obeyed, OA. The graph in this region is a straight line through the origin. The extension x is directly proportional to the applied force F. There are two ways to find the work done.

Method 1

We can think about the average force needed to produce an extension x. The average force is half the final force F, and so we can write:

elastic potential energy = work done

$$\text{elastic potential energy} = \frac{\text{final force}}{2} \times \text{extension}$$

$$\text{elastic potential energy} = \tfrac{1}{2}Fx$$

$$\text{or} \qquad E = \tfrac{1}{2}Fx$$

Method 2

The other way to find the elastic potential energy is to recognise that we can get the same answer by finding the area under the graph. The area shaded in Figure 7.15 is a triangle whose area is given by:

$$\text{area} = \tfrac{1}{2} \times \text{base} \times \text{height}$$

This again gives:

$$\text{elastic potential energy} = \tfrac{1}{2}Fx$$

$$\text{or} \qquad E = \tfrac{1}{2}Fx$$

The **work done** in **stretching** or **compressing** a material is always equal to the area under the graph of force against extension. This is true whatever the shape of the graph, provided we draw the graph with extension on the horizontal axis. If the graph is not a straight line, we cannot use the Fx relationship, so we have to resort to counting squares or some other technique to find the answer.

WORKED EXAMPLE

3 Figure 7.16 shows a simplified version of a force–extension graph for a piece of metal. Find the elastic potential energy when the metal is stretched to its elastic limit, and the total work that must be done to break the metal.

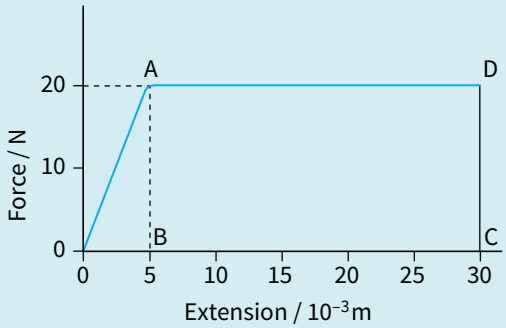


Figure 7.16 For Worked example 3.

Step 1 The elastic potential energy when the metal is stretched to its elastic limit is given by the area under the graph up to the elastic limit. The graph is a straight line up to $x = 5.0\,\text{mm}$, $F = 20\,\text{N}$, so the elastic potential energy is the area of triangle OAB:

$$\text{elastic potential energy} = \tfrac{1}{2}Fx$$
$$= \tfrac{1}{2} \times 20 \times 5.0 \times 10^{-3}$$
$$= 0.050\,\text{J}$$

Step 2 To find the work done to break the metal, we need to add on the area of the rectangle ABCD:

$$\text{work done} = \text{total area under the graph}$$
$$= 0.05 + (20 \times 25 \times 10^{-3})$$
$$= 0.05 + 0.50 = 0.55\,\text{J}$$

Note that the elastic potential energy relates to the elastic part of the graph (i.e. up to the elastic limit), so we can only consider the force–extension graph up to the elastic limit.

There is an alternative equation for elastic potential energy. We know that, according to Hooke's law (page 104), applied force F and extension x are related by $F = kx$, where k is the force constant. Substituting for F gives:

$$\text{elastic potential energy} = \tfrac{1}{2}Fx = \tfrac{1}{2} \times kx \times x$$

$$\text{elastic potential energy} = \tfrac{1}{2}kx^2$$

QUESTIONS

15 A force of 12 N extends a length of rubber band by 18 cm. Estimate the energy stored in this rubber band. Explain why your answer can only be an estimate.

16 A spring has a force constant of $4800\,\text{N m}^{-1}$. Calculate the elastic potential energy when it is compressed by 2.0 mm.

17 Figure 7.17 shows force–extension graphs for two materials. For each of the following questions, explain how you deduce your answer from the graphs.

- **a** State which polymer has the greater stiffness.
- **b** State which polymer requires the greater force to break it.
- **c** State which polymer requires the greater amount of work to be done in order to break it.

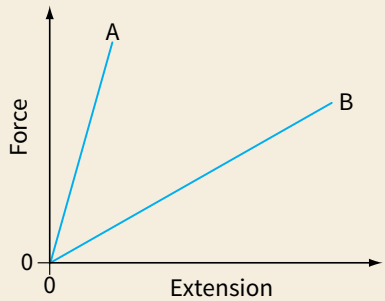


Figure 7.17 Force–extension graph for two polymers.

Summary

- Density is defined as the mass per unit volume of a substance:

  $$\text{density} = \frac{\text{mass}}{\text{volume}}$$

- Pressure is defined as the normal force acting per unit cross-sectional area:

  $$\text{pressure} = \frac{\text{force}}{\text{area}}$$

- Pressure in a fluid increases with depth:

  $p = \rho g h$

- Hooke's law states that the extension of a material is directly proportional to the applied force. For a spring or a wire, $F = kx$, where k is the force constant. The force constant has units of N m^{-1}.

- Stress is defined as:

  $$\text{stress} = \frac{\text{force}}{\text{cross-sectional area}}$$

  $$\sigma = \frac{F}{A}$$

- Strain is defined as:

  $$\text{strain} = \frac{\text{extension}}{\text{original length}}$$

  $$\text{or} \quad \varepsilon = \frac{x}{L}$$

- To describe the behaviour of a material under tensile and compressive forces, we have to draw a graph of stress against strain. The gradient of the initial linear section of the graph is equal to the Young modulus. The Young modulus is an indication of the stiffness of the material.

- The Young modulus E is given by:

  $$E = \frac{\text{stress}}{\text{strain}} = \frac{\sigma}{\varepsilon}$$

- The unit of the Young modulus is pascal (Pa) or N m^{-2}.

- The area under a force–extension graph is equal to the work done by the force.

- For a spring or a wire obeying Hooke's law, the elastic potential energy E is given by:

  $E = \frac{1}{2}kx \qquad F_x = \frac{1}{2}kx^2$

End-of-chapter questions

1 **a** **i** Define **density**. [1]

ii State the base units in the SI system in which density is measured. [1]

b **i** Define **pressure**. [1]

ii State the base units in the SI system in which pressure is measured. [1]

2 Sketch a force–extension graph for a spring which has a spring constant of 20 N m^{-1} and which obeys Hooke's law for forces up to 5 N. Your graph should cover forces between 0 and 6 N and show values on both axes. [3]

3 Two springs, each with a spring constant 20 N m^{-1}, are connected in series. Draw a diagram of the two springs connected in series and determine the total extension if a mass, with weight 2.0 N, is hung on the combined springs. [5]

4 A sample of fishing line is 1.0 m long and is of circular cross-section of radius 0.25 mm. When a weight is hung on the line, the extension is 50 mm and the stress is 2.0×10^8 Pa. Calculate:

a the cross-sectional area of the line [1]

b the weight added [2]

c the strain in the line [2]

d the Young modulus. [2]

5 Figure 7.18 shows the force–extension graph for a metal wire of length 2.0 m and cross-sectional area 1.5×10^{-7} m^2.

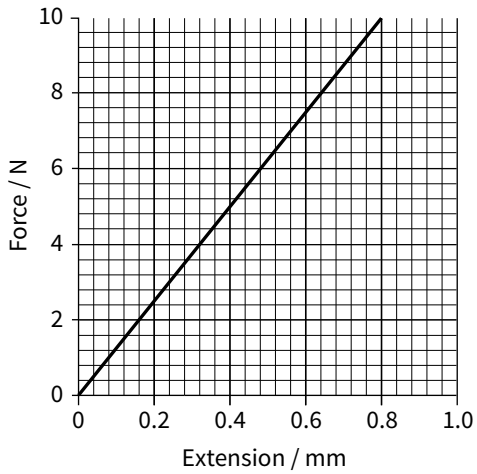


Figure 7.18 For End-of-chapter Question 5.

a Calculate the Young modulus. [3]

b Determine the energy stored in the wire when the extension is 0.8 mm. [2]

c Calculate the work done in stretching the wire between 0.4 mm and 0.8 mm. [2]

6 **a** Explain how a manometer may be adapted to measure large pressures. [2]

b Explain why the pressure measured by a manometer does not depend on the cross-sectional area of the tube used in the manometer. [2]

7 **a** Figure 7.19 shows the stress–strain curves for three different materials, P, Q and R.
State and explain which material has the greatest Young modulus. [2]

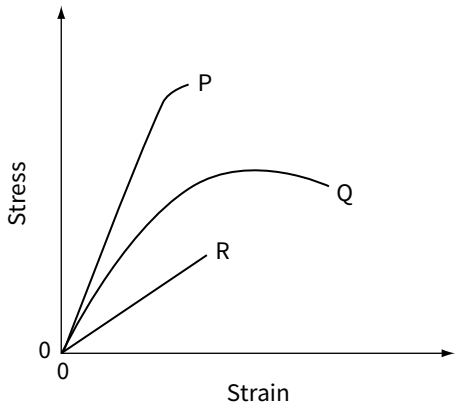


Figure 7.19 For End-of-chapter Question 7.

b Describe an experiment to determine the Young modulus for a material in the form of a wire.
Include a labelled diagram and explain how you would make the necessary measurements.
Show how you would use your measurements to calculate the Young modulus. [7]

8 **a** State the meaning of **tensile stress** and **tensile strain**. [2]

b A vertical steel wire of length 1.6 m and cross-sectional area $1.3 \times 10^{-6}\,m^2$ carries a weight of 60 N.
The Young modulus for steel is 2.1×10^{11} Pa. Calculate:

i the stress in the wire [2]

ii the strain in the wire [2]

iii the extension produced in the wire by the weight. [2]

9 To allow for expansion in the summer when temperatures rise, a steel railway line laid in cold weather is pre-stressed by applying a force of 2.6×10^5 N to the rail of cross-sectional area $5.0 \times 10^{-3}\,m^2$. If the railway line is not pre-stressed then a strain of 1.4×10^{-5} is caused by each degree Celsius rise in temperature.
The Young modulus of the steel is 2.1×10^{11} Pa.

a State and explain whether the force applied to the rail when it is laid should be tensile or compressive. [2]

b Calculate:

i the strain produced when the rail is laid [3]

ii the temperature rise when the rail becomes unstressed. [2]

10 Figure 7.20 shows the stress–strain graph for a metal wire.

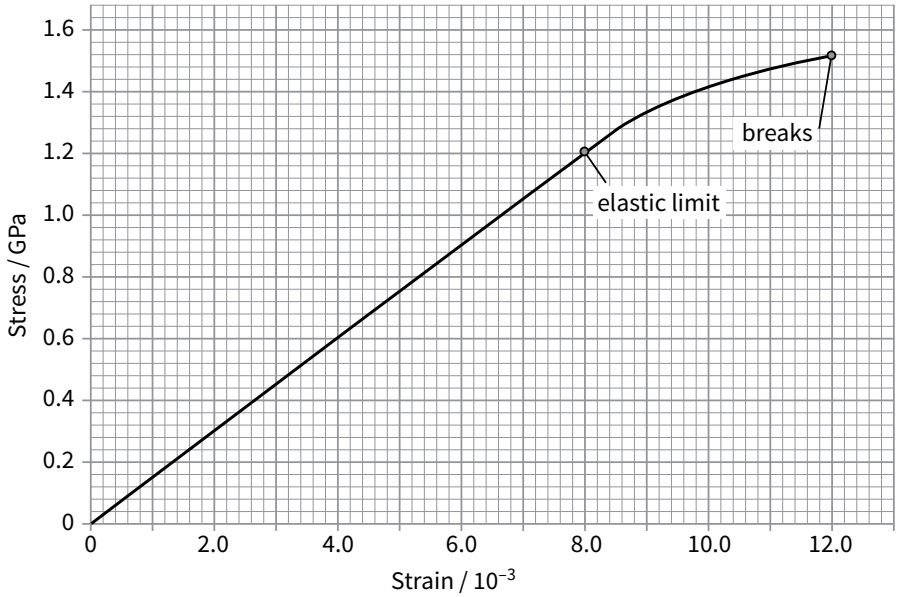


Figure 7.20 For End-of-chapter Question 10.

The wire has a diameter of 0.84 mm and a natural length of 3.5 m. Use the graph to determine:

a the Young modulus of the wire [3]

b the extension of the wire when the stress is 0.6 GPa [2]

c the force which causes the wire to break, assuming that the cross-sectional area of the wire remains constant [3]

d the energy stored when the wire has a stress of 0.6 GPa. [3]

11 Figure 7.21 shows the force–extension graph for a spring.

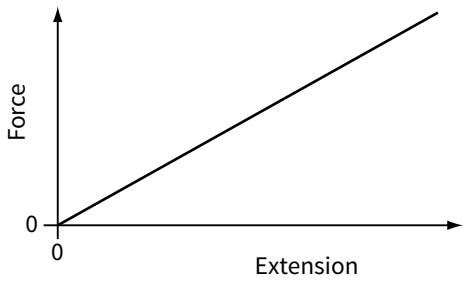


Figure 7.21 For End-of-chapter Question 11.

a State what is represented by:

i the gradient of the graph [1]

ii the area under the graph. [1]

b The spring has force constant $k = 80\,\text{N m}^{-1}$. The spring is compressed by 0.060 m and placed between two trolleys that run on a friction-free, horizontal track. Each trolley has a mass of 0.40 kg. When the spring is released the trolleys fly apart with equal speeds but in opposite directions.

i How much energy is stored in the spring when it is compressed by 0.060 m? [2]

ii Explain why the two trolleys must fly apart with equal speeds. [2]

iii Calculate the speed of each trolley. [2]

12 **a** Liquid of density ρ fills a cylinder of base area A and height h.

i Using the symbols provided, state the mass of liquid in the container. [1]

ii Using your answer to **i**, derive a formula for the pressure exerted on the base of the cylinder. [2]

b A boy stands on a platform of area 0.050 m^2 and a manometer measures the pressure created in a flexible plastic container by the weight W of the boy (Figure 7.22).

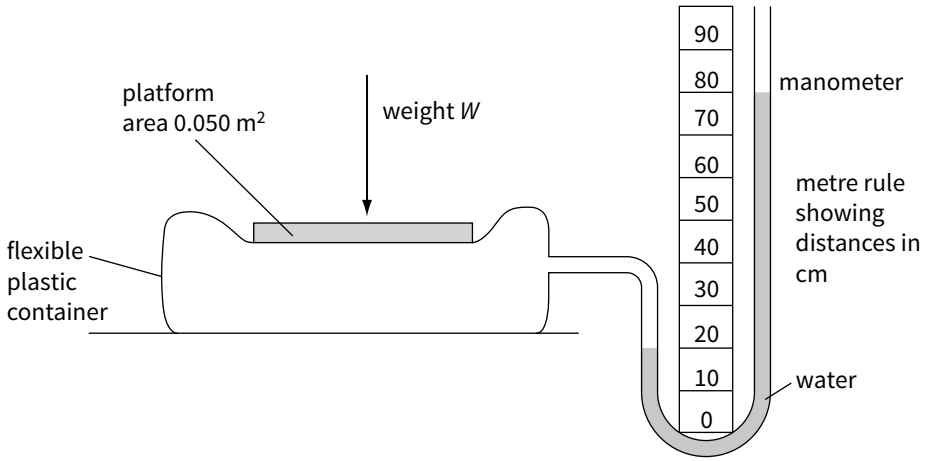


Figure 7.22 For End-of-chapter Question 12.

The density of water is 1000 kg m^{-3}. Determine:

i the pressure difference between the inside of the plastic container and the atmosphere outside [2]

ii the weight W of the boy. [2]

13 Figure 7.23 shows water in a container filled to a depth of 0.50 m. The density of water is 1000 kg m^{-3}.

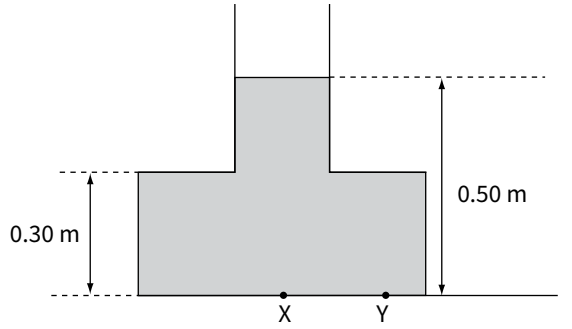


Figure 7.23 For End-of-chapter Question 13.

a Calculate the pressure at X on the base of the container. [2]

b Explain why the pressure at X must be equal to the pressure at Y. [1]

c Explain why the force downwards on the base of the container is larger than the weight of the liquid in the container. [2]

14 **a** Describe why the pressure of the atmosphere is less at the top of a mountain than at the bottom. [1]

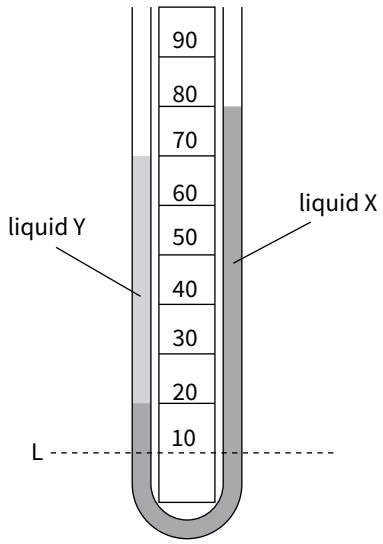


Figure 7.24 For End-of-chapter Question 14.

b Figure 7.24 shows a U-tube, open at both ends, which contains two different liquids, X and Y, that do not mix. The numbers on the metre rule are distances in centimetres. The density of liquid Y is 800 kg m^{-3}.

i Explain how Figure 7.24 shows that liquid Y has a greater density than liquid X. [2]

ii Calculate the density of liquid X. [3]

iii Explain why the pressure in the U-tube is the same on both sides of the manometer at level L. [1]

iv Calculate the pressure caused by liquid in the U-tube at level L. [2]

Chapter 8: Electric fields

Learning outcomes

You should be able to:

- show an understanding of the concept of an electric field
- define electric field strength
- draw field lines to represent an electric field
- calculate the strength of a uniform electric field
- calculate the force on a charge in a uniform electric field
- describe how charged particles move in a uniform electric field

Electricity in nature

The lower surface of a thundercloud is usually negatively charged. When lightning strikes (Figure 8.1), an intense electric current is sent down to the ground below. You may have noticed a 'strobe' effect – this is because each lightning strike usually consists of four or five flashes at intervals of 50 milliseconds or so. You will already know a bit about electric (or electrostatic) fields, from your experience of static electricity in everyday life, and from your studies in science. In this chapter, you will learn how we can make these ideas more formal. We will look at how electric forces are caused, and how we can represent their effects in terms of electric fields. Then we will find mathematical ways of calculating electric forces and field strengths.



Figure 8.1 Lightning flashes, dramatic evidence of natural electric fields.

Attraction and repulsion

Static electricity can be useful – it is important in the process of photocopying, in dust precipitation to clean up industrial emissions, and in crop-spraying, among many other applications. It can also be a nuisance. Who hasn't experienced a shock, perhaps when getting out of a car or when touching a door handle? Static electric charge has built up and gives us a shock when it discharges.

We explain these effects in terms of **electric charge**. Simple observations in the laboratory give us the following picture:

- Objects are usually electrically neutral (uncharged), but they may become electrically charged, for example when one material is rubbed against another.
- There are two types of charge, which we call positive and negative.
- Opposite types of charge attract one another; like charges repel (Figure 8.2).
- A charged object may also be able to attract an uncharged one; this is a result of electrostatic induction.

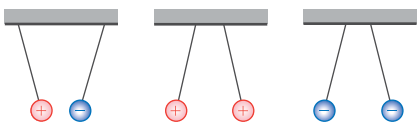


Figure 8.2 Attraction and repulsion between electric charges.

These observations are macroscopic. They are descriptions of phenomena that we can observe in the laboratory, without having to consider what is happening on the microscopic scale, at the level of particles such as atoms and electrons. However, we can give a more subtle explanation if we consider the microscopic picture of static electricity.

Using a simple model, we can consider matter to be made up of three types of particles: electrons (which have negative charge), protons (positive) and neutrons (neutral). An uncharged object has equal numbers of protons and electrons, whose charges therefore cancel out.

When one material is rubbed against another, there is friction between them, and electrons may be rubbed off one material onto the other (Figure 8.3). The material that has gained electrons is now negatively charged, and the other material is positively charged.

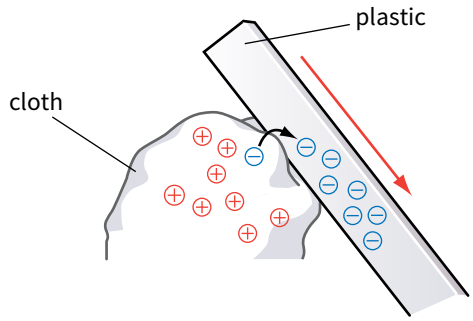


Figure 8.3 Friction can transfer electrons between materials.

If a positively charged object is brought close to an uncharged one, the electrons in the second object may be attracted. We observe this as a force of attraction between the two objects. (This is known as electrostatic induction.)

It is important to appreciate that it is usually electrons that are involved in moving within a material, or from one material to another. This is because electrons, which are on the outside of atoms, are less strongly held within a material than are protons. They may be free to move about within a material (like the conduction electrons in a metal), or they may be relatively weakly bound within atoms.

BOX 8.1: Investigating electric fields

If you rub a strip of plastic so that it becomes charged, and then hold it close to your hair, you feel your hair being pulled upwards. The influence of the charged plastic spreads into the space around it; we say that there is an **electric field** around the charge. To produce an electric field, we need unbalanced charges (as with the charged plastic). To observe the field, we need to put something in it that will respond to the field (as your hair responds). There are two simple ways in which you can do this in the laboratory. The first uses a charged strip of gold foil, attached to an insulating handle (Figure 8.4). The second uses grains of a material such as semolina; these line up in an electric field (Figure 8.5), rather like the way in which iron filings line up in a magnetic field (Figure 8.5).

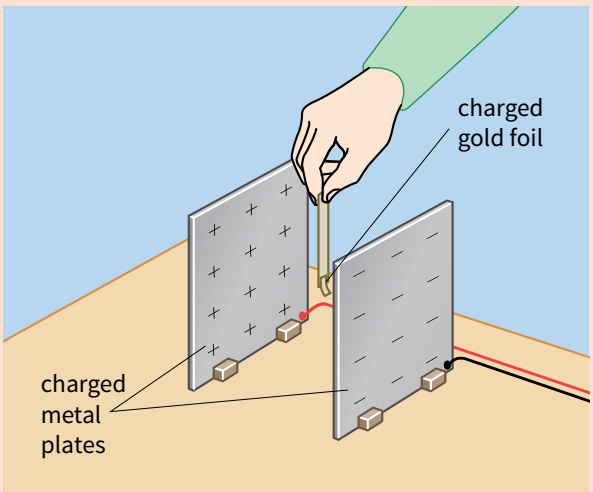


Figure 8.4 Investigating the electric field between two charged metal plates.

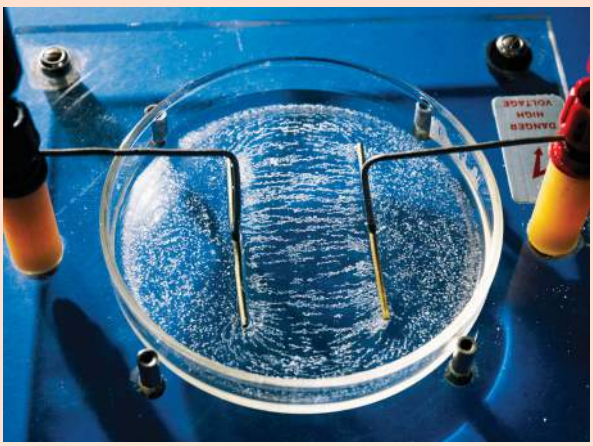


Figure 8.5 Apparatus showing a uniform electric field between two parallel charged plates.

The concept of an electric field

A charged object experiences a force in an electric field. This is what an electric field is. We say that there is an electric field anywhere where an electric charge experiences a force. An electric field is a **field of force**.

This is a rather abstract idea. You will be more familiar with the idea of a 'field of force' from your experience of magnets. There is a magnetic field around a permanent magnet; another magnet placed nearby will experience a force. You have probably plotted the field lines used to represent the field around a magnet. There is a third type of force field which we are all familiar with, because we live in it – a gravitational field. Our weight is the force exerted on us by the gravitational field of the Earth. So we have:

- electric fields – act on objects with electric charge
- magnetic fields – act on magnetic materials, magnets and moving charges (including electric currents)
- gravitational fields – act on objects with mass.

Later we will see that the electric force and the magnetic force are closely linked. They are generally considered as a single entity, the electromagnetic force.

Representing electric fields

We can draw electric fields in much the same way that we can draw magnetic fields, by showing **field lines** (sometimes called lines of force). The three most important field shapes are shown in Figure 8.6.

As with magnetic fields, this representation tells us two things about the field: its direction (from the direction of the lines), and how strong it is (from the separation of the lines). The arrows go from positive to negative; they tell us the direction of the force on a positive charge in the field.

- A uniform field has the same strength at all points. Example: the electric field between oppositely charged parallel plates.
- A radial field spreads outwards in all directions. Example: the electric field around a point charge or a charged sphere.

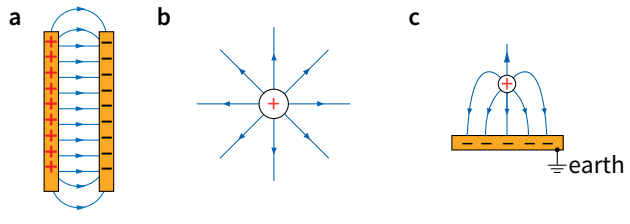


Figure 8.6 Field lines are drawn to represent an electric field. They show the direction of the force on a positive charge placed at a point in the field. **a** A uniform electric field is produced between two oppositely charged plates. **b** A radial electric field surrounds a charged sphere. **c** The electric field between a charged sphere and an earthed plate.

QUESTIONS

1 Which of the three field diagrams in Figure 8.7 represents:

a two positive charges repelling each other?

b two negative charges?

c two opposite charges?

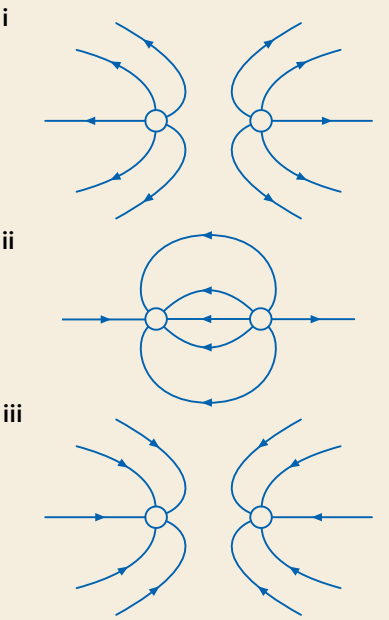


Figure 8.7 Electric fields between charges – see Question 1.

2 Many molecules are described as polar; that is, they have regions that are positively or negatively charged, though they are neutral overall. Draw a diagram to show how sausage-shaped polar molecules like those shown in Figure 8.8 might realign themselves in a solid.



Figure 8.8 Polar molecules – see Question 2.

3 Figure 8.9 shows the electric field pattern between a thundercloud and a building. State and explain where the electric field strength is greatest.

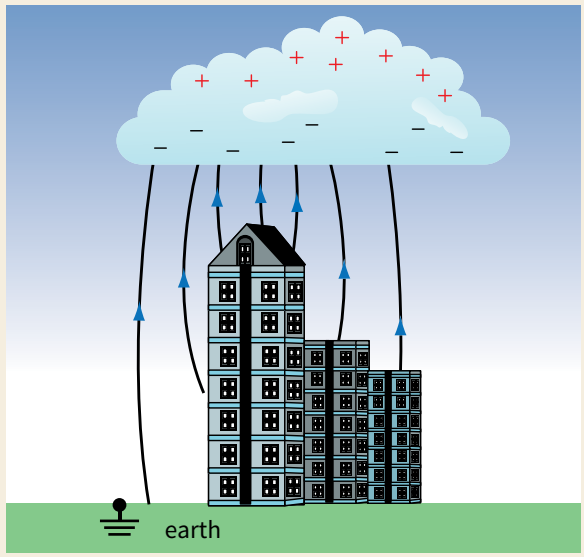


Figure 8.9 Predict where the electric field will be strongest – that's where lightning may strike.

We can draw electric fields for other arrangements. Note the symbol for an earth, which is assumed to be uncharged (i.e. at zero volts).

Electric field strength

For an electric field, we define electric field strength E as follows:

> The **electric field strength** at a point is the force per unit charge exerted on a stationary positive charge at that point.

So, to define electric field strength, we imagine putting a positive test charge $+Q$ in the field and measuring the electric force F that it feels (Figure 8.10). (If you have used a charged gold leaf to investigate a field, this illustrates the principle of testing the field with a charge.)

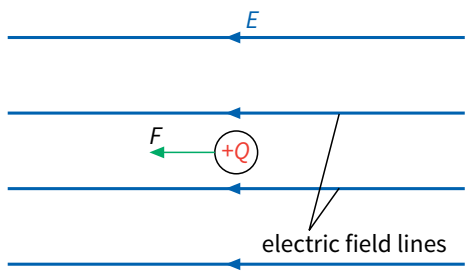


Figure 8.10 A field of strength E exerts force F on charge $+Q$.

From this definition, we can write an equation for E:

$$E = \frac{F}{Q}$$

It follows that the units of electric field strength are newtons per coulomb ($N\,C^{-1}$).

The strength of a uniform field

You can set up a uniform field between two parallel metal plates by connecting them to the terminals of a high-voltage power supply (Figure 8.11). The strength of the field between them depends on two factors:

- the voltage V between the plates – the higher the voltage, the stronger the field: $E \propto V$
- the separation d between the plates – the greater their separation, the weaker the field: $E \propto \frac{1}{d}$

These factors can be combined to give an equation for E:

$$E = -\frac{V}{d}$$

Worked example 1 shows a derivation of this. Note that the minus sign is necessary because, in Figure 8.11, the voltage V increases towards the right while the force F acts in the opposite direction, towards the left. E is a vector quantity. In calculations, we are often interested in the **magnitude** of the electric field strength, hence we can write:

$$E = \frac{V}{d}$$

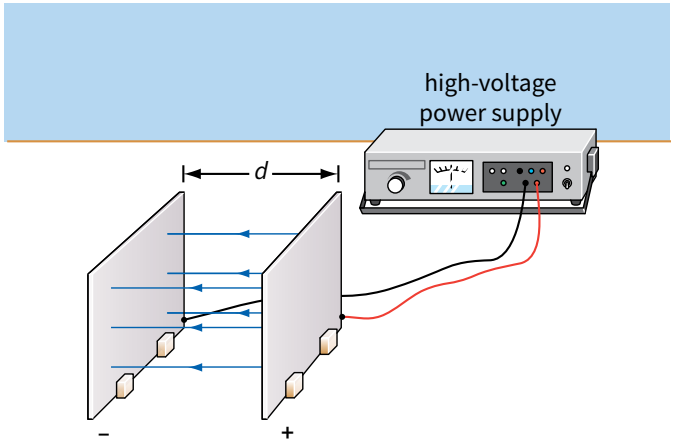


Figure 8.11 There is a uniform field between two parallel, charged plates.

From this equation, we can see that we can write the units of electric field strength as volts per metre ($V\,m^{-1}$). Note:

$$1\,V\,m^{-1} = 1\,N\,C^{-1}$$

Worked example 2 shows how to solve problems involving uniform fields.

QUESTIONS

4 Figure 8.12 shows an arrangement of parallel plates, each at a different voltage. The electric field lines are shown in the space between the first pair. Copy and complete the diagram to show the electric field lines in the other two spaces.

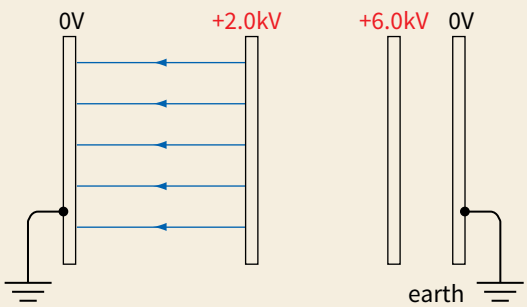


Figure 8.12 An arrangement of parallel plates – see Question 4.

5 Calculate the electric field strength at a point where a charge of 20 mC experiences a force vertically downwards of 150 N.

6 Calculate the electric field strength between two parallel charged plates, separated by 40 cm and with a potential difference between them of 1000 V.

7 An electron is situated in a uniform electric field. The electric force that acts on it is 8×10^{-16} N. What is the strength of the electric field?
(Electron charge $e = 1.6 \times 10^{-19}$ C.)

QUESTION

8 Air is usually a good insulator. However, a spark can jump through dry air when the electric field strength is greater than about 40 000 V cm^{-1}. This is called electrical breakdown. The spark shows that electrical charge is passing through the air – a current is flowing. (Do not confuse this with a chemical spark such as you might see when watching fireworks; in that case, small particles of a chemical substance are burning quickly.)

a A Van de Graaff generator (Figure 8.13) is able to make sparks jump across a 4 cm gap. What is the voltage produced by the generator?

b The highest voltage reached by the live wire of a conventional mains supply is 325 V. In theory (but DO NOT try this), how close would you have to get to a live wire to get a shock from it?

c Estimate the voltage of a thundercloud from which lightning strikes the ground 100 m below.

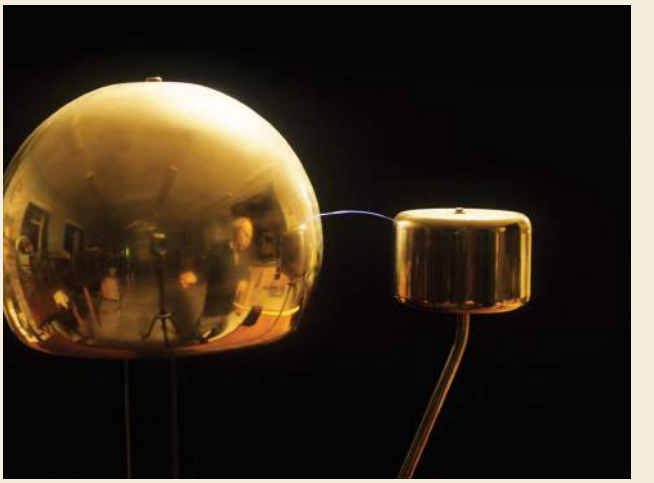


Figure 8.13 A Van de Graaff generator produces voltages sufficient to cause sparks in air.

WORKED EXAMPLES

1 Two metal plates are separated by a distance d. The potential difference between the plates is V. A positive charge Q is pulled at a constant speed with a constant force F from the negative plate all the way to the positive plate. Using the definition for electric field strength and the concept of work done, show that the magnitude of the electric field strength E is given by the equation:

$$E = \frac{V}{d}$$

Step 1 We have:

work done on charge = energy transformed

From their definitions, we can write:

work done = force × distance or $W = Fd$

energy transformed = VQ

Step 2 Substituting gives:

$Fd = VQ$

and rearranging gives:

$$\frac{F}{Q} = \frac{V}{d}$$

Step 3 The left-hand side of the equation is the electric field strength E. Hence:

$$E = \frac{V}{d}$$

2 Two parallel metal plates separated by 2.0 cm have a potential difference of 5.0 kV. Calculate the electric force acting on a dust particle between the plates that has a charge of 8.0×10^{-19} C.

Step 1 Write down the quantities given in the question:

$d = 2.0 \times 10^{-2}$ m

$V = 5.0 \times 10^{3}$ V

$Q = 8.0 \times 10^{-19}$ C

Hint: *When you write down the quantities it is important to include the units and to change them into base units. We have used powers of ten to do this.*

Step 2 To calculate the force F, you first need to determine the strength of the electric field:

$$E = \frac{V}{d}$$

$$E = \frac{50 \times 10^{3}}{2.0 \times 10^{-2}} = 2.5 \times 10^{5}\ \text{V m}^{-1}$$

Step 3 Now calculate the force on the dust particle:

$F = EQ$

$F = 2.5 \times 10^{5} \times 8.0 \times 10^{-19}$

$= 2.0 \times 10^{-13}$ N

Force on a charge

Now we can calculate the force F on a charge Q in the uniform field between two parallel plates. We have to combine the general equation for field strength $E = \frac{F}{Q}$ with the equation for the strength of a uniform field $E = -\frac{V}{d}$. This gives:

$$F = QE = -\frac{QV}{d}$$

For an electron with charge $-e$, this becomes:

$$F = \frac{eV}{d}$$

Figure 8.14 shows a situation where this force is important. A beam of electrons is entering the space between two charged parallel plates. How will the beam move?

We have to think about the force on a single electron. In the diagram, the upper plate is negative relative to the lower plate, and so the electron is pushed downwards. (You can think of this simply as the negatively charged electron being attracted by the positive plate, and repelled by the negative plate.)

If the electron were stationary, it would accelerate directly downwards. However, in this example, the electron is moving to the right. Its horizontal velocity will be unaffected by the force, but as it moves sideways it will also accelerate downwards. It will follow a curved path, as shown. This curve is a parabola.

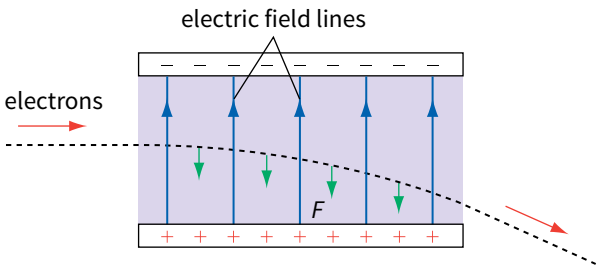


Figure 8.14 The parabolic path of a moving electron in a uniform electric field.

Note that the force on the electron is the same at all points between the plates, and it is always in the same direction (downwards, in this example).

This situation is equivalent to a ball being thrown horizontally in the Earth's uniform gravitational field (Figure 8.15). It continues to move at a steady speed horizontally, but at the same time it accelerates downwards. The result is the familiar curved trajectory shown. For the electron described above, the force of gravity is tiny – negligible compared to the electric force on it.

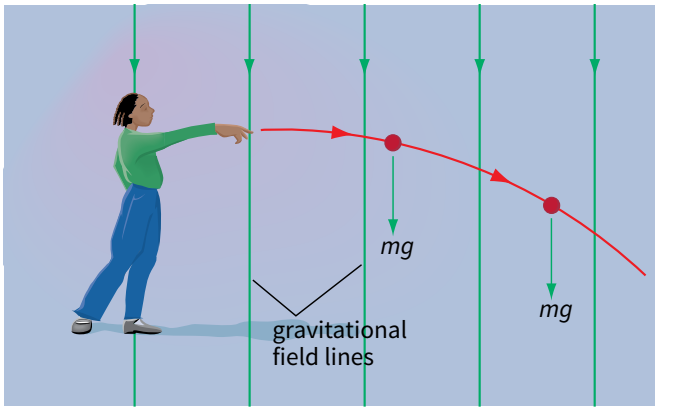


Figure 8.15 A ball, thrown in the uniform gravitational field of the Earth, follows a parabolic path.

QUESTIONS

9 In Figure 8.16, two parallel plates are shown, separated by 25 cm.

- **a** Copy the diagram and draw field lines to represent the field between the plates.
- **b** What is the potential difference between points A and B?
- **c** What is the electric field strength at C, and at D?
- **d** Calculate the electric force on a charge of +5 μC placed at C. In which direction does the force act?

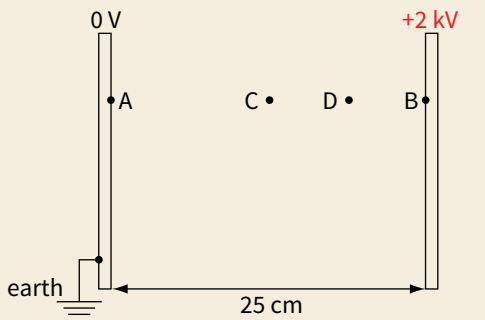


Figure 8.16 Two parallel, charged plates.

10 A particle of charge +2 μC is placed between two parallel plates, 10 cm apart and with a potential difference of 5 kV between them. Calculate the field strength between the plates, and the force exerted on the charge.

QUESTIONS

11 We are used to experiencing accelerations that are usually less than $10\,m\,s^{-2}$. For example, when we fall, our acceleration is about $9.81\,m\,s^{-2}$. When a car turns a corner sharply at speed, its acceleration is unlikely to be more than $5\,m\,s^{-2}$. However, if you were an electron, you would be used to experiencing much greater accelerations than this. Calculate the acceleration of an electron in a television tube where the electric field strength is $50\,000\,V\,cm^{-1}$.
(Electron charge $-e = -1.6 \times 10^{-19}\,C$;
electron mass $m_e = 9.11 \times 10^{-31}\,kg$.)

12 **a** Use a diagram to explain how the electric force on a charged particle could be used to separate a beam of electrons (e^-) and positrons (e^+) into two separate beams. (A positron is a positively charged particle that has the same mass as an electron but opposite charge. Positron–electron pairs are often produced in collisions in a particle accelerator.)

b Explain how this effect could be used to separate ions that have different masses and charges.

Summary

- An electric field is a field of force, created by electric charges, and can be represented by electric field lines.
- The strength of the field is the force acting per unit positive charge at a point in the field, $E = \frac{F}{Q}$.
- In a uniform field (e.g. between two parallel charged plates), the force on a charge is the same at all points; the strength of the field is given by $E = -\frac{V}{d}$.
- An electric charge moving initially at right-angles to a uniform electric field follows a parabolic path.

End-of-chapter questions

1 Figure 8.4 on page 118 shows apparatus used to investigate the field between a pair of charged, parallel plates.

a Explain why the piece of gold foil deflects in the manner shown. [1]

b State and explain what would be observed if the gold foil momentarily touched the negatively charged plate. [2]

2 A charged dust particle in an electric field experiences a force of 4.4×10^{-13} N. The charge on the particle is 8.8×10^{-17} C. Calculate the electric field strength. [2]

3 Calculate the potential difference that must be applied across a pair of parallel plates, placed 4 cm apart, to produce an electric field of 4000 V m^{-1}. [2]

4 A potential difference of 2.4 kV is applied across a pair of parallel plates. The electric field strength between the plates is 3.0×10^{4} V m^{-1}.

a Calculate the separation of the plates. [2]

b The plates are now moved so that they are 2.0 cm apart. Calculate the electric field strength produced in this new position. [2]

5 A variable power supply is connected across a pair of parallel plates. The potential difference across the plates is doubled and the distance between the plates is decreased to one-third of the original. State by what factor the electric field changes. Explain your reasoning. [3]

6 Figure 8.17 shows a positively charged sphere hanging by an insulating thread close to an earthed metal plate.

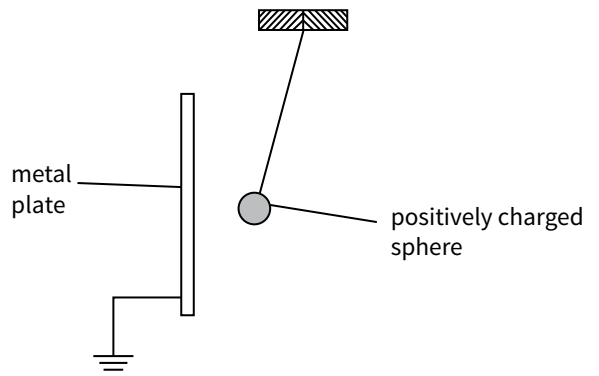


Figure 8.17 For End-of-chapter Question 6.

a Copy Figure 8.17 and draw **five** lines to show the electric field near the plate and the sphere. [3]

b Explain why the sphere is attracted towards the metal plate. [2]

c The sphere is now replaced with a similar negatively charged sphere.

i Explain what would be observed when the sphere is brought near to the earthed metal plate. [2]

ii Describe any changes to the electric field that would occur. [1]

7 Figure 8.18 shows a proton as it moves between two charged parallel plates. The charge on the proton is $+1.6 \times 10^{-19}$ C.

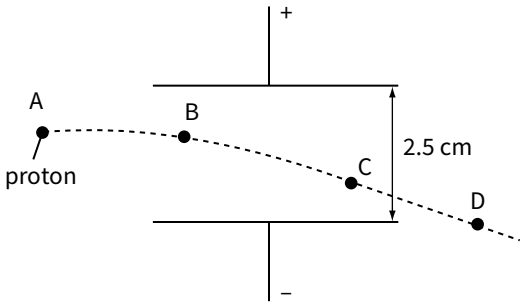


Figure 8.18 For End-of-chapter Question 7.

a Copy Figure 8.18 and draw the electric field between the parallel plates. [2]

The force on the proton when it is at position B is 6.4×10^{-14} N.

b In which direction does the force on the proton act when it is at position B? [1]

c What will be the magnitude of the force on the proton when it is at position C? [1]

d Calculate the electric field strength between the plates. [2]

e Calculate the potential difference between the plates. [2]

8 **a** Define what is meant by the **electric field strength** at a point. [2]

In a particle accelerator a proton, initially at rest, is accelerated between two metal plates, as shown in Figure 8.19.

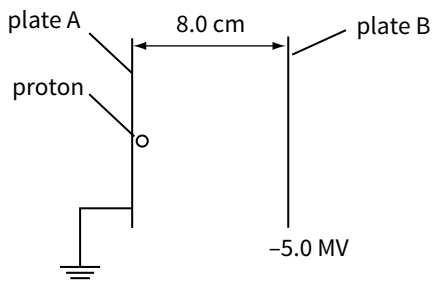


Figure 8.19 For End-of-chapter Question 8.

b Calculate the force on the proton due to the electric field. [3]

c Calculate the work done on the proton by the electric field when it moves from plate A to plate B. [2]

d State the energy gained by the proton. [1]

e Assuming that all this energy is converted to kinetic energy of the proton, calculate the speed of the proton when it reaches plate B. [3]

(Charge on a proton = $+1.6 \times 10^{-19}$ C; mass of a proton = 1.7×10^{-27} kg.)

9 a Figure 8.20 shows the structure of a spark plug in an internal combustion engine. Figure 8.21 is an enlarged version of the end of the spark plug, showing some of the lines of force representing the electric field.

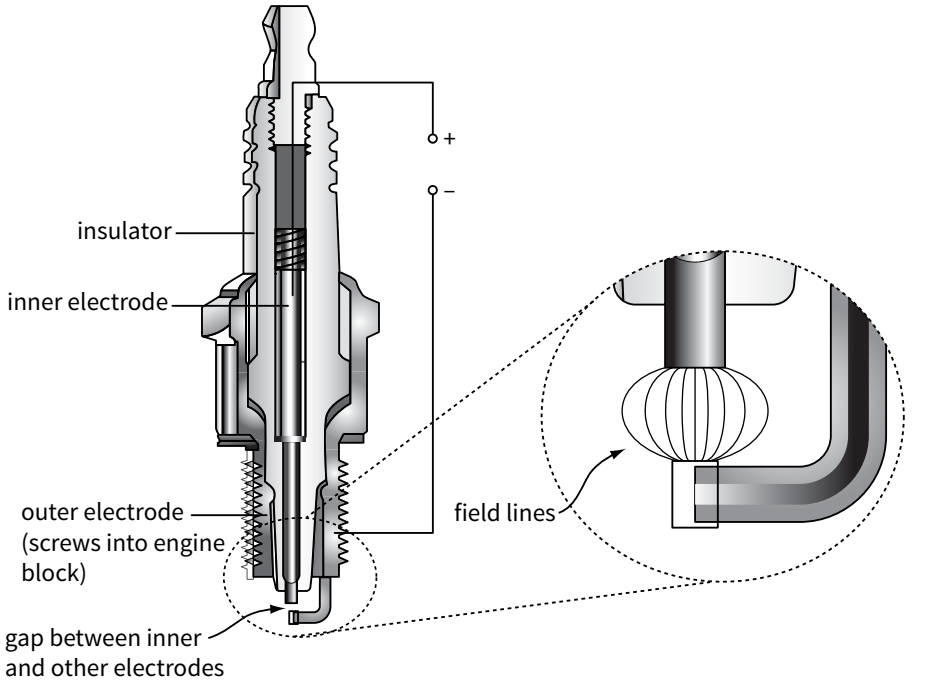


Figure 8.20

Figure 8.21

a i Copy the field lines from Figure 8.21. On your copy, draw arrows on the lines of force to show the direction of the field. [1]

ii What evidence does the diagram give that the field is strongest near the tip of the inner electrode? [1]

b The gap between the inner and outer electrodes is 1.25 mm and a field strength of 5.0×10^6 N C^{-1} is required for electrical breakdown. Estimate the minimum potential difference that must be applied across the inner and outer electrodes for a spark to be produced. (You may treat the two electrodes as a pair of parallel plates.) [2]

c When an electron is accelerated through a potential drop of approximately 20 V it will have sufficient energy to ionise a nitrogen atom. Show that an electron must move 4.0 μm to gain this energy. [2]



Chapter 9: Electric current, potential difference and resistance

Learning outcomes

You should be able to:

- show an understanding of the nature of electric current
- define charge and the coulomb
- solve problems using the equation $Q = It$
- derive $I = nAve$
- solve problems involving the mean drift velocity of charge carriers
- define potential difference, e.m.f. and the volt
- use energy considerations to distinguish between p.d. and e.m.f.
- define resistance and the ohm
- solve problems using $V = IR$
- solve problems concerning energy and power in electric circuits

Developing ideas

Electricity plays a vital part in our lives. We use electricity as a way of transferring energy from place to place – for heating, lighting and making things move. For people in a developing nation, the arrival of a reliable electricity supply marks a great leap forward. In Kenya, a micro-hydroelectric scheme has been built on Kabiri Falls, on the slopes of Mount Kenya. Although this produces just 14 kW of power, it has given work to a number of people, as shown in Figures 9.1, 9.2 and 9.3.



Figure 9.1 An operator controls the water inlet at the Kabiri Falls power plant. The generator is on the right.



Figure 9.2 A metal workshop uses electrical welding equipment. This allows rapid repairs to farmers' machinery.



Figure 9.3 A hairdresser can now work in the evenings, thanks to electrical lighting.

Circuit symbols and diagrams

Before we go on to study electricity we need to introduce the concept of circuit diagrams. It is impossible to draw anything but the simplest circuits in a pictorial manner. To make it possible to draw complex circuits, a shorthand method using standard circuit symbols is used. You will have met many circuit components and their symbols in your previous studies. Some are shown in Table 9.1 and Figure 9.4.

The symbols in Table 9.1 are a small part of a set of internationally agreed conventional symbols for electrical components. It is essential that scientists, engineers, manufacturers and others around the world use the same symbol for a particular component. In addition, many circuits are now designed by computers and these need a universal language in which to work and to present their results.

Symbol	Component name
	connecting lead
	cell
	battery of cells
	fixed resistor
	power supply
	junction of conductors
	crossing conductors (no connection)
	filament lamp
V	voltmeter
A	ammeter
	switch
	variable resistor
	microphone
	loudspeaker
	fuse
	earth
	alternating signal
	capacitor
	thermistor
	light-dependent resistor (LDR)
	semiconductor diode
	light-emitting diode (LED)

Table 9.1 Electrical components and their circuit symbols.

The International Electrotechnical Commission (IEC) is the body which establishes agreements on such things as electrical symbols, as well as safety standards, working practices and so on. The circuit symbols used here form part of an international standard known as IEC 60617. Because this is a shared 'language', there is less likelihood that misunderstandings will arise between people working in different organisations and different countries.

What's in a word?

Electricity is a rather tricky word. In everyday life, its meaning may be rather vague – sometimes we use it to mean electric current; at other times, it may mean electrical energy or electrical power. In this chapter and the ones which follow, we will avoid using the word electricity and try to develop the correct usage of these more precise scientific terms.

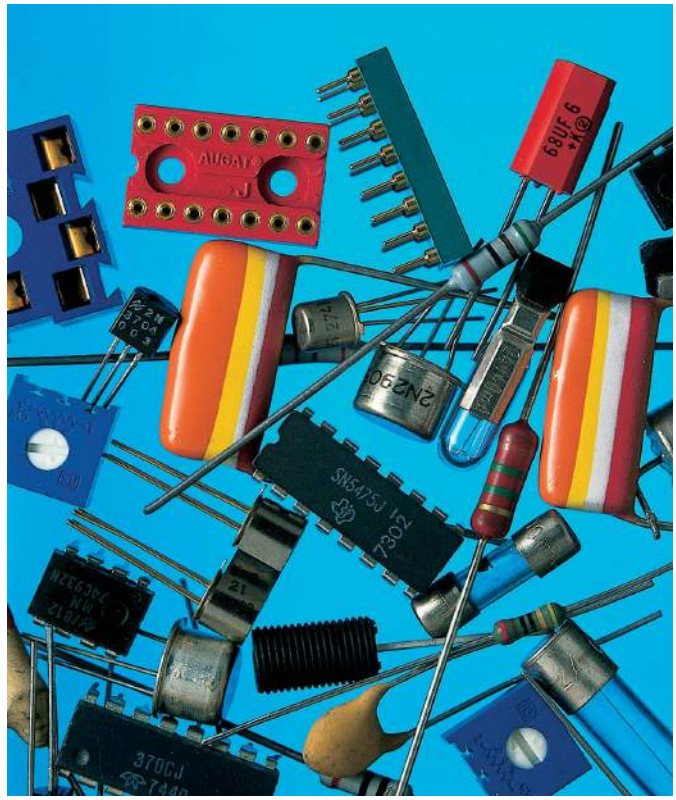


Figure 9.4 A selection of electrical components, including resistors, fuses, capacitors and microchips.

Electric current

You will have carried out many practical activities involving electric current. For example, if you connect a wire to a cell (Figure 9.5), there will be current in the wire. And of course you make use of electric currents every day of your life – when you switch on a lamp or a computer, for example.

In the circuit of Figure 9.5, the direction of the current is from the positive terminal of the cell, around the circuit to the negative terminal. This is a scientific convention: the direction of current is from positive to negative, and hence the current may be referred to as **conventional current**. But what is going on inside the wire?

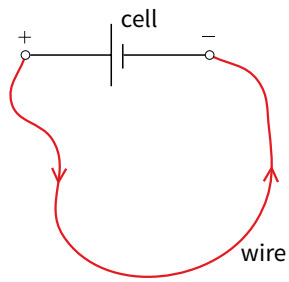


Figure 9.5 There is current in the wire when it is connected to a cell.

A wire is made of metal. Inside a metal, there are negatively charged electrons which are free to move about. We call these **conduction** or **free** electrons, because they are the particles which allow a metal to conduct an electric current. The atoms of a metal bind tightly together; they usually form a regular array, as shown in Figure 9.6. In a typical metal, such as copper or silver, one electron from each atom breaks free to become a conduction electron. The atom remains as a positively charged ion. Since there are equal numbers of free electrons (negative) and ions (positive), the metal has no overall charge – it is neutral.

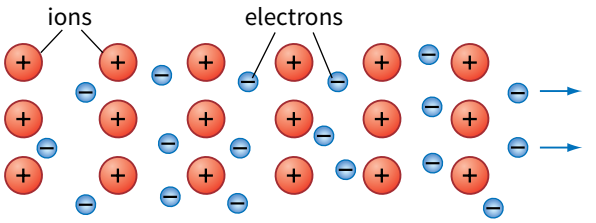


Figure 9.6 In a metal, conduction electrons are free to move among the fixed positive ions. A cell connected across the ends of the metal causes the electrons to drift towards its positive terminal.

When the cell is connected to the wire, it exerts an electrical force on the conduction electrons that makes them travel along the length of the wire. Since electrons are negatively charged, they flow away from the negative terminal of the cell and towards the positive terminal. This is in the opposite direction to conventional current. This may seem a bit odd; it comes about because the direction of conventional current was chosen long before anyone had any idea what was going on inside a piece of metal carrying a current. If the names positive and negative had originally been allocated the other way round, we would now label electrons as positively charged, and conventional current and electron flow would be in the same direction.

Note that there is a current at all points in the circuit as soon as the circuit is completed. We do not have to wait for charge to travel around from the cell. This is because the charged electrons are already present throughout the metal before the cell is connected.

We can use the idea of an electric field to explain why charge flows almost instantly. Connect the terminals of a cell to the two ends of a wire and we have a complete circuit. The cell produces an electric field in the wire; the field lines are along the wire, from the positive terminal to the negative. This means that there is a force on each electron in the wire, so each electron starts to move and the current exists almost instantly.

Charge carriers

Sometimes a current is a flow of positive charges – for example, a beam of protons produced in a particle accelerator. The current is in the same direction as the particles. Sometimes a current is due to both positive and negative charges – for example, when charged particles flow through a solution. A solution which conducts is called an **electrolyte** and it contains both positive and negative ions. These move in opposite directions when the solution is connected to a cell (Figure 9.7). Any charged particles which contribute to an electric current are known as **charge carriers**; these can be electrons, protons or ions.

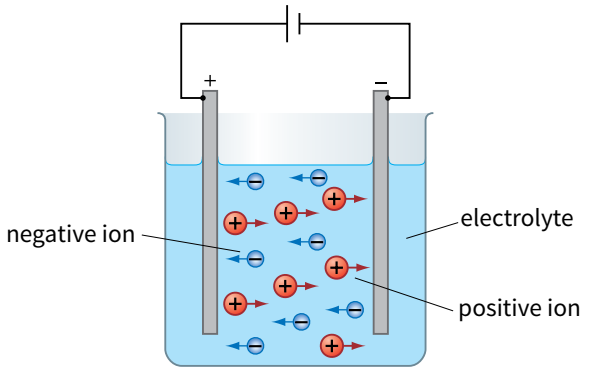


Figure 9.7 Both positive and negative charges are free to move in a solution. Both contribute to the electric current.

QUESTIONS

1 Look at Figure 9.7 and state the direction of the conventional current in the electrolyte (towards the left, towards the right or in both directions at the same time).

2 Figure 9.8 shows a circuit with a conducting solution having both positive and negative ions.

a Copy the diagram and draw in a cell between points A and B. Clearly indicate the positive and negative terminals of the cell.

b Add an arrow to show the direction of the conventional current in the solution.

c Add arrows to show the direction of the conventional current in the connecting wires.

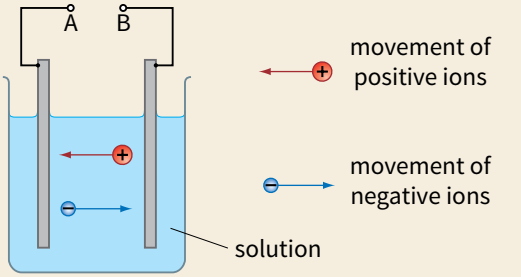


Figure 9.8 For Question 2.

Current and charge

When charged particles flow past a point in a circuit, we say that there is a current in the circuit. Electrical current is measured in **amperes** (A). So how much charge is moving when there is a current of 1 A? Charge is measured in **coulombs** (C). For a current of 1 A, the rate at which charge passes a point in a circuit is 1 C in a time of 1 s. Similarly, a current of 2 A gives a charge of 2 C in a time of 1 s. A current of 3 A gives a charge of 6 C in a time of 2 s, and so on. The relationship between charge, current and time may be written as the following word equation:

$$\text{current} = \frac{\text{charge}}{\text{time}}$$

This equation explains what we mean by current.

Electric current is the rate of flow of electric charge past a point.

The equation for current can be rearranged to give an equation for charge:

$$\text{charge} = \text{current} \times \text{time}$$

This gives us the definition of the unit of charge, the coulomb.

One coulomb is the charge which flows past a point in a circuit in a time of 1 s when the current is 1 A.

In symbols, the charge flowing past a point is given by the relationship:

$$\Delta Q = I\Delta t$$

where ΔQ is the charge which flows during a time Δt and I is the current.

Note that the ampere and the coulomb are both SI units; the ampere is a base unit while the coulomb is a derived unit (see page 40).

QUESTIONS

3 The current in a circuit is 0.40 A. Calculate the charge that passes a point in the circuit in a period of 15 s.

4 Calculate the current that gives a charge flow of 150 C in a time of 30 s.

5 In a circuit, a charge of 50 C passes a point in 20 s. Calculate the current in the circuit.

6 A car battery is labelled '50 A h'. This means that it can supply a current of 50 A for one hour.

a For how long could the battery supply a continuous current of 200 A needed to start the car?

b Calculate the charge which flows past a point in the circuit in this time.

Charged particles

As we have seen, current is the flow of charged particles called charge carriers. But how much charge does each particle carry?

Electrons each carry a tiny negative charge of approximately -1.6×10^{-19} C. This charge is represented by $-e$. The magnitude of the charge is known as the **elementary charge**. This charge is so tiny that you would need about six million million million electrons – that's 6 000 000 000 000 000 000 of them – to have a charge equivalent to one coulomb.

elementary charge $e = 1.6 \times 10^{-19}$ C

Protons are positively charged, with a charge $+e$. This is equal and opposite to that of an electron. Ions carry charges that are multiples of $+e$ and $-e$.

WORKED EXAMPLES

1 There is a current of 10 A through a lamp for 1.0 hour. Calculate how much charge flows through the lamp in this time.

Step 1 We need to find the time t in seconds:

$\Delta t = 60 \times 60 = 3600$ s

Step 2 We know the current $I = 10$ A, so the charge which flows is:

$\Delta Q = I\Delta t = 10 \times 3600 = 36\,000$ C $= 3.6 \times 10^4$ C

2 Calculate the current in a circuit when a charge of 180 C passes a point in a circuit in 2.0 minutes.

Step 1 Rearranging $Q = It$ gives:

$I = \frac{\Delta Q}{\Delta t}$ (or current $= \frac{\text{charge}}{\text{time}}$)

Step 2 With time in seconds, we then have:

current $I = \frac{180}{120} = 1.5$ A

Because electric charge is carried by particles, it must come in amounts which are multiples of e. So, for example, 3.2×10^{-19} C is possible, because this is $+2e$, but 2.5×10^{-19} C is impossible, because this is not an integer multiple of e.

We say that charge is 'quantised'; this means that it can only come in amounts which are integer multiples of the elementary charge. If you are studying chemistry, you will know that ions have charges of $\pm e$, $\pm 2e$, etc. The only exception is in the case of the fundamental particles called quarks, which are the building blocks from which particles such as protons and neutrons are made. These have charges of $\pm\frac{1}{3}e$ or $\pm\frac{2}{3}e$. However, quarks always appear in twos or threes in such a way that their combined charge is zero or a multiple of e.

QUESTIONS

7 Calculate the number of protons which would have a charge of one coulomb.
(Proton charge = $+1.6 \times 10^{-19}$ C.)

8 Which of the following quantities of electric charge is possible? Explain how you know.
6.0×10^{-19} C, 8.0×10^{-19} C, 10.0×10^{-19} C

An equation for current

Copper, silver and gold are good conductors of electric current. There are large numbers of conduction electrons in a copper wire – as many conduction electrons as there are atoms. The number of conduction electrons per unit volume (e.g. in 1 m^3 of the metal) is called the **number density** and has the symbol n. For copper, the value of n is about 10^{29} m^{-3}.

Figure 9.9 shows a length of wire, with cross-sectional area A, along which there is a current I. How fast do the electrons have to travel? The following equation allows us to answer this question:

$$I = nAvq$$

Here, v is called the **mean drift velocity** of the electrons and q is the charge of each particle carrying the current. Since these are usually electrons, we can replace q by e, where e is the elementary charge. The equation then becomes:

$$I = nAve$$

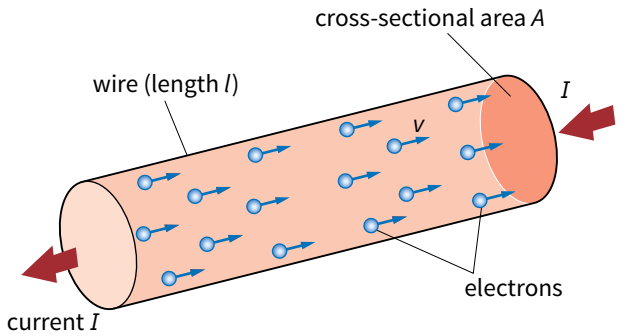


Figure 9.9 A current I in a wire of cross-sectional area A. The charge carriers are mobile conduction electrons with mean drift velocity v.

Deriving $I = nAve$

Look at the wire shown in Figure 9.9. Its length is l. We imagine that all of the electrons shown travel at the same speed v along the wire.

Now imagine that you are timing the electrons to determine their speed. You start timing when the first electron emerges from the right-hand end of the wire. You stop timing when the last of the electrons shown in the diagram emerges. (This is the electron shown at the left-hand end of the wire in the diagram.) Your timer shows that this electron has taken time t to travel the distance l.

In the time t, all of the electrons in the length l of wire have emerged from the wire. We can calculate how many electrons this is, and hence the charge that has flowed in time t:

$$\text{number of electrons} = \text{number density} \times \text{volume of wire}$$
$$= n \times A \times l$$

$$\text{charge of electrons} = \text{number} \times \text{electron charge}$$
$$= n \times A \times l \times e$$

We can find the current I because we know that this is the charge that flows in time t, and current = charge/time:

$$I = n \times A \times l \times e / t$$

Substituting v for l / t gives

$$I = nAve$$

The moving charge carriers that make up a current are not always electrons. They might, for example, be ions (positive or negative) whose charge q is a multiple of e. Hence we can write a more general version of the equation as

$$I = nAvq$$

Worked example 3 shows how to use this equation to calculate a typical value of v.

WORKED EXAMPLE

3 Calculate the mean drift velocity of the electrons in a copper wire of cross-sectional area $5.0 \times 10^{-6}\,m^2$ carrying a current of 1.0 A. The electron number density for copper is $8.5 \times 10^{28}\,m^{-3}$.

Step 1 Rearrange the equation $I = nAve$ to make v the subject:

$$v = \frac{I}{nAe}$$

Step 2 Substitute values and calculate v:

$$v = \frac{1.0}{8.5 \times 10^{28} \times 5.0 \times 10^{-6} \times 1.6 \times 10^{-19}}$$

$$= 1.47 \times 10^{-5}\,m\,s^{-1}$$

$$= 0.015\,mm\,s^{-1}$$

Slow flow

It may surprise you to find that, as suggested by the result of Worked example 3, electrons in a copper wire drift at a fraction of a millimetre per second. To understand this result fully, we need to closely examine how electrons behave in a metal. The conduction electrons are free to move around inside the metal. When the wire is connected to a battery or an external power supply, each electron within the metal experiences an electrical force that causes it to move towards the positive end of the battery. The electrons randomly collide with the fixed but vibrating metal ions. Their journey along the metal is very haphazard. The actual velocity of an electron between collisions is of the order of magnitude $10^5\,m\,s^{-1}$, but its haphazard journey causes it to have a **drift velocity** towards the positive end of the battery. Since there are billions of electrons, we use the term **mean drift velocity** v of the electrons.

Figure 9.10 shows how the mean drift velocity of electrons varies in different situations. We can understand this using the equation:

$$v = \frac{I}{nAe}$$

- If the current increases, the drift velocity v must increase. That is:
  $v \propto I$
- If the wire is thinner, the electrons move more quicklyfor a given current. That is:
  $v \propto \frac{1}{A}$
  There are fewer electrons in a thinner piece of wire, so an individual electron must travel more quickly.
- In a material with a lower density of electrons (smaller n), the mean drift velocity must be greater for a given current. That is:
  $v \propto \frac{1}{n}$

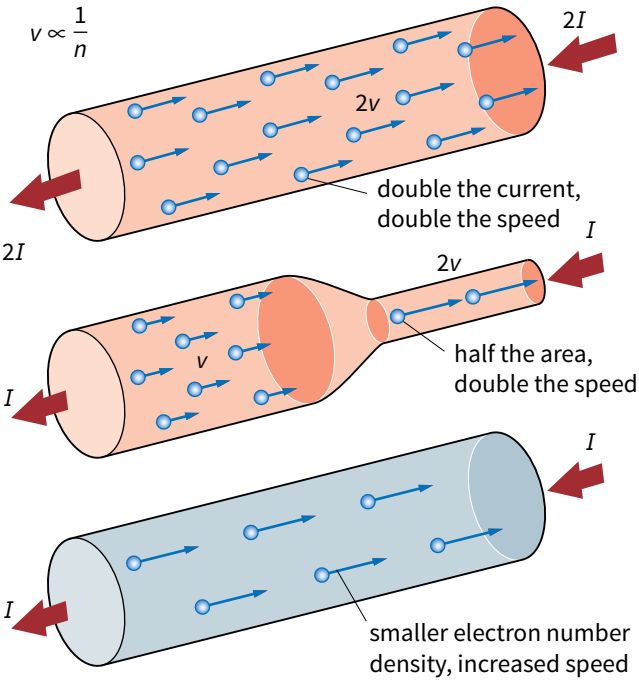


Figure 9.10 The mean drift velocity of electrons depends on the current, the cross-sectional area and the electron density of the material.

QUESTIONS

9 Calculate the current in a gold wire of cross-sectional area $2.0\,mm^2$ when the mean drift velocity of the electrons in the wire is $0.10\,mm\,s^{-1}$. The electron number density for gold is $5.9 \times 10^{28}\,m^{-3}$.

10 Calculate the mean drift velocity of electrons in a copper wire of diameter 1.0 mm with a current of 5.0 A. The electron number density for copper is $8.5 \times 10^{28}\,m^{-3}$.

11 A length of copper wire is joined in series to a length of silver wire of the same diameter. Both wires have a current in them when connected to a battery. Explain how the mean drift velocity of the electrons will change as they travel from the copper into the silver. Electron number densities:

copper $n = 8.5 \times 10^{28}\,m^{-3}$

silver $n = 5.9 \times 10^{28}\,m^{-3}$.

It may help you to picture how the drift velocity of electrons changes by thinking about the flow of water in a river. For a high rate of flow, the water moves fast – this corresponds to a greater current I. If the course of the river narrows, it speeds up – this corresponds to a smaller cross-sectional area A.

Metals have a high electron number density – typically of the order of 10^{28} or $10^{29}\,m^{-3}$. Semiconductors, such as silicon and germanium, have much lower values of n – perhaps $10^{23}\,m^{-3}$. In a semiconductor, electron mean drift velocities are typically a million times greater than those in metals for the same current. Electrical insulators, such as rubber and plastic, have very few conduction electrons per unit volume to act as charge carriers.

The meaning of voltage

The term **voltage** is often used in a rather casual way. In everyday life, the word is used in a less scientific and often incorrect sense – for example, 'A big voltage can go through you and kill you.' In this section, we will consider a bit more carefully just what we mean by voltage and potential difference in relation to electric circuits.

Look at the simple circuit in Figure 9.11. Assume the power supply has negligible internal resistance. (We look at internal resistance later in Chapter 11). The three voltmeters are measuring three voltages or potential differences. With the switch open, the voltmeter placed across the supply measures 12 V. With the switch closed, the voltmeter across the power supply still measures 12 V and the voltmeters placed across the resistors measure 8 V and 4 V. You will not be surprised to see that the voltage across the power supply is equal to the sum of the voltages across the resistors.

Earlier in this chapter we saw that electric current is the rate of flow of electric charge. Figure 9.12 shows the same circuit as in Figure 9.11, but here we are looking at the movement of one coulomb (1 C) of charge round the circuit. Electrical energy is transferred to the charge by the power supply. The charge flows round the circuit, transferring some of its electrical energy to heat in the first resistor, and the rest to heat in the second resistor.

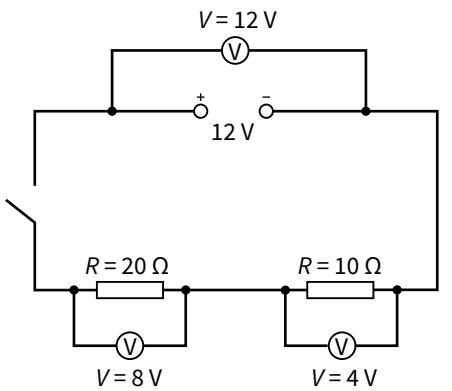


Figure 9.11 Measuring voltages in a circuit. Note that each voltmeter is connected **across** the component.

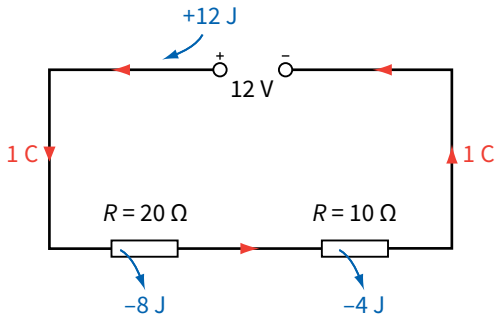


Figure 9.12 Energy transfers as 1 C of charge flows round a circuit. This circuit is the same as that shown in Figure 9.11.

The voltmeter readings indicate the energy transferred to the component by each unit of charge. The voltmeter placed across the power supply measures the e.m.f. of the supply, whereas the voltmeters placed across the resistors measure the potential difference (p.d.) across these components. The terms e.m.f. and potential difference have different meanings – so you have to be very vigilant.

The term **potential difference** is used when charges **lose** energy by transferring electrical energy to other forms of energy in a component. Potential difference, V, is defined as the energy transferred per unit charge.

> The potential difference between two points, A and B, is the energy per unit charge as charge moves from point A to point B.

A power supply or a battery transfers energy to electrical charges in a circuit. The **e.m.f.**, E, of the supply is also defined as the energy transferred per unit charge.

> e.m.f. is defined as the total work done per unit charge when charge flows round a complete circuit.

Note that e.m.f. stands for electromotive force. This is a misleading term. It has nothing at all to do with force. This term is a legacy from the past and we are stuck with it! It is best to forget where it comes from and simply use the term e.m.f.

Electrical resistance

If you connect a lamp to a battery, a current in the lamp causes it to glow. But what determines the size of the current? This depends on two factors:

- the potential difference or voltage *V* across the lamp – the greater the potential difference, the greater the current for a given lamp
- the resistance *R* of the lamp – the greater the resistance, the smaller the current for a given potential difference.

Now we need to think about the meaning of **electrical resistance**. The resistance of any component is defined as the ratio of the potential difference to the current. As a word equation, this is written as:

$$\text{resistance} = \frac{\text{potential difference}}{\text{current}}$$

or

$$R = \frac{V}{I}$$

where R is the resistance of the component, V is the potential difference across the component and I is the current in the component. You can rearrange the equation above to give:

$$I = \frac{V}{R} \quad \text{or} \quad V = IR$$

Table 9.2 summarises these quantities and their units.

Quantity	Symbol for quantity	Unit	Symbol for unit
current	*I*	ampere (amp)	A
voltage (p.d., e.m.f.)	*V*	volt	V
resistance	*R*	ohm	Ω

Table 9.2 Basic electrical quantities, their symbols and SI units. Take care to understand the difference between *V* (in italics) meaning the quantity voltage and V meaning the unit volt.

Defining the ohm

The unit of resistance, the ohm, can be determined from the equation that defines resistance:

$$\text{resistance} = \frac{\text{potential difference}}{\text{current}}$$

The ohm is equivalent to '1 volt per ampere'. That is:

$$1\,\Omega = 1\,\text{V}\,\text{A}^{-1}$$

The ohm is the resistance of a component when a potential difference of 1 volt drives a current of 1 ampere through it.

WORKED EXAMPLE

4 Calculate the current in a lamp given that its resistance is 15 Ω and the potential difference across its ends is 3.0 V.

Step 1 Here we have $V = 3.0\,\text{V}$ and $R = 15\,\Omega$.

Step 2 Substituting in $I = \frac{V}{R}$ gives:

current $I = \frac{3.0}{15} = 0.20\,\text{A}$

So the current in the lamp is 0.20 A.

QUESTIONS

12 A car headlamp bulb has a resistance of 36 Ω. Calculate the current in the lamp when connected to a '12 V' battery.

13 You can buy lamps of different brightness to fit in light fittings at home (Figure 9.13). A '100 watt' lamp glows more brightly than a '60 watt' lamp. Explain which of the lamps has the higher resistance.



Figure 9.13 Both of these lamps work from the 230 V mains supply, but one has a higher resistance than the other. For Question 13.

14 **a** Calculate the potential difference across a motor carrying a current of 1.0 A and having a resistance of 50 Ω.

b Calculate the potential difference across the same motor when the current is doubled. Assume its resistance remains constant.

15 Calculate the resistance of a lamp carrying a current of 0.40 A when connected to a 230 V supply.

BOX 9.1: Determining resistance

As we have seen, the equation for resistance is:

$$R = \frac{V}{I}$$

To determine the resistance of a component, we therefore need to measure both the potential difference V across it and the current I through it. To measure the current, we need an ammeter. To measure the potential difference, we need a voltmeter. Figure 9.14 shows how these meters should be connected to determine the resistance of a metallic conductor, such as a length of wire.

- The ammeter is connected **in series** with the conductor, so that there is the same current in both.
- The voltmeter is connected across (**in parallel** with) the conductor, to measure the potential difference across it.

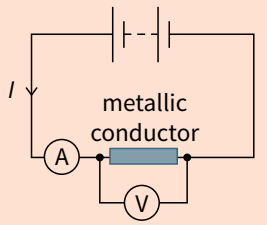


Figure 9.14 Connecting an ammeter and a voltmeter to determine the resistance of a metallic conductor in a circuit.

QUESTION

16 In Figure 9.14 the reading on the ammeter is 2.4 A and the reading on the voltmeter is 6.0 V. Calculate the resistance of the metallic conductor.

Electrical power

The rate at which energy is transferred is known as power. Power P is measured in watts (W). (If you are not sure about this, refer back to Chapter 5, where we looked at the concept of power in relation to forces and work done.)

$$\text{power} = \frac{\text{energy transferred}}{\text{time taken}}$$

$$P = \frac{W}{\Delta t}$$

where P is the power and W is the energy transferred in a time Δt. Take care not to confuse W for energy transferred or work done with W for watts.

The rate at which energy is transferred in an electrical component is related to two quantities:

- the current I in the component
- the potential difference V across the component.

We can derive an equation for electrical power from the equations we have met so far. The amount of energy W transferred by a charge ΔQ when it moves through a potential difference V is given by:

$$W = V\Delta Q$$

Hence:

$$P = \frac{W}{\Delta t} = \frac{V\Delta Q}{\Delta t} = V\left(\frac{\Delta Q}{\Delta t}\right)$$

The ratio of charge to time, $\frac{\Delta Q}{\Delta t}$, is the current I in the component. Therefore:

$$P = VI$$

As a word equation, we have:

power = potential difference × current

and in units:

watts = amps × volts

WORKED EXAMPLE

5 Calculate the rate at which energy is transferred by a 230 V mains supply which provides a current of 8.0 A to an electric heater.

Step 1 Use the equation for power:

$P = VI$

with $V = 230$ V and $I = 8.0$ A.

Step 2 Substitute values:

$P = 8 \times 230 = 1840$ W (1.84 kW)

QUESTIONS

17 Calculate the current in a 60 W light bulb when it is connected to a 230 V power supply.

18 A large power station supplies electrical energy to the grid at a voltage of 25 kV. Calculate the output power of the station when the current it supplies is 40 kA.

Fuses

A fuse is a device which is fitted in an electric circuit; it is usually there to protect the wiring from excessive currents. For example, the fuses in a domestic fuse box will 'blow' if the current is too large. High currents cause wires to get hot, and this can lead to damaged wires, fumes from melting insulation, and even fires.

Fuses (Figure 9.15) are usually marked with their current rating; that is, the maximum current which they will permit. Inside the fuse cartridge is a thin wire which gets hot and melts if the current exceeds this value. This breaks the circuit and stops any hazardous current. Worked example 7 shows how an appropriate fuse is chosen.



Figure 9.15 Fuses of different current ratings.

QUESTION

19 An electric cooker is usually connected to the mains supply in a separate circuit from other appliances, because it draws a high current. A particular cooker is rated at 10 kW, 230 V.

a Calculate the current in the cooker when it is fully switched on.

b Suggest a suitable current rating for the fuse for this cooker.

WORKED EXAMPLES

6 a A power station produces 20 MW of power at a voltage of 200 kV. Calculate the current supplied to the grid cables.

Step 1 Here we have P and V and we have to find I, so we can use $P = VI$.

Step 2 Rearranging the equation and substituting the values we know gives:

$$\text{current } I = \frac{P}{V} = \frac{20 \times 10^6}{200 \times 10^3} = 100\,\text{A}$$

Hint: *Remember to convert megawatts into watts and kilovolts into volts.*

So the power station supplies a current of 100 A.

b The grid cables are 15 km long, with a resistance per unit length of $0.20\,\Omega\,\text{km}^{-1}$. How much power is wasted as heat in these cables?

Step 1 First we must calculate the resistance of the cables:

resistance $R = 15\,\text{km} \times 0.20\,\Omega\,\text{km}^{-1} = 3.0\,\Omega$

Step 2 Now we know I and R and we want to find P. We can use $P = I^2R$:

power wasted as heat, $P = I^2R = (100)^2 \times 3.0$

$= 3.0 \times 10^4\,\text{W}$

$= 30\,\text{kW}$

Hence, of the 20 MW of power produced by the power station, 30 kW is wasted – just 0.15%.

7 An electric kettle is rated at 2.5 kW, 230 V. Determine a suitable current rating of the fuse to put in the three-pin plug. Choose from 1 A, 5 A, 13 A, 30 A.

Step 1 Calculate the current in the kettle in normal operation. Rearranging $P = VI$ to make I the subject gives:

$$I = \frac{P}{V}$$

So:

$$I = \frac{2500}{230} = 10.9\,\text{A}$$

Step 2 Now we know that the normal current in the kettle is 10.9 A. We must choose a fuse with a **slightly higher** rating than this. Therefore the value of the fuse rating is 13 A.

Hint: *A 5 A fuse would not be suitable because it would melt as soon as the kettle is switched on. A 30 A fuse would allow more than twice the normal current before blowing, which would not provide suitable protection.*

Power and resistance

A current I in a resistor of resistance R transfers energy to it. The resistor dissipates the energy as heat. The p.d. V across the resistor is given by $V = IR$. Combining this with the equation for power, $P = VI$, gives us two further forms of the equation for power dissipated in the resistor:

$$P = I^2R$$

$$P = \frac{V^2}{R}$$

Which form of the equation we use in any particular situation depends on the information we have available to us. This is illustrated in Worked examples 6a and 6b, which relate to a power station and to the grid cables which lead from it (Figure 9.16).



Figure 9.16 A power station and electrical transmission lines. How much electrical power is lost as heat in these cables? (See Worked examples 6a and 6b.)

QUESTIONS

20 A calculator is powered by a 3.0 V battery. The calculator's resistance is 20 kΩ. Calculate the power transferred to the calculator.

21 An energy-efficient light bulb is labelled '230 V, 15 W'. This means that when connected to the 230 V mains supply it is fully lit and changes electrical energy to heat and light at the rate of 15 W. Calculate:

- **a** the current which flows through the bulb when fully lit
- **b** its resistance when fully lit.

22 Calculate the resistance of a 100 W light bulb that draws a current of 0.43 A from a power supply.

Calculating energy

We can use the relationship for power as energy transferred per unit time and the equation for electrical power to find the energy transferred in a circuit.
Since:

power = current × voltage

and:

energy = power × time

we have:

energy transferred = current × voltage × time

$$W = IV\Delta t$$

Working in SI units, this gives energy transferred in joules.

QUESTIONS

23 A 12 V car battery can supply a current of 10 A for 5.0 hours. Calculate how many joules of energy the battery transfers in this time.

24 A lamp is operated for 20 s. The current in the lamp is 10 A. In this time, it transfers 400 J of energy to the lamp. Calculate:

- **a** how much charge flows through the lamp
- **b** how much energy each coulomb of charge transfers to the lamp
- **c** the p.d. across the lamp.

Summary

- Electric current is the rate of flow of charge. In a metal this is due to the flow of electrons. In an electrolyte, the flow of positive and negative ions produces the current.
- The direction of conventional current is from positive to negative; the direction of electron flow is from negative to positive.
- The SI unit of charge is the coulomb (C). One coulomb is the charge which passes a point when a current of 1 A flows for 1 s.

  charge = current × time $\quad (\Delta Q = I\Delta t)$
- The elementary charge $e = 1.6 \times 10^{-19}$ C.
- The current I in a conductor of cross-sectional area A depends on the mean drift velocity v of the charge carriers and their number density n.

  $I = nAvq$
- The term potential difference (p.d.) is used when charges lose energy in a component. It is defined as the energy transferred per unit charge.

  $V = \frac{W}{\Delta Q}$ or $W = V\Delta Q$
- The term electromotive force (e.m.f.) is used when charges gain electrical energy from a battery or similar device. It is also defined as the energy transferred per unit charge.

  $E = \frac{W}{\Delta Q}$ or $W = E\Delta Q$
- A volt is a joule per coulomb. That is, $1\text{ V} = 1\text{ J C}^{-1}$.
- Power is the rate of energy transfer. In electrical terms, power is the product of voltage and current. That is, $P = VI$.
- Resistance is defined as the ratio of voltage to current. That is:

  $\text{resistance} = \frac{\text{voltage}}{\text{current}} \quad \left(R = \frac{V}{I}\right)$
- The resistance of a component is 1 ohm when a potential difference of 1 volt is produced per ampere.
- For a resistance R, the power dissipated is given by:

  $P = I^2R$ or $P = \frac{V^2}{R}$
- Energy transferred in a circuit in a time Δt is given by:

  $W = IV\Delta t$

End-of-chapter questions

1 Calculate the charge which passes through a lamp when there is a current of 150 mA for 40 minutes. [3]

2 A generator produces a current of 40 A. How long will it take for a total of 2000 C to flow through the output? [2]

3 In a lightning strike there is an average current of 30 kA, which lasts for 2000 μs. Calculate the charge which is transferred in this process. [3]

4 **a** A lamp of resistance 15 Ω is connected to a battery of e.m.f. 4.5 V. Calculate the current in the lamp. [2]

b Calculate the resistance of the filament of an electric heater which takes a current of 6.5 A when it is connected across a mains supply of 230 V. [2]

c Calculate the voltage which is require to drive a current of 2.4 A through a wire of resistance 3.5 Ω. [2]

5 A battery of e.m.f. 6 V produces a steady current of 2.4 A for 10 minutes. Calculate:

a the charge which it supplied [2]

b the energy that it transferred. [2]

6 Calculate the energy gained by an electron when it is accelerated through a potential difference of 50 kV. (Charge on the electron = -1.6×10^{-19} C.) [2]

7 A woman has available 1 A, 3 A, 5 A, 10 A and 13 A fuses. Explain which fuse she should use for a 120 V, 450 W hairdryer. [3]

8 Figure 9.17 shows the electrolysis of copper chloride.

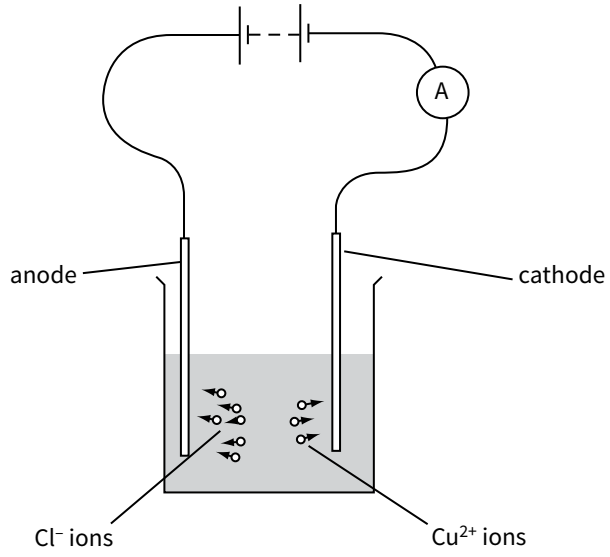


Figure 9.17 For End-of-chapter Question 8.

a **i** On a copy of the diagram, mark the direction of the conventional current in the electrolyte. Label it conventional current. [1]

ii Mark the direction of the electron flow in the connecting wires. Label this electron flow. [1]

b In a time period of 8 minutes, 3.6×10^{16} chloride (Cl^-) ions are neutralised and liberated at the anode and 1.8×10^{16} copper (Cu^{2+}) ions are neutralised and deposited on the cathode.

i Calculate the total charge passing through the electrolyte in this time. [2]

ii Calculate the current in the circuit. [2]

9 Figure 9.18 shows an electron tube. Electrons moving from the cathode to the anode constitute a current. The current in the ammeter is 4.5 mA.

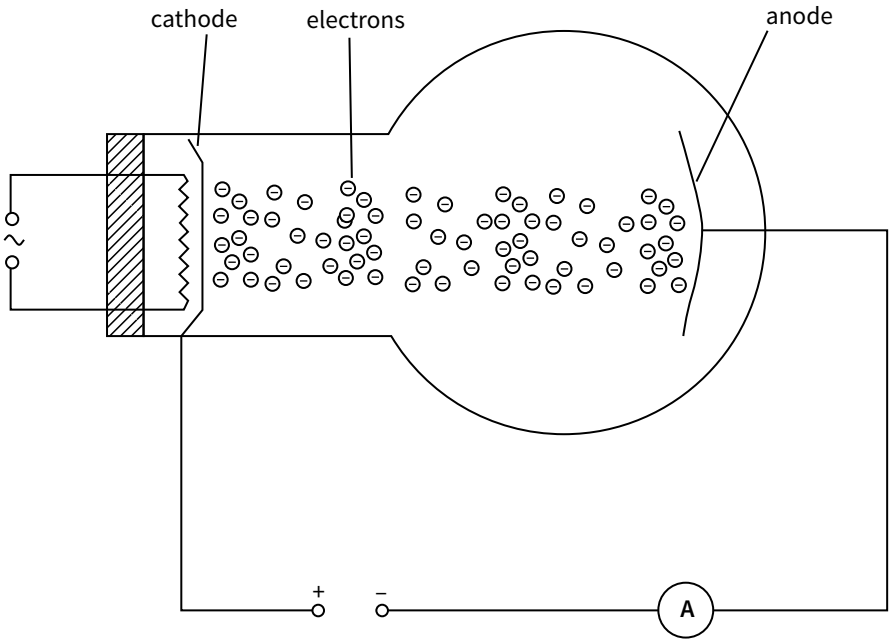


Figure 9.18 For End-of-chapter Question 9.

a Calculate the charge passing through the ammeter in 3 minutes. [3]

b Calculate the number of electrons which hit the anode in 3 minutes. [3]

c The potential difference between the cathode and the anode is 75 V. Calculate the energy gained by an electron as it travels from the cathode to the anode. [2]

10 A length of copper track on a printed circuit board has a cross-sectional area of $5.0 \times 10^{-8}\,m^2$. The current in the track is 3.5 mA. You are provided with some useful information about copper:

$1\,m^3$ of copper has a mass of $8.9 \times 10^3\,kg$

54 kg of copper contains 6.0×10^{26} atoms

In copper, there is roughly one electron liberated from each copper atom.

a Show that the electron number density n for copper is about $10^{29}\,m^{-3}$. [2]

b Calculate the mean drift velocity of the electrons. [3]

11 **a** Explain the difference between **potential difference** and **e.m.f.** [2]

b A battery has negligible internal resistance, an e.m.f. of 12.0 V and a capacity of 100 A h (ampere-hours). Calculate:

i the total charge that it can supply [2]

ii the total energy that it can transfer. [2]

c The battery is connected to a 27 W lamp. Calculate the resistance of the lamp. [3]

12 Some electricity-generating companies use a unit called the kilowatt-hour (kW h) to calculate energy bills. 1 kW h is the energy a kilowatt appliance transfers in 1 hour.

a Show that 1 kW h is equal to 3.6 MJ. [2]

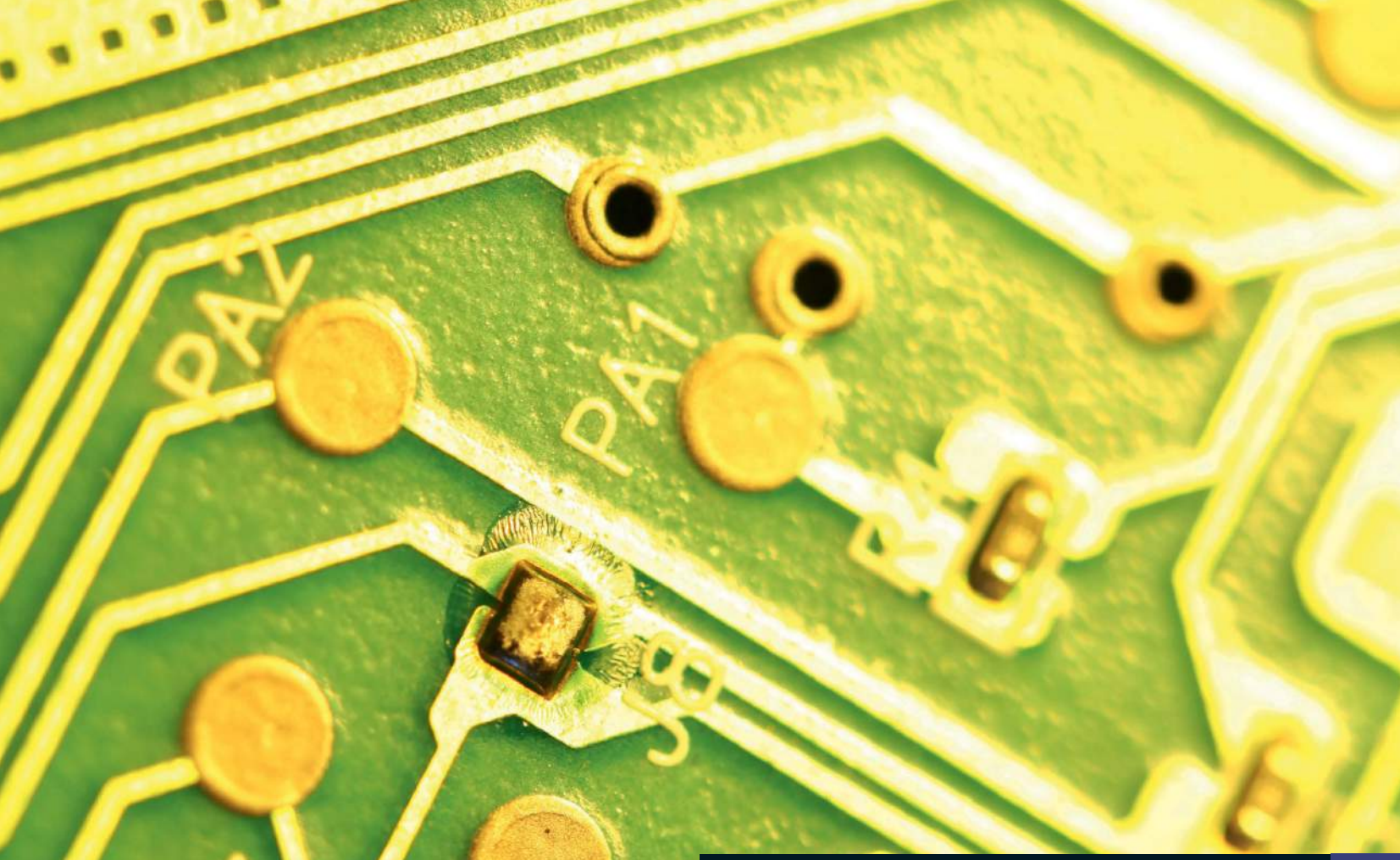
b An electric shower heater is rated at 230 V, 9.5 kW.

i Calculate the current it will take from the mains supply. [2]

ii Suggest why the shower requires a separate circuit from other appliances. [1]

iii Suggest a suitable current rating for the fuse in this circuit. [1]

c Calculate the energy transferred when a boy uses the shower for 5 minutes. [2]



Chapter 10: Kirchhoff's laws

Learning outcomes

You should be able to:

- recall and apply Kirchhoff's laws
- use Kirchhoff's laws to derive the formulae for the combined resistance of two or more resistors in series and in parallel
- recognise that ammeters are connected in series within a circuit and therefore should have low resistance
- recognise that voltmeters are connected in parallel across a component, or components, and therefore should have high resistance

Circuit design

Over the years, electrical circuits have become increasingly complex, with more and more components combining to achieve very precise results (Figure 10.1). Such circuits typically include power supplies, sensing devices, potential dividers and output devices. At one time, circuit designers would start with a simple circuit and gradually modify it until the desired result was achieved. This is impossible today when circuits include many hundreds or thousands of components.

Instead, electronics engineers (Figure 10.2) rely on computer-based design software which can work out the effect of any combination of components. This is only possible because computers can be programmed with the equations that describe how current and voltage behave in a circuit. These equations, which include Ohm's law and Kirchhoff's two laws, were established in the 18th century, but they have come into their own in the 21st century through their use in computer-aided design (CAD) systems.



Figure 10.1 A complex electronic circuit – this is the circuit board which controls a computer's hard drive.



Figure 10.2 A computer engineer uses a computer-aided design (CAD) software tool to design a circuit which will form part of a microprocessor, the device at the heart of every computer.

Kirchhoff's first law

You should be familiar with the idea that current may divide up where a circuit splits into two separate branches. For example, a current of 5.0 A may split at a junction or a point in a circuit into two separate currents of 2.0 A and 3.0 A. The total amount of current remains the same after it splits. We would not expect some of the current to disappear, or extra current to appear from nowhere. This is the basis of **Kirchhoff's first law**, which states that:

> The sum of the currents entering any point in a circuit is equal to the sum of the currents leaving that same point.

This is illustrated in Figure 10.3. In the first part, the current into point P must equal the current out, so:

$$I_1 = I_2$$

In the second part of the figure, we have one current coming into point Q, and two currents leaving. The current divides at Q. Kirchhoff's first law gives:

$$I_1 = I_2 + I_3$$

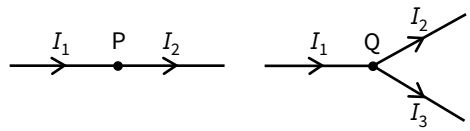


Figure 10.3 Kirchhoff's first law: current is conserved because charge is conserved.

Kirchhoff's first law is an expression of the **conservation of charge**. The idea is that the total amount of charge entering a point must exit the point. To put it another way, if a billion electrons enter a point in a circuit in a time interval of 1.0 s, then one billion electrons must exit this point in 1.0 s. The law can be tested by connecting ammeters at different points in a circuit where the current divides. You should recall that an ammeter must be connected in series so the current to be measured passes **through** it.

QUESTIONS

1 Use Kirchhoff's first law to deduce the value of the current I in Figure 10.4.

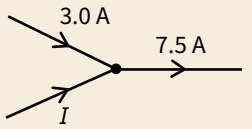


Figure 10.4 For Question 1.

2 In Figure 10.5, calculate the current in the wire X. State the direction of this current (towards P or away from P).

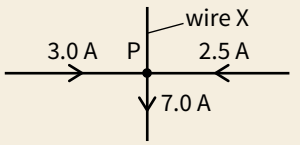


Figure 10.5 For Question 2.

Formal statement of Kirchhoff's first law

We can write Kirchhoff's first law as an equation:

$$\Sigma I_{in} = \Sigma I_{out}$$

Here, the symbol Σ (Greek letter sigma) means 'the sum of all', so ΣI_{in} means 'the sum of all currents entering into a point' and ΣI_{out} means 'the sum of all currents leaving that point'. This is the sort of equation which a computer program can use to predict the behaviour of a complex circuit.

Kirchhoff's second law

This law deals with e.m.f.s and voltages in a circuit. We will start by considering a simple circuit which contains a cell and two resistors of resistances R_1 and R_2 (Figure 10.8). Since this is a simple series circuit, the current I must be the same all the way around, and we need not concern ourselves further with Kirchhoff's first law. For this circuit, we can write the following equation:

$$E = IR_1 + IR_2$$

e.m.f. of battery = sum of p.d.s across the resistors

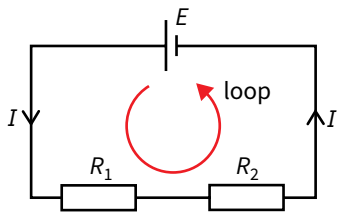


Figure 10.8 A simple series circuit.

You should not find these equations surprising. However, you may not realise that they are a consequence of applying **Kirchhoff's second law** to the circuit. This law states that:

> The sum of the e.m.f.s around any loop in a circuit is equal to the sum of the p.d.s around the loop.

QUESTIONS

3 Calculate ΣI_{in} and ΣI_{out} in Figure 10.6. Is Kirchhoff's first law satisfied?

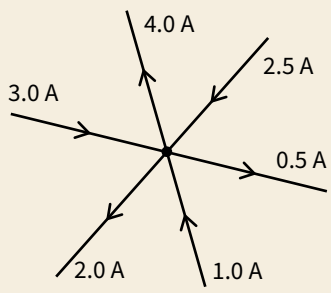


Figure 10.6 For Question 3.

4 Use Kirchhoff's first law to deduce the value and direction of the current I in Figure 10.7.

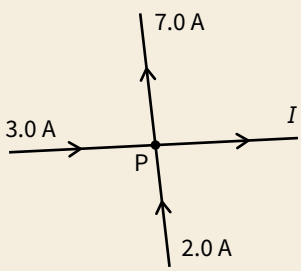


Figure 10.7 For Question 4.

WORKED EXAMPLE

1 Use Kirchhoff's laws to find the current in the circuit in Figure 10.9.

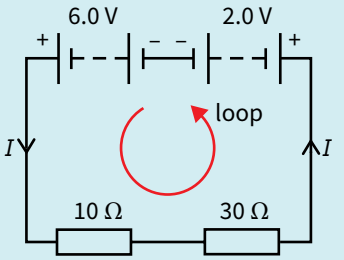


Figure 10.9 A circuit with two opposing batteries.

This is a series circuit so the current is the same all the way round the circuit.

Step 1 We calculate the sum of the e.m.f.s:

sum of e.m.f.s = 6.0 V − 2.0 V = 4.0 V

The batteries are connected in opposite directions so we must consider one of the e.m.f.s as negative.

Step 2 We calculate the sum of the p.d.s.

sum of p.d.s = $(I \times 10) + (I \times 30) = 40I$

Step 3 We equate these:

$4.0 = 40I$

and so $I = 0.1$ A

No doubt, you could have solved this problem without formally applying Kirchhoff's second law, but you will find that in more complex problems the use of these laws will help you to avoid errors.

You will see later (page 148) that Kirchhoff's second law is an expression of the conservation of energy. We shall look at another example of how this law can be applied, and then look at how it can be applied in general.

QUESTION

5 Use Kirchhoff's second law to deduce the p.d. across the resistor of resistance R in the circuit shown in Figure 10.10, and hence find the value of R. (Assume the battery of e.m.f. 10 V has negligible internal resistance.)

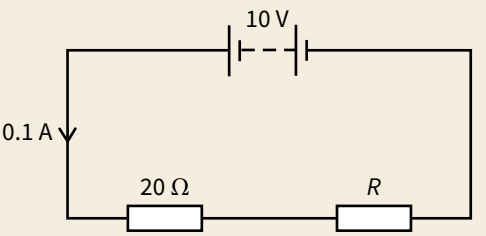


Figure 10.10 Circuit for Question 5.

An equation for Kirchhoff's second law

In a similar manner to the formal statement of the first law, the second law can be written as an equation:

$$\Sigma E = \Sigma V$$

where ΣE is the sum of the e.m.f.s and ΣV is the sum of the potential differences.

Applying Kirchhoff's laws

Figure 10.11 shows a more complex circuit, with more than one 'loop'. Again there are two batteries and two resistors. The problem is to find the current in each resistor. There are several steps in this; Worked example 2 shows how such a problem is solved.

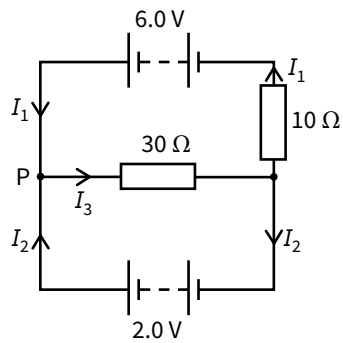


Figure 10.11 Kirchhoff's laws are needed to determine the currents in this circuit.

Signs and directions

Caution is necessary when applying Kirchhoff's second law. You need to take account of the ways in which the sources of e.m.f. are connected and the directions of the currents. Figure 10.12 shows one loop from a larger complicated circuit to illustrate this point. Only the components and currents in this particular are shown.

WORKED EXAMPLE

2 Calculate the current in each of the resistors in the circuit shown in Figure 10.11.

Step 1 Mark the currents flowing. The diagram shows I_1, I_2 and I_3.

Hint: *It does not matter if we mark these flowing in the wrong directions, as they will simply appear as negative quantities in the solutions.*

Step 2 Apply Kirchhoff's first law. At point P, this gives:

$I_1 + I_2 = I_3$ (1)

Step 3 Choose a loop and apply Kirchhoff's second law. Around the upper loop, this gives:

$6.0 = (I_3 \times 30) + (I_1 \times 10)$ (2)

Step 4 Repeat step 3 around other loops until there are the same number of equations as unknown currents. Around the lower loop, this gives:

$2.0 = I_3 \times 30$ (3)

We now have three equations with three unknowns (the three currents).

Step 5 Solve these equations as simultaneous equations. In this case, the situation has been chosen to give simple solutions. Equation 3 gives $I_3 = 0.067\,\text{A}$, and substituting this value in equation 2 gives $I_1 = 0.400\,\text{A}$. We can now find I_2 by substituting in equation 1:

$I_2 = I_3 - I_1 = 0.067 - 0.400 = -0.333\,\text{A}$

$\approx -0.33\,\text{A}$

Thus I_2 is negative – it is in the opposite direction to the arrow shown in Figure 11.11.

Note that there is a third 'loop' in this circuit; we could have applied Kirchhoff's second law to the outermost loop of the circuit. This would give a fourth equation:

$6 - 2 = I_1 \times 10$

However, this is not an independent equation; we could have arrived at it by subtracting equation 3 from equation 2.

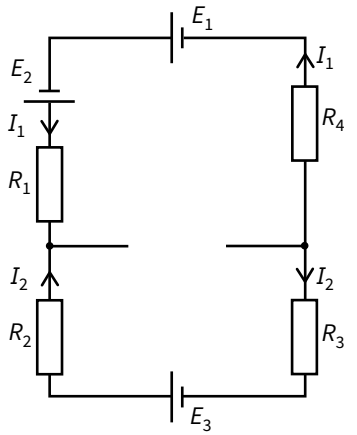


Figure 10.12 A loop extracted from a complicated circuit.

e.m.f.s

Starting with the cell of e.m.f. E_1 and working **anticlockwise** around the loop (because E_1 is 'pushing current' anticlockwise):

sum of e.m.f.s $= E_1 + E_2 - E_3$

Note that E_3 is opposing the other two e.m.f.s.

p.d.s

Starting from the same point, and working **anticlockwise** again:

sum of p.d.s $= I_1R_1 - I_2R_2 - I_2R_3 + I_1R_4$

Note that the direction of current I_2 is clockwise, so the p.d.s that involve I_2 are negative.

QUESTION

6 You can use Kirchhoff's second law to find the current I in the circuit shown in Figure 10.13. Choosing the best loop can simplify the problem.

a Which loop in the circuit should you choose?

b Calculate the current I.

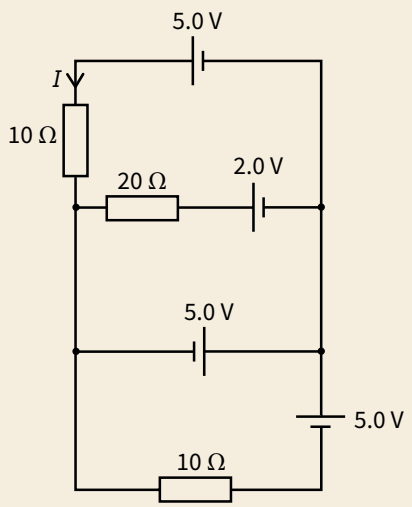


Figure 10.13 Careful choice of a suitable loop can make it easier to solve problems like this. For Question 6.

QUESTION

7 Use Kirchhoff's second law to deduce the resistance *R* of the resistor shown in the circuit loop of Figure 10.14.

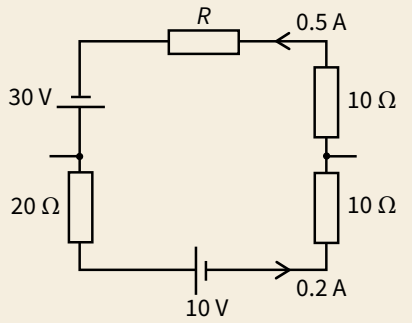


Figure 10.14 For Question 7.

Conservation of energy

Kirchhoff's second law is a consequence of the principle of conservation of energy. If a charge, say 1 C, moves around the circuit, it **gains** energy as it moves through each source of e.m.f. and loses energy as it passes through each p.d. If the charge moves all the way round the circuit, so that it ends up where it started, it must have the same energy at the end as at the beginning. (Otherwise we would be able to create energy from nothing simply by moving charges around circuits.) So:

energy gained passing through sources of e.m.f.
= energy lost passing through components with p.d.s

You should recall that an e.m.f. in volts is simply the energy gained per 1 C of charge as it passes through a source. Similarly, a p.d. is the energy lost per 1 C as it passes through a component.

1 volt = 1 joule per coulomb

Hence we can think of Kirchhoff's second law as:

energy gained per coulomb around loop
= energy lost per coulomb around loop

Here is another way to think of the meaning of e.m.f. A 1.5 V cell gives 1.5 J of energy to each coulomb of charge which passes through it. The charge then moves round the circuit, transferring the energy to components in the circuit. The consequence is that, by driving 1 C of charge around the circuit, the cell transfers 1.5 J of energy. Hence the e.m.f. of a source simply tells us the amount of energy (in joules) transferred by the source in driving unit charge (1 C) around a circuit.

QUESTIONS

8 Use the idea of the energy gained and lost by a 1 C charge to explain why two 6 V batteries connected together in series can give an e.m.f. of 12 V or 0 V, but connected in parallel they give an e.m.f. of 6 V.

9 Apply Kirchhoff's laws to the circuit shown in Figure 10.15 to determine the current that will be shown by the ammeters A_1, A_2 and A_3.

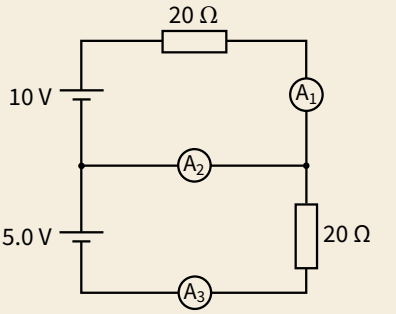


Figure 10.15 Kirchhoff's laws make it possible to deduce the ammeter readings.

Resistor combinations

You are already familiar with the formulae used to calculate the combined resistance *R* of two or more resistors connected in series or in parallel. To derive these formulae we have to make use of Kirchhoff's laws.

Resistors in series

Take two resistors of resistances R_1 and R_2 connected in series (Figure 10.16). According to Kirchhoff's first law, the current in each resistor is the same. The p.d. *V* across the combination is equal to the sum of the p.d.s across the two resistors:

$$V = V_1 + V_2$$

Since $V = IR$, $V_1 = IR_1$ and $V_2 = IR_2$, we can write:

$$IR = IR_1 + IR_2$$

Cancelling the common factor of current *I* gives:

$$R = R_1 + R_2$$

For three or more resistors, the equation for total resistance *R* becomes:

$$R = R_1 + R_2 + R_3 + \cdots$$

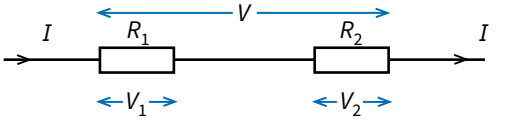


Figure 10.16 Resistors in series.

QUESTIONS

10 Calculate the combined resistance of two 5 Ω resistors and a 10 Ω resistor connected in series.

11 The cell shown in Figure 10.17 provides an e.m.f. of 2.0 V. The p.d. across one lamp is 1.2 V. Determine the p.d. across the other lamp.

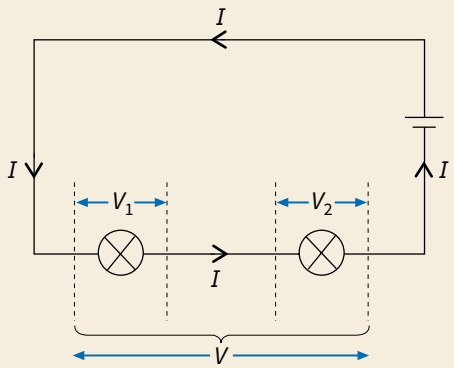


Figure 10.17 A series circuit for Question 11.

12 You have five 1.5 V cells. How would you connect all five of them in series to give an e.m.f. of:

a 7.5 V? **b** 1.5 V? **c** 4.5 V?

Resistors in parallel

For two resistors of resistances R_1 and R_2 connected in parallel (Figure 10.18), we have a situation where the current divides between them. Hence, using Kirchhoff's first law, we can write:

$$I = I_1 + I_2$$

If we apply Kirchhoff's second law to the loop that contains the two resistors, we have:

$$I_1R_1 - I_2R_2 = 0\,\text{V}$$

(because there is no source of e.m.f. in the loop).

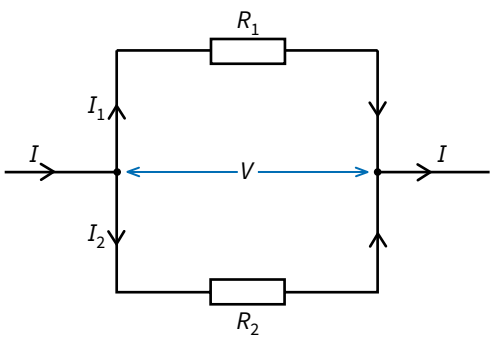


Figure 10.18 Resistors connected in parallel.

This equation states that the two resistors have the same p.d. V across them. Hence we can write:

$$I = \frac{V}{R}$$

$$I_1 = \frac{V}{R_1}$$

$$I_2 = \frac{V}{R_2}$$

Substituting in $I = I_1 + I_2$ and cancelling the common factor V gives:

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$$

For three or more resistors, the equation for total resistance R becomes:

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \cdots$$

To summarise, when components are connected in parallel:

- all have the same p.d. across their ends
- the current is shared between them
- we use the reciprocal formula to calculate their combined resistance.

WORKED EXAMPLE

3 Two 10 Ω resistors are connected in parallel. Calculate the total resistance.

Step 1 We have $R_1 = R_2 = 10\,\Omega$, so:

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$$

$$\frac{1}{R} = \frac{1}{10} + \frac{1}{10} = \frac{2}{10} = \frac{1}{5}$$

Step 2 Inverting both sides of the equation gives:

$R = 5\,\Omega$

Hint: *Take care not to forget this step! Nor should you write* $\frac{1}{R} = \frac{1}{5} = 5\,\Omega$, *as then you are saying* $\frac{1}{5} = 5$*).*

You can also determine the resistance as follows:

$$R = (R_1^{-1} + R_2^{-1})^{-1}$$

$$= (10^{-1} + 10^{-1})^{-1} = 5\,\Omega$$

QUESTIONS

13 Calculate the total resistance of four 10 Ω resistors connected in parallel.

14 Calculate the resistances of the following combinations:

a 100 Ω and 200 Ω in series

b 100 Ω and 200 Ω in parallel

c 100 Ω and 200 Ω in series and this in parallel with 200 Ω.

15 Calculate the current drawn from a 12 V battery of negligible internal resistance connected to the ends of the following:

a 500 Ω resistor

b 500 Ω and 1000 Ω resistors in series

c 500 Ω and 1000 Ω resistors in parallel.

16 You are given one 200 Ω resistor and two 100 Ω resistors. What total resistances can you obtain by connecting some, none, or all of these resistors in various combinations?

Solving problems with parallel circuits

Here are some useful ideas which may prove helpful when you are solving problems with parallel circuits (or checking your answers to see whether they seem reasonable).

- When two or more resistors are connected in parallel, their combined resistance is smaller than any of their individual resistances. For example, three resistors of 2 Ω, 3 Ω and 6 Ω connected together in parallel have a combined resistance of 1 Ω. This is less than the smallest of the individual resistances. This comes about because, by connecting the resistors in parallel, you are providing extra pathways for the current. Since the combined resistance is lower than the individual resistances, it follows that connecting two or more resistors in parallel will increase the current drawn from a supply. Figure 10.19 shows a hazard which can arise when electrical appliances are connected in parallel.
- When components are connected in parallel, they all have the same p.d. across them. This means that you can often ignore parts of the circuit which are not relevant to your calculation.
- Similarly, for resistors in parallel, you may be able to calculate the current in each one individually, then add them up to find the total current. This may be easier than working out their combined resistance using the reciprocal formula. (This is illustrated in Question 19.)

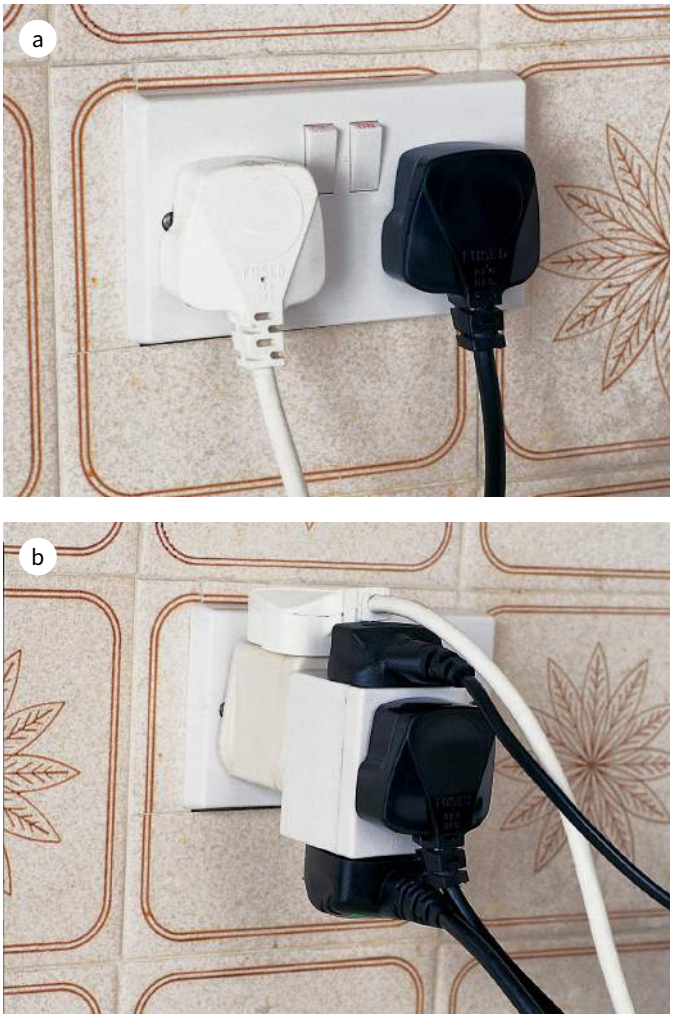


Figure 10.19 **a** Correct use of an electrical socket. **b** Here, too many appliances (resistances) are connected in parallel. This reduces the total resistance and increases the current drawn, to the point where it becomes dangerous.

QUESTION

17 Three resistors of resistances 20 Ω, 30 Ω and 60 Ω are connected together in parallel. Select which of the following gives their combined resistance:

110 Ω, 50 Ω, 20 Ω, 10 Ω

(No need to do the calculation!)

QUESTIONS

18 In the circuit in Figure 10.20 the battery of e.m.f. 10 V has negligible internal resistance. Calculate the current in the 20 Ω resistor shown in the circuit.

19 Determine the current drawn from the battery in Figure 10.20.

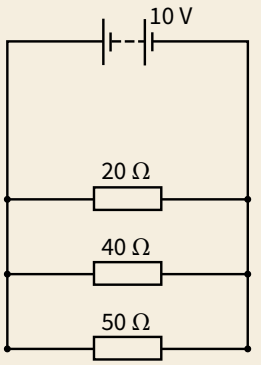


Figure 10.20 Circuit diagram for Questions 18 and 19.

20 What value of resistor must be connected in parallel with a 20 Ω resistor so that their combined resistance is 10 Ω?

21 You are supplied with a number of 100 Ω resistors. Describe how you could combine the minimum number of these to make a 250 Ω resistor.

22 Calculate the current at each point (A–E) in the circuit shown in Figure 10.21.

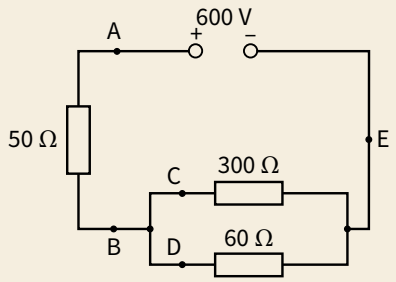


Figure 10.21 For Question 22.

BOX 10.1: Ammeters and voltmeters

Ammeters and voltmeters are connected differently in circuits (Figure 10.22). Ammeters are always connected in series, since they measure the current in a circuit. For this reason, an ammeter should have as low a resistance as possible so that as little energy as possible is dissipated in the ammeter itself. Inserting an ammeter with a higher resistance could significantly reduce the current flowing in the circuit. The ideal resistance of an ammeter is zero. Digital ammeters have very low resistances.

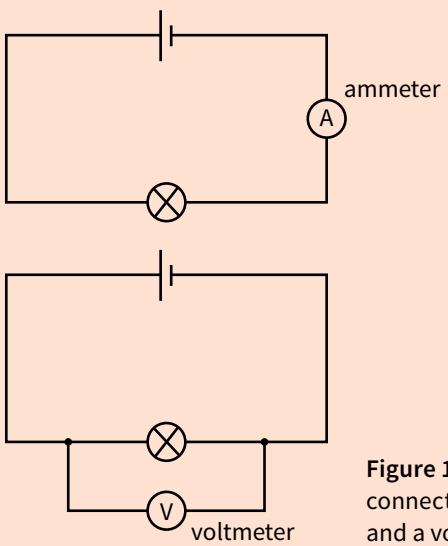


Figure 10.22 How to connect up an ammeter and a voltmeter.

Voltmeters measure the potential difference between two points in the circuit. For this reason, they are connected in parallel (i.e. between the two points), and they should have a very high resistance to take as little current as possible. The ideal resistance of a voltmeter would be infinite. In practice, voltmeters have typical resistance of about 1 MΩ. A voltmeter with a resistance of 10 MΩ measuring a p.d. of 2.5 V will take a current of 2.5×10^{-7} A and dissipate just 0.625 μJ of heat energy from the circuit every second.

Figure 10.23 shows some measuring instruments.

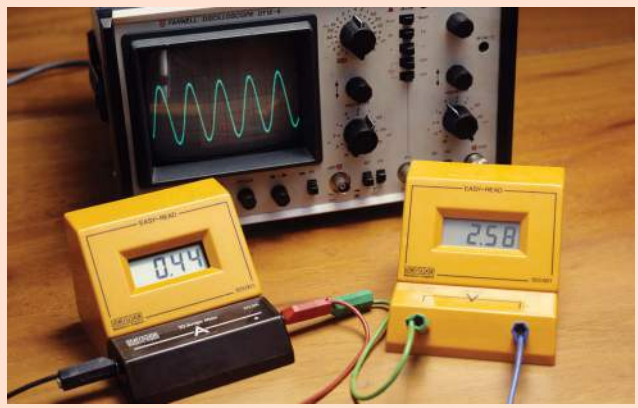


Figure 10.23 Electrical measuring instruments: an ammeter, a voltmeter and an oscilloscope. The oscilloscope can display rapidly changing voltages.

QUESTION

23 **a** A 10 V power supply of negligible internal resistance is connected to a 100 Ω resistor. Calculate the current in the resistor.

b An ammeter is now connected in the circuit, to measure the current. The resistance of the ammeter is 5.0 Ω. Calculate the ammeter reading.

Summary

- Kirchhoff's first law states that the sum of the current currents entering any point in a circuit is equal to the sum of the currents leaving that point.
- Kirchhoff's second law states that the sum of the e.m.f.s around any loop in a circuit is equal to the sum of the p.d.s around the loop.
- The combined resistance of resistors in series is given by the formula:

  $R = R_1 + R_2 + \cdots$
- The combined resistance of resistors in parallel is given by the formula:

  $\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \cdots$
- Ammeters have a low resistance and are connected in series in a circuit.
- Voltmeters have a high resistance and are connected in parallel in a circuit.

End-of-chapter questions

1 Use Kirchhoff's first law to calculate the unknown currents in the examples in Figure 10.24. In each example state the direction of the current. [4]

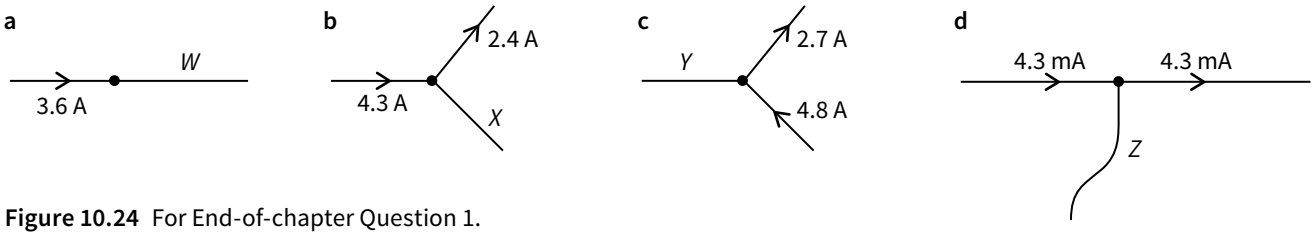


Figure 10.24 For End-of-chapter Question 1.

2 Figure 10.25 shows a part of a circuit.

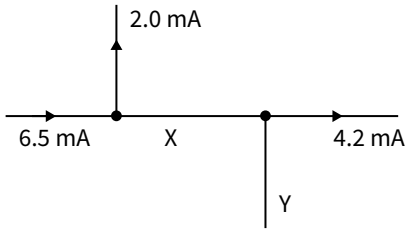


Figure 10.25 For End-of-chapter Question 2.

Copy the circuit and write in the currents at X and at Y, and show their directions. [2]

3 Figure 10.26 shows four circuits. Find the unknown potential difference (or differences) in each case. [5]

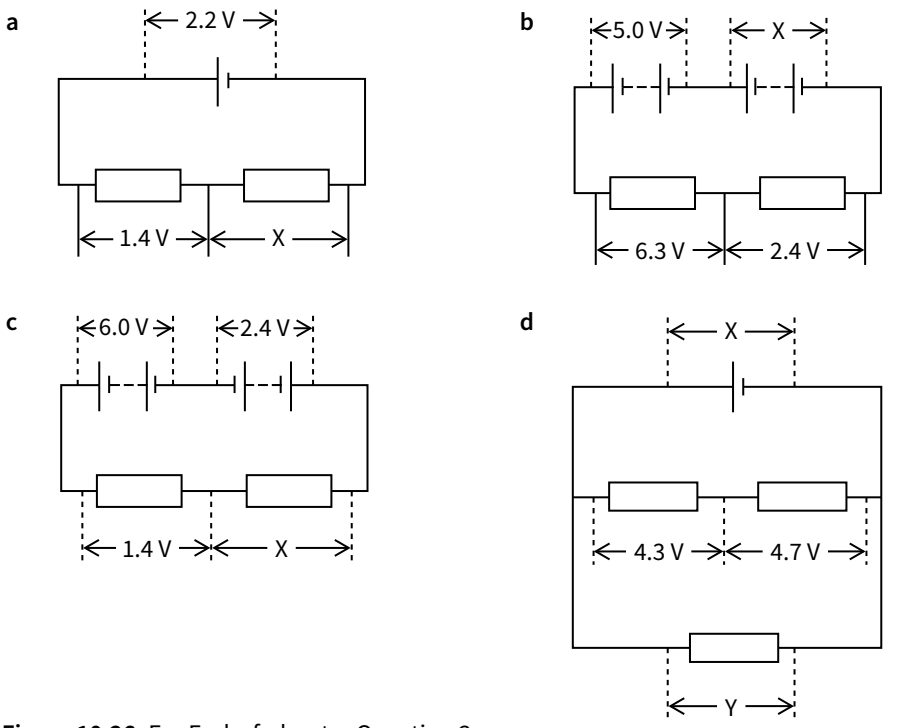


Figure 10.26 For End-of-chapter Question 3.

4 A filament lamp and a 220 Ω resistor are connected in series to a battery of e.m.f. 6.0 V. The battery has negligible internal resistance. A high-resistance voltmeter placed across the resistor measures 1.8 V.

Calculate:

a the current drawn from the battery [1]

b the p.d. across the lamp [1]

c the total resistance of the circuit [1]

d the number of electrons passing through the battery in a time of 1.0 minute. [4]

(The elementary charge is 1.6×10^{-19} C.)

5 The circuit diagram in Figure 10.27 shows a 12 V power supply connected to some resistors.

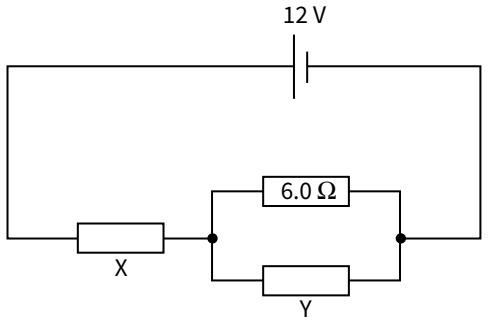


Figure 10.27 For End-of-chapter Question 5.

The current in the resistor X is 2.0 A and the current in the 6.0 Ω resistor is 0.5 A. Calculate:

a the current in resistor Y [1]

b the resistance of resistor Y [2]

c the resistance of resistor X. [2]

6 **a** Explain the difference between the terms **e.m.f.** and **potential difference**. [2]

b Figure 10.28 shows a circuit containing batteries and resistors. You may assume that the batteries have negligible internal resistance.

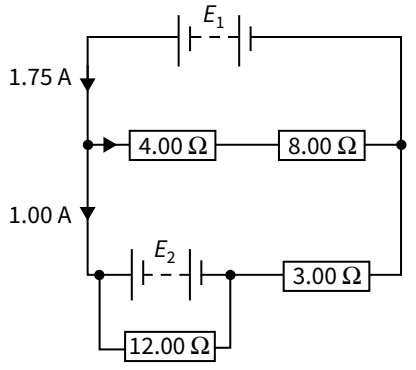


Figure 10.28 For End-of-chapter Question 6b.

i Use Kirchhoff's first law to find the current in the 4.00 Ω and 8.00 Ω resistors. [1]

ii Calculate the e.m.f. of E_1. [2]

iii Calculate the value of E_2. [2]

iv Calculate the current in the 12.00 Ω resistor. [2]

7 **a** Explain why an ammeter is designed to have a low resistance. [2]

A student builds the circuit in Figure 10.29, using a battery of negligible internal resistance.
The reading on the voltmeter is 9.0 V.

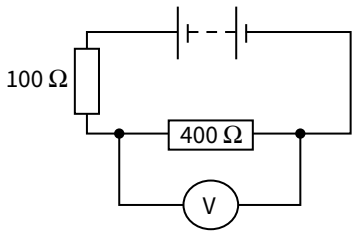


Figure 10.29 For End-of-chapter Question 7a.

b **i** The voltmeter has a resistance of 1200 Ω. Calculate the e.m.f. of the battery. [4]

ii The student now repeats the experiment using a voltmeter of resistance 12 kΩ. Show that the reading on this voltmeter would be 9.5 V. [3]

iii Refer to your answers to i and ii and explain why a voltmeter should have as high a resistance as possible. [2]

8 **a** Explain what is meant by the **resistance** of a resistor. [1]

b Figure 10.30 shows a network of resistors connected to a cell of e.m.f. 6.0 V.

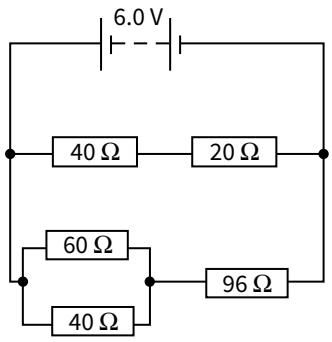


Figure 10.30 For End-of-chapter Question 8b.

Show that the resistance of the network of resistors is 40 Ω. [3]

c Calculate the current in the 60 Ω resistor. [3]

Chapter 11: Resistance and resistivity

Learning outcomes

You should be able to:

- state Ohm's law
- sketch and explain the *I*–*V* characteristics for various components
- sketch the temperature characteristic for an NTC thermistor
- solve problems involving the resistivity of a material

Superconductivity

As metals are cooled, their resistance decreases. It was discovered as long ago as 1911 that when mercury was cooled using liquid helium to 4.1 K (4.1 degrees above absolute zero), its resistance suddenly fell to zero. This phenomenon was named **superconductivity**. Other metals, such as lead at 7.2 K, also become superconductors.

When charge flows in a superconductor it can continue in that superconductor without the need for any potential difference and without dissipating any energy. This means that large currents can occur without the unwanted heating effect that would occur in a normal metallic or semiconducting conductor.

Initially superconductivity was only of scientific interest and had little practical use, as the liquid helium that was required to cool the superconductors is very expensive to produce. In 1986 it was discovered that particular ceramics became superconducting at much higher temperatures – above 77 K, the boiling point of liquid nitrogen. This meant that liquid nitrogen, which is readily available, could be used to cool the superconductors and expensive liquid helium was no longer needed. Consequently superconductor technology became a feasible proposition.

Uses of superconductors

The JR-Maglev train in Japan's Yamanashi province floats above the track using superconducting magnets (Figure 11.1). This means that not only is the heating effect of the current in the magnet coils reduced to zero – it also means that the friction between the train and the track is eliminated and that the train can reach incredibly high speeds of up to 581 km h^{-1}.



Figure 11.1 The Japanese JR-Maglev train, capable of speeds approaching 600 km h^{-1}.

Particle accelerators, such as the Large Hadron Collider (LHC) at the CERN research facility in Switzerland, accelerate beams of charged particles to very high energies by making them orbit around a circular track many times. The particles are kept moving in the circular path by very strong magnetic fields produced by electromagnets whose coils are made from superconductors. Much of our understanding of the fundamental nature of matter is from doing experiments in which beams of these very high speed particles are made to collide with each other.

Magnetic resonance imaging (MRI) was developed in the 1940s. It is used by doctors to examine internal organs without invasive surgery. Superconducting magnets can be made much smaller than conventional magnets, and this has enabled the magnetic fields produced to be much more precise, resulting in better imaging. You will find out more about MRI in Chapter 27.

The *I–V* characteristic for a metallic conductor

In Chapter 9 we saw how we could measure the resistance of a resistor using a voltmeter and ammeter. In this section we are going to investigate the variation of the current, and hence resistance, as the potential difference across a conductor changes.

The potential difference across a metal conductor can be altered using a variable power supply or by placing a variable resistor in series with the conductor. This allows us to measure the current at different potential differences across the conductor. The results of such a series of measurements shown graphically in Figure 11.2.

Look at the graph of Figure 11.2. Such a graph is known as an ***I–V* characteristic**. The points are slightly scattered, but they clearly lie on a straight line. A line of best fit has been drawn. You will see that it passes through the origin of the graph. In other words, the current I is directly proportional to the voltage V.

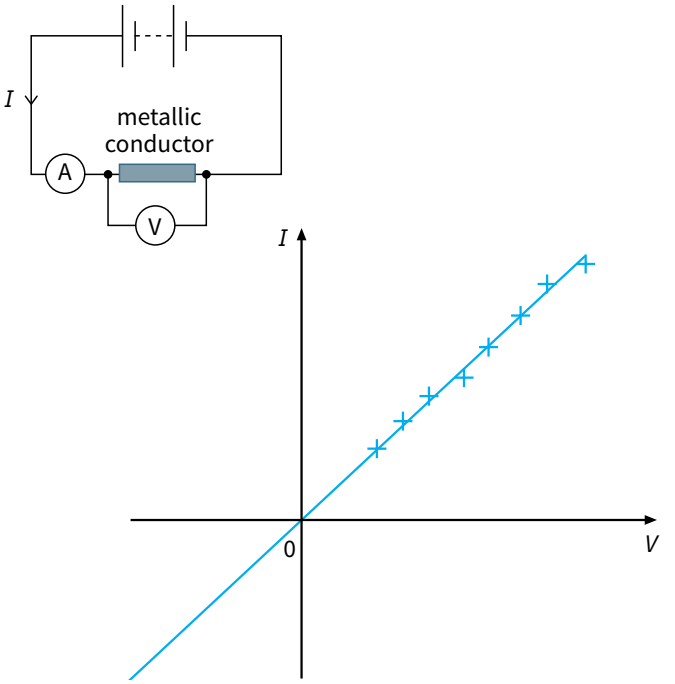


Figure 11.2 To determine the resistance of a component, you need to measure both current and potential difference.

The straight-line graph passing through the origin shows that the resistance of the conductor remains constant. If you double the current, the voltage will also double. However, its resistance, which is the ratio of the voltage to the current, remains the same. Instead of using:

$$R = \frac{V}{I}$$

to determine the resistance, for a graph of I against V which is a straight line passing through the origin you can also use:

$$\text{resistance} = \frac{1}{\text{gradient of graph}}$$

(This will give a more accurate value for R than if you were to take a single experimental data point. Take care! You can only find resistance from the gradient if the I–V graph is a straight line through the origin.)

By reversing the connections to the resistor, the p.d. across it will be reversed, i.e. negative. The current will flow in the opposite direction – it is also negative. The graph is symmetrical, showing that if a p.d. of, say, 2.0 V produces a current of 0.5 A, then a p.d. of –2.0 V will produce a current of –0.5 A. This is true for most simple metallic conductors but is not true for some electronic components, such as diodes.

You get results similar to those shown in Figure 11.2 for a commercial **resistor**. Resistors have different resistances, hence the gradient of the I–V graph will be different for different resistors.

QUESTION

1 Table 11.1 shows the results of an experiment to measure the resistance of a carbon resistor whose resistance is given by the manufacturer as 47 Ω ± 10%.

a Plot a graph to show the I–V characteristic of this resistor.

b Do the points appear to fall on a straight line which passes through the origin of the graph?

c Use the graph to determine the resistance of the resistor.

d Does the value of the resistance fall within the range given by the manufacturer?

Potential difference / V	Current / A
2.1	0.040
4.0	0.079
6.3	0.128
7.9	0.192
10.0	0.202
12.1	0.250

Table 11.1 Potential difference V and current I data for Question 1.

Ohm's law

For the metallic conductor whose I–V characteristic is shown in Figure 11.2, the current in it is directly proportional to the p.d. across it. This means that its resistance is independent of both the current and the p.d. This is because the ratio $\frac{V}{I}$ is a constant. Any component which behaves like this is described as an **ohmic component**, and we say that it obeys **Ohm's law**. The statement of Ohm's law is very precise and you must not confuse this with the equation '$\frac{V}{I} = R$'.

Ohm's law

A conductor obeys Ohm's law if the current in it is directly proportional to the potential difference across its ends.

QUESTION

2 An electrical component allows a current of 10 mA through it when a voltage of 2.0 V is applied. When the voltage is increased to 8.0 V, the current becomes 60 mA. Does the component obey Ohm's law? Give numerical values for the resistance to justify your answer.

Resistance and temperature

A conductor that does not obey Ohm's law is described as **non-ohmic**. An example is a filament lamp. Figure 11.3 shows such a lamp; you can clearly see the wire filament glowing as the current passes through it. Figure 11.4 shows the *I*–*V* characteristic for a similar lamp.

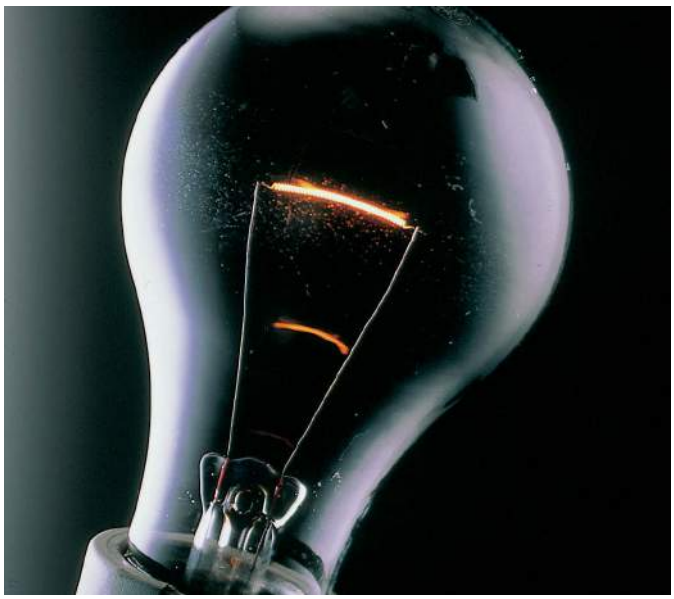


Figure 11.3 The metal filament in a lamp glows as the current passes through it. It also feels warm. This shows that the lamp produces both heat and light.

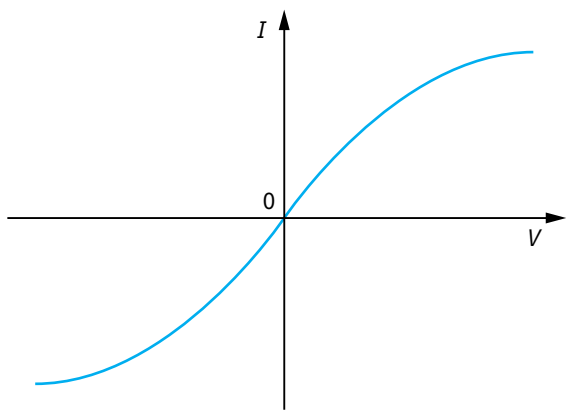


Figure 11.4 The *I*–*V* characteristic for a filament lamp.

There are some points you should notice about the graph in Figure 11.4:

- The line passes through the origin (as for an ohmic component).
- For very small currents and voltages, the graph is roughly a straight line.
- At higher voltages, the line starts to curve. The current is a bit less than we would have expected from a straight line. This suggests that the lamp's resistance has increased. You can also tell that the resistance has increased because the ratio $\frac{V}{I}$ is larger for higher voltages than for low voltages.

The fact that the graph of Figure 11.4 is not a straight line shows that the resistance of the lamp depends on the temperature of its filament. Its resistance may increase by a factor as large as ten between when it is cold and when it is brightest (when its temperature may be as high as 1750 °C).

Thermistors

Thermistors are components that are designed to have a resistance which changes rapidly with temperature. Thermistors ('**therm**al res**istors**') are made from metal oxides such as those of manganese and nickel. There are two distinct types of thermistor:

- Negative temperature coefficient (NTC) thermistors – the resistance of this type of thermistor decreases with increasing temperature. Those commonly used for physics teaching may have a resistance of many thousands of ohms at room temperature, falling to a few tens of ohms at 100 °C. You should become familiar with the properties of NTC thermistors.
- Positive temperature coefficient (PTC) thermistors – the resistance of this type of thermistor rises abruptly at a definite temperature, usually around 100–150 °C.

The change in their resistance with temperature gives thermistors many uses:

- Water temperature sensors in cars and ice sensors on aircraft wings – if ice builds up on the wings, the thermistor 'senses' this temperature drop and a small heater is activated to melt the ice.
- Baby alarms – the baby rests on an air-filled pad, and as he or she breathes, air from the pad passes over a thermistor, keeping it cool; if the baby stops breathing, the air movement stops, the thermistor warms up and an alarm sounds.
- Fire sensors – a rise in temperature activates an alarm.
- Overload protection in electric razor sockets – if the razor overheats, the thermistor's resistance rises rapidly and cuts off the circuit.

QUESTIONS

3 The two graphs in Figure 11.5 show the *I*–*V* characteristics of a metal wire at two different temperatures, θ_1 and θ_2.

a Calculate the resistance of the wire at each temperature.

b State which is the higher temperature, θ_1 or θ_2.

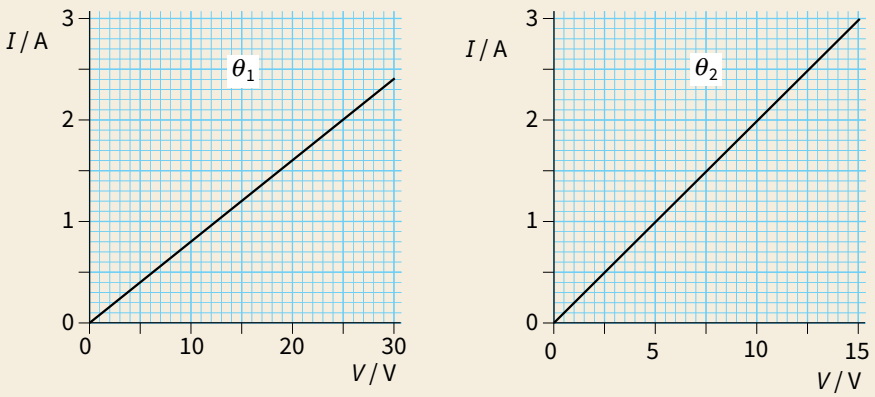


Figure 11.5 *I*–*V* graphs for a wire at two different temperatures. For Question 3.

4 The graph of Figure 11.6 shows the *I*–*V* characteristics of two electrical components, a filament lamp and a length of steel wire.

a Identify which curve relates to each component.

b State the voltage at which both have the same resistance.

c Determine the resistance at the voltage stated in **b**.

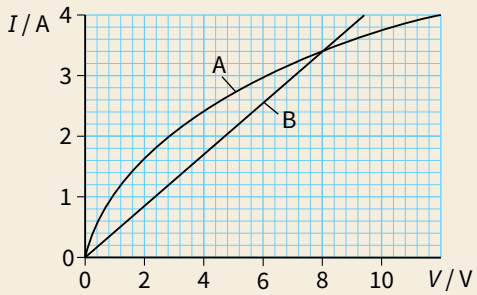


Figure 11.6 For Question 4.

Diodes

The semiconductor diode is another example of a non-ohmic conductor. A diode is any component that allows electric current in only one direction. Nowadays, most diodes are made of semiconductor materials. One type, the light-emitting diode or LED, gives out light when it conducts.

Figure 11.7 shows the *I*–*V* characteristic for a diode. There are some points you should notice about this graph.

- We have included positive and negative values of current and voltage. This is because, when connected one way round (positively biased), the diode conducts and has a fairly low resistance. Connected the other way round (negatively biased), it allows only a tiny current and has almost infinite resistance.
- For positive voltages less than about 0.6 V, the current is almost zero and hence the diode has almost infinite resistance. It starts to conduct suddenly at its **threshold voltage**. The resistance of the diode decreases dramatically for voltages greater than 0.6 V.

The resistance of a diode depends on the potential difference across it. From this we can conclude that it does not obey Ohm's law; it is a non-ohmic component.

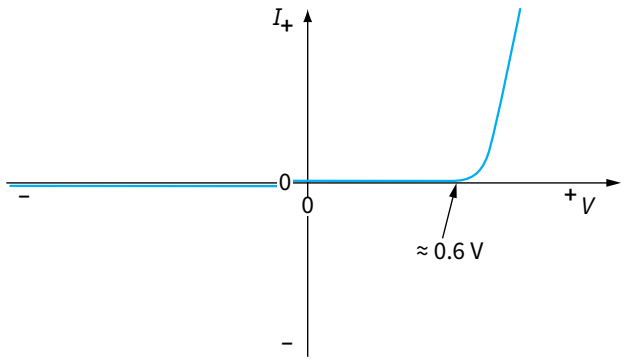


Figure 11.7 The current against potential difference (*I*–*V*) characteristic for a diode. The graph is not a straight line. A diode does not obey Ohm's law.

Diodes are used as rectifiers. They allow current to pass in one direction only and so can be used to convert alternating current into direct current. (There is more about this in Chapter 29.) Most modern diodes are made from silicon and will start conducting when there is a potential difference of about 0.6 V across them. You need to remember this key 0.6 V value.

LEDs have traditionally been used as indicator lamps to show when an appliance is switched on. Newer versions, some of which produce white light, are replacing filament lamps, for example in traffic lights and torches (flashlights) – see Figure 11.8. Although they are more expensive to manufacture, they are more energy-efficient and hence cheaper to run, so that the overall cost is less.

The threshold voltage at which an LED starts to conduct and emit light is higher than 0.6 V and depends on the colour of light it emits, but may be taken to be about 2 V.



Figure 11.8 This torch has seven white LEDs, giving a brighter, whiter light than a traditional filament lamp.

Understanding the origin of resistance

To understand a little more about the origins of resistance, it is helpful to look at how the resistance of a pure metal wire changes as its temperature is increased. This is shown in the graph of Figure 11.9. You will see that the resistance of the pure metal increases linearly as the temperature increases from 0 °C to 100 °C. Compare this with the graph of Figure 11.10 for an NTC thermistor; the thermistor's resistance decreases very dramatically over a narrow temperature range.

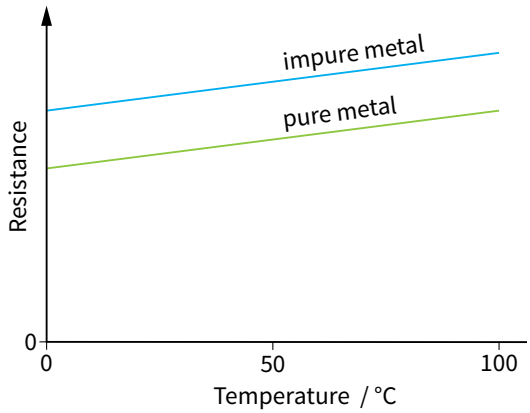


Figure 11.9 The resistance of a metal increases gradually as its temperature is increased. The resistance of an impure metal wire is greater than that of a pure metal wire of the same dimensions.

QUESTIONS

5 The graph in Figure 11.10 was obtained by measuring the resistance R of a particular thermistor as its temperature θ changed.

a Determine its resistance at:

i 20 °C

ii 45 °C.

b Determine the temperature when its resistance is:

i 5000 Ω

ii 2000 Ω.

6 A student connects a circuit with an NTC thermistor, a filament lamp and a battery in series. The lamp glows dimly. The student warms the thermistor with a hair dryer. What change will the student notice in the brightness of the lamp? Explain your answer.

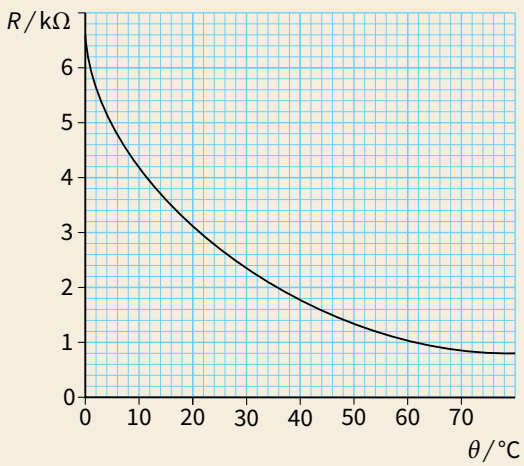


Figure 11.10 The resistance of an NTC thermistor decreases as the temperature increases. For Question 5.

Figure 11.10 also shows how the resistance of the metal changes if it is slightly impure. The resistance of an impure metal is greater than that of the pure metal and follows the same gradual upward slope. The resistance of a metal changes in this gradual way over a wide range of temperatures – from close to absolute zero up to its melting point, which may be over 2000 °C.

This suggests that there are two factors which affect the resistance of a metal:

- the temperature
- the presence of impurities.

Figure 11.11 shows a simple model which explains what happens in a metal when electrons flow through it.

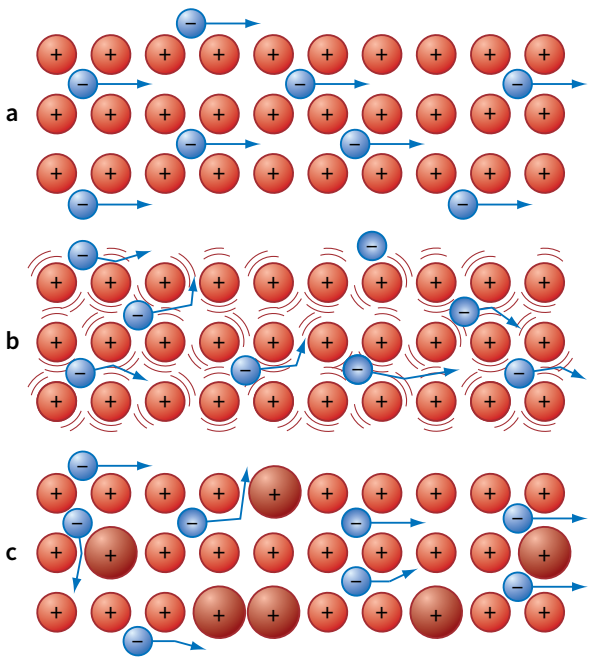


Figure 11.11 A model of the origins of resistance in a metal. **a** At low temperatures, electrons flow relatively freely. **b** At higher temperatures, the electrons are obstructed by the vibrating ions and they make very frequent collisions with the ions. **c** Impurity atoms can also obstruct the free flow of electrons.

In a metal, a current is due to the movement of free electrons. At low temperatures, they can move easily past the positive ions (Figure 11.11a). However, as the temperature is raised, the ions vibrate with larger amplitudes. The electrons collide more frequently with the vibrating ions, and this decreases their mean drift velocity. They lose energy to the vibrating ions (Figure 11.11b).

If the metal contains impurities, some of the atoms will be of different sizes (Figure 11.11c). Again, this disrupts the free flow of electrons. In colliding with impurity atoms, the electrons lose energy to the vibrating atoms.

You can see that electrons tend to lose energy when they collide with vibrating ions or impurity atoms. They give up energy to the metal, so it gets hotter. The resistance of the metal increases with the temperature of the wire because of the decrease in the mean drift velocity of the electrons.

Conduction in semiconductors is different. At low temperatures, there are few delocalised, or free, electrons. For conduction to occur, electrons must have sufficient energy to free themselves from the atom they are bound to. As the temperature increases, a few electrons gain enough energy to break free of their atoms to become conduction electrons. The number of conduction electrons thus increases and so the material becomes a better conductor. At the same time, there are more electron–ion collisions, but this effect is small compared with the increase in the number of conduction electrons.

QUESTION

7 The resistance of a metal wire changes with temperature. This means that a wire could be used to sense changes in temperature, in the same way that a thermistor is used.

a Suggest **one** advantage a thermistor has over a metal wire for this purpose.

b Suggest **one** advantage a metal wire has over a thermistor.

Resistivity

The resistance of a particular wire depends on its size and shape. A long wire has a greater resistance than a short one, provided it is of the same thickness and material. A thick wire has less resistance than a thin one. For a metal in the shape of a wire, R depends on the following factors:

- length L
- cross-sectional area A
- the material the wire is made from
- the temperature of the wire.

At a constant temperature, the resistance is directly proportional to the length of the wire and inversely proportional to its cross-sectional area. That is:

$$\text{resistance} \propto \text{length}$$

and

$$\text{resistance} \propto \frac{1}{\text{cross-sectional area}}$$

We can see how these relate to the formulae for adding resistors in series and in parallel:

- If we double the length of a wire it is like connecting two identical resistors in series; their resistances add to give double the resistance. The resistance is proportional to the length.
- Doubling the cross-sectional area of a wire is like connecting two identical resistors in parallel; their combined resistance is halved (since $\frac{1}{R_{total}} = \frac{1}{R} + \frac{1}{R}$). Hence the resistance is inversely proportional to the cross-sectional area.

Combining the two proportionalities for length and cross-sectional area, we get:

$$\text{resistance} \propto \frac{\text{length}}{\text{cross-sectional area}}$$

or

$$R \propto \frac{L}{A}$$

But the resistance of a wire also depends on the material it is made of. Copper is a better conductor than steel, steel is a better conductor than silicon, and so on. So if we are to determine the resistance R of a particular wire, we need to take into account its length, its cross-sectional area and the material. The relevant property of the material is its **resistivity**, for which the symbol is ρ (Greek letter rho).

The word equation for resistance is:

$$\text{resistance} = \frac{\text{resistivity} \times \text{length}}{\text{cross-sectional area}}$$

$$R = \frac{\rho L}{A}$$

We can rearrange this equation to give an equation for resistivity. The resistivity of a material is defined by the following word equation:

$$\text{resistivity} = \frac{\text{resistance} \times \text{cross-sectional area}}{\text{length}}$$

$$\rho = \frac{RA}{L}$$

Values of the resistivities of some typical materials are shown in Table 11.2. Notice that the units of resistivity are ohm metres (Ω m); this is not the same as ohms per metre.

Material	Resistivity / Ω m	Material	Resistivity / Ω m
silver	1.60×10^{-8}	mercury	69.0×10^{-8}
copper	1.69×10^{-8}	graphite	800×10^{-8}
nichrome[a]	1.30×10^{-8}	germanium	0.65
aluminium	3.21×10^{-8}	silicon	2.3×10^{3}
lead	20.8×10^{-8}	Pyrex glass	10^{12}
manganin[b]	44.0×10^{-8}	PTFE[d]	10^{13}–10^{16}
eureka[c]	49.0×10^{-8}	quartz	5×10^{16}

(a) Nichrome – an alloy of nickel, copper and aluminium used in electric heaters because it does not oxidise at 1000 °C.
(b) Manganin – an alloy of 84% copper, 12% manganese and 4% nickel.
(c) Eureka (constantan) – an alloy of 60% copper and 40% nickel.
(d) Poly(tetrafluoroethene) or Teflon.

Table 11.2 Resistivities of various materials at 20 °C.

WORKED EXAMPLE

1 Find the resistance of a 2.6 m length of eureka wire with cross-sectional area 2.5×10^{-7} m^2.

Step 1 Use the equation for resistance:

$$\text{resistance} = \frac{\text{resistivity} \times \text{length}}{\text{area}}$$

$$R = \frac{\rho L}{A}$$

Step 2 Substitute values from the question and use the value for ρ from Table 11.2:

$$R = \frac{49.0 \times 10^{-8} \times 2.6}{2.5 \times 10^{-7}} = 5.1\,\Omega$$

So the wire has a resistance of 5.1 Ω.

Resistivity and temperature

Resistivity, like resistance, depends on temperature. For a metal, resistivity increases with temperature. As we saw above, this is because there are more frequent collisions between the conduction electrons and the vibrating ions of the metal.

QUESTIONS

8 Use the resistivity value quoted in Table 11.2 to calculate the lengths of 0.50 mm diameter manganin wire needed to make resistance coils with resistances of:

a 1.0 Ω

b 5.0 Ω

c 10 Ω.

9 1.0 cm^3 of copper is drawn out into the form of a long wire of cross-sectional area 4.0×10^{-7} m^2. Calculate its resistance. (Use the resistivity value for copper from Table 11.2.)

10 A 1.0 m length of copper wire has a resistance of 0.50 Ω.

a Calculate the resistance of a 5.0 m length of the same wire.

b What will be the resistance of a 1.0 m length of copper wire having half the diameter of the original wire?

11 A piece of steel wire has a resistance of 10 Ω. It is stretched to twice its original length. Compare its new resistance with its original resistance.

Summary

- A conductor obeys Ohm's law if the current in it is directly proportional to the potential difference across its ends.
- Ohmic components include a wire at constant temperature and a resistor.
- Non-ohmic components include a filament lamp and a light-emitting diode.
- A semiconductor diode allows current in one direction only.
- As the temperature of a metal increases, so does its resistance.
- A thermistor is a component which shows a rapid change in resistance over a narrow temperature range. The resistance of an NTC thermistor decreases as its temperature is increased.
- The resistivity ρ of a material is defined as $\rho = \frac{RA}{L}$, where R is the resistance of a wire of that material, A is its cross-sectional area and L is its length. The unit of resistivity is the ohm metre (Ω m).

End-of-chapter questions

1 The graph in Figure 11.12 shows the I–V characteristic of an electrical component.

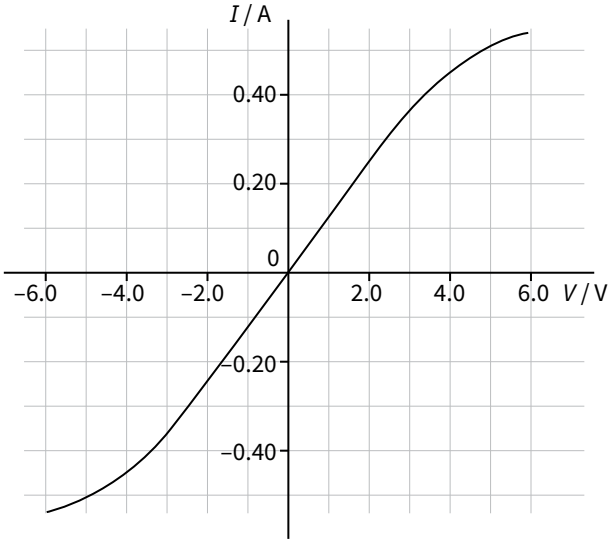


Figure 11.12 For End-of-chapter Question 1.

a Calculate the resistance of the component when the potential difference across it is:

i 2.0 V [2]

ii 5.0 V. [1]

b Suggest what the component is. [1]

2 A student connects an NTC thermistor to a battery and an ammeter. He places the thermistor in a beaker of water and gradually heats the water from 10 °C to its boiling point, recording the value of the current as he does so. He then plots a graph of the current in the thermistor against the temperature of the water.

a Sketch the graph you would expect the student to obtain from the experiment. [1]

b Explain how the student could now use the thermistor as a thermometer. [2]

3 **a** Describe the difference between the conduction processes in copper and in silicon, a semiconductor. [2]

b Explain why the resistance of a metallic conductor increases with temperature while that of a semiconductor decreases. [4]

4 A nichrome wire has a length of 1.5 m and a cross-sectional area of 0.080 mm^2.The resistivity of nichrome is $1.30 \times 10^{-8}\,\Omega\,m$.

a Calculate the resistance of the wire. [2]

b Calculate the length of this wire which would be needed to make an element of an electric heater of resistance 30 Ω. [2]

5 Figure 11.13 shows a circuit.

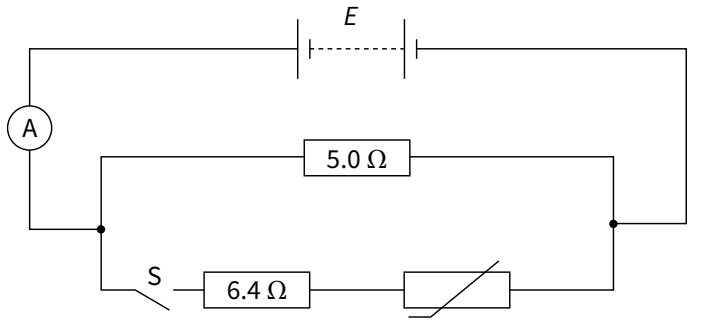


Figure 11.13 For End-of-chapter Question 5.

a When switch S is open the current in ammeter A is 0.48 A. Calculate the e.m.f. of the battery. You may assume the battery has negligible internal resistance. [2]

b When switch S is closed the current in the ammeter increases to 0.72 A.

i Deduce the current in the 6.4 Ω resistor. [1]

ii State the current in the thermistor. [1]

c State and explain how the reading on the ammeter changes when the temperature of the thermistor is increased. [3]

6 **a** Explain why the resistance of a metal increases when its temperature increases. [2]

b State **two** other factors which determine the resistance of a stated length of wire. [2]

c When a potential difference of 1.5 V is applied across a 5.0 m length of insulated copper wire, a current of 0.24 A is measured in it.

i Calculate the resistance of the length of wire. [2]

ii The resistivity of copper is $1.69 \times 10^{-8}\,\Omega\,m$. Calculate the diameter of the wire. [3]

d The wire is now made into a tight bundle. State and explain how you would expect the current in it to change. [3]

7 Figure 11.14 shows a piece of silicon of width 32 mm and length 36 mm. The resistance of the silicon between the points P and Q is 1.1 MΩ. Silicon has a resistivity of $2.3 \times 10^{3}\,\Omega\,m$.

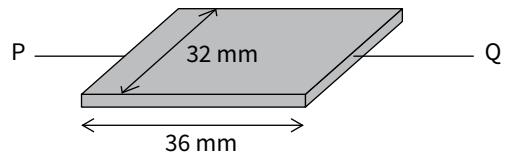


Figure 11.14 For End-of-chapter Question 7.

a Calculate the thickness of the piece of silicon. [3]

b Calculate the current which would pass through the silicon if a potential difference of 12 V were applied across P and Q. [2]

c Discuss how the current would change if it were large enough to cause the silicon to become significantly warmer. [3]

8 A student is investigating the properties of a semiconducting diode. Figure 11.15 shows the circuit she builds.

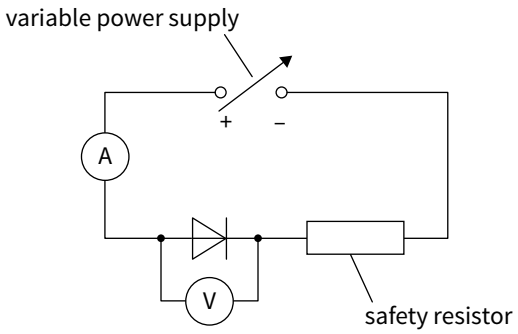


Figure 11.15 For End-of-chapter Question 8.

a **i** Sketch a graph to show how the current in the diode would vary as the voltage across it is increased from 0 V to 1.0 V. [1]

ii The supply is now connected in the reverse direction and once more the potential difference across the diode is increased from 0 V to 1.0 V. Complete the *I*–*V* graph. [1]

b Suggest why the safety resistor is required. [2]

c When the potential difference across the safety resistor is 1.4 V, the current in it is 20 mA. Calculate the resistance of the safety resistor. [2]

9 **a** Explain what is meant by an **ohmic conductor**. [2]

b **i** Sketch a graph of resistance *R* against voltage *V* for a wire of pure iron kept at constant temperature. Label this line X. [1]

ii Sketch a graph of resistance *R* against voltage *V* for a second wire of impure iron, of the same diameter and the same length, which is kept at the same temperature. Label this line Y. [1]

iii Explain how the graphs would change if the wires were kept at a higher, but still constant, temperature. [2]

c Deduce how the resistance of a wire made of pure iron would change if both the diameter and the length were doubled. [3]

Chapter 12: Practical circuits

Learning outcomes

You should be able to:

- explain the effects of internal resistance on terminal p.d. and power output of a source of e.m.f.
- explain the use of potential divider circuits
- solve problems involving the potentiometer as a means of comparing voltages

The first electrical cell – an historical mystery

The Italian Alessandro Volta (Figure 12.1a) is generally credited with inventing the first battery. He devised it after his friend and rival Luigi Galvani had shown that a (dead) frog's leg could be made to twitch if an electrically charged plate was connected to it. Volta's battery consisted of alternate discs of copper and zinc, separated by felt soaked in brine – see Figure 12.1b.

However, there is evidence that earlier technologists may have beaten him by over 1000 years. In 1936 a small pot was discovered during an archaeological dig near Baghdad. The pot was sealed with pitch, and inside the pot there was a copper cylinder surrounding an iron rod. When filled with an acid, perhaps vinegar, a potential difference of around 1.5 volts could be produced between the copper and the iron.

It has been suggested that this battery might have been used to electroplate metal objects with gold. So did Volta really invent the battery, or did he just rekindle an art that had been lost for more than a millennium?

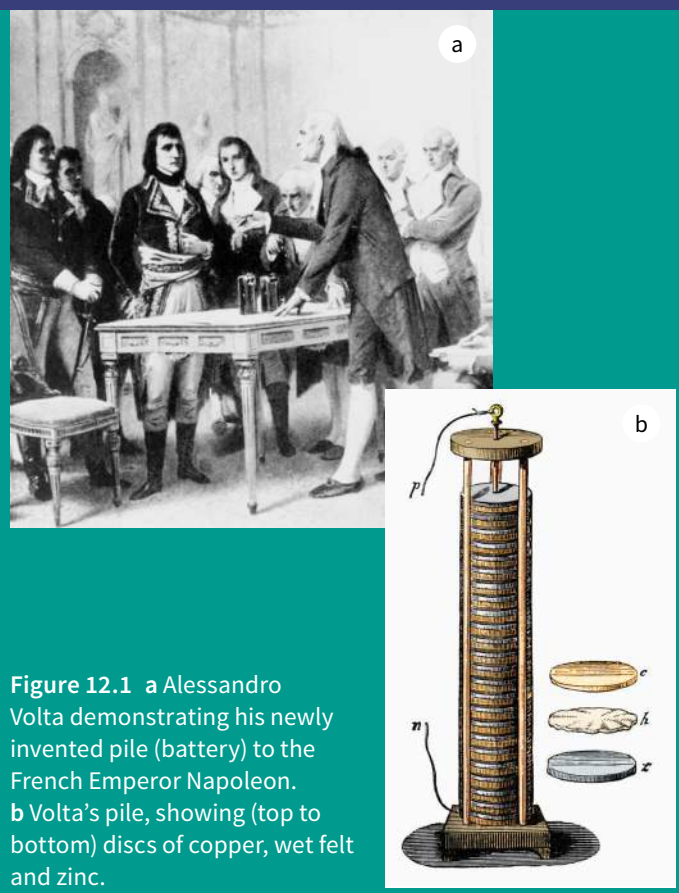


Figure 12.1 **a** Alessandro Volta demonstrating his newly invented pile (battery) to the French Emperor Napoleon. **b** Volta's pile, showing (top to bottom) discs of copper, wet felt and zinc.

Internal resistance

You will be familiar with the idea that, when you use a power supply or other source of e.m.f., you cannot assume that it is providing you with the exact voltage across its terminals as suggested by the value of its e.m.f. There are several reasons for this. For example, the supply may not be made to a high degree of precision, batteries become flat, and so on. However, there is a more important factor, which is that all sources of e.m.f. have an **internal resistance**. For a power supply, this may be due to the wires and components inside, whereas for a cell the internal resistance is due to the chemicals within it. Experiments show that the voltage across the terminals of the power supply depends on the circuit of which it is part. In particular, the voltage across the power supply terminals decreases if it is required to supply more current.

Figure 12.2 shows a circuit you can use to investigate this effect, and a sketch graph showing how the voltage across the terminals of a power supply might decrease as the supplied current increases.

The charges moving round a circuit have to pass through the external components **and** through the internal resistance of the power supply. These charges gain electrical energy from the power supply. This energy is lost as heat as the charges pass through the external components and through the internal resistance of the power supply. Power supplies and batteries get warm when they are being used. (Try using a cell to light a small torch bulb; feel the cell before connecting to the bulb, and then feel it again after the bulb has been lit for about 15 seconds.)

The reason for this heating effect is that some of the electrical potential energy of the charges is transformed to internal energy as they do work against the internal resistance of the cell.

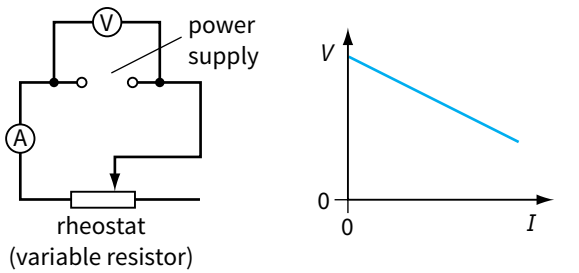


Figure 12.2 **a** A circuit for determining the e.m.f. and internal resistance of a supply; **b** typical form of results.

It can often help to solve problems if we show the internal resistance r of a source of e.m.f. explicitly in circuit diagrams (Figure 12.3). Here, we are representing a cell as if it were a 'perfect' cell of e.m.f. E, together with a separate resistor of resistance r. The dashed line enclosing E and r represents the fact that these two are, in fact, a single component.

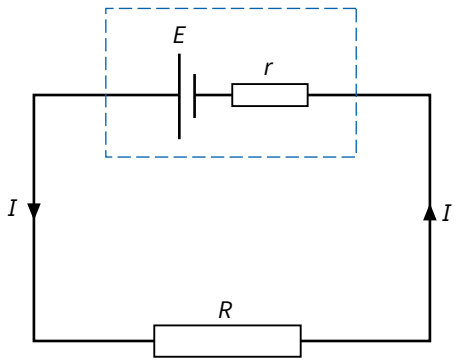


Figure 12.3 It can be helpful to show the internal resistance r of a cell (or a supply) in a circuit diagram.

Now we can determine the current when this cell is connected to an external resistor of resistance R. You can see that R and r are in series with each other. The current I is the same for both of these resistors. The combined resistance of the circuit is thus $R + r$, and we can write:

$$E = I(R + r) \quad \text{or} \quad E = IR + Ir$$

We cannot measure the e.m.f. E of the cell directly, because we can only connect a voltmeter across its terminals. This **terminal p.d.** V across the cell is always the same as the p.d. across the external resistor. Therefore, we have:

$$V = IR$$

This will be less than the e.m.f. E by an amount Ir. The quantity Ir is the potential difference across the internal resistor and is referred to as the **lost volts**. If we combine these two equations, we get:

$$V = E - Ir$$

or

terminal p.d. = e.m.f. – 'lost volts'

The 'lost volts' indicates the energy transferred to the internal resistance of the supply. If you short-circuit a battery with a piece of wire, a large current will flow, and the battery will get warm as energy is transferred within it. This is also why you may damage a power supply by trying to make it supply a larger current than it is designed to give.

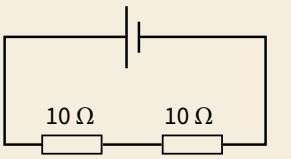
QUESTIONS

1 A battery of e.m.f. 5.0 V and internal resistance 2.0 Ω is connected to an 8.0 Ω resistor. Draw a circuit diagram and calculate the current in the circuit.

2 **a** Calculate the current in each circuit in Figure 12.4.

b Calculate also the 'lost volts' for each cell, and the terminal p.d.

i $E = 3.0$ V, $r = 4.0$ Ω



ii $E = 3.0$ V, $r = 4.0$ Ω

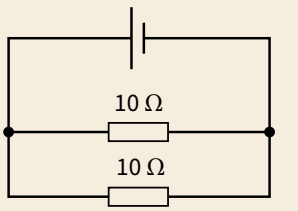


Figure 12.4 For Question 2.

3 Four identical cells, each of e.m.f. 1.5 V and internal resistance 0.10 Ω, are connected in series. A lamp of resistance 2.0 Ω is connected across the four cells. Calculate the current in the lamp.

WORKED EXAMPLE

1 There is a current of 0.40 A when a battery of e.m.f. 6.0 V is connected to a resistor of 13.5 Ω. Calculate the internal resistance of the cell.

Step 1 Substitute values from the question in the equation for e.m.f.:

$E = 6.0$ V, $I = 0.40$ A, $R = 13.5$ Ω

$E = IR + Ir$

$6.0 = 0.40 \times 13.5 + 0.40 \times r$

$= 5.4 + 0.40r$

Step 2 Rearrange the equation to make r the subject and solve:

$6.0 - 5.4 = 0.40r$

$0.60 = 0.40r$

$$r = \frac{0.60}{0.40} = 1.5\ \Omega$$

BOX 12.1: Determining e.m.f. and internal resistance

You can get a good idea of the e.m.f. of an isolated power supply or a battery by connecting a digital voltmeter across it. A digital voltmeter has a very high resistance (~$10^7\,\Omega$), so only a tiny current will pass through it. The 'lost volts' will then only be a tiny fraction of the e.m.f. If you want to determine the internal resistance r as well as the e.m.f. E, you need to use a circuit like that shown in Figure 12.2. When the variable resistor is altered, the current in the circuit changes, and measurements can be recorded of the circuit current I and terminal p.d. V. The internal resistance r can be found from a graph of V against I (Figure 12.5).

Compare the equation $V = E - Ir$ with the equation of a straight line $y = mx + c$. By plotting V on the y-axis and I on the x-axis, a straight line should result. The intercept on the y-axis is E, and the gradient is $-r$.

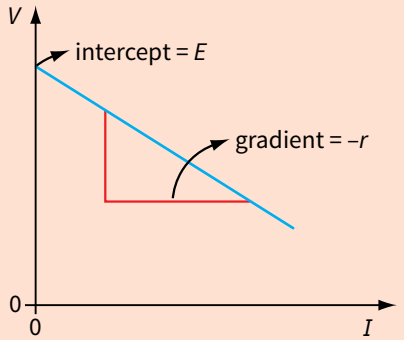


Figure 12.5 E and r can be found from this graph.

QUESTIONS

4 When a high-resistance voltmeter is placed across an isolated battery, its reading is 3.0 V. When a 10 Ω resistor is connected across the terminals of the battery, the voltmeter reading drops to 2.8 V. Use this information to determine the internal resistance of the battery.

5 The results of an experiment to determine the e.m.f. E and internal resistance r of a power supply are shown in Table 12.1. Plot a suitable graph and use it to find E and r.

V / V	1.43	1.33	1.18	1.10	0.98
I / A	0.10	0.30	0.60	0.75	1.00

Table 12.1 Results for Question 5.

The effects of internal resistance

You cannot ignore the effects of internal resistance. Consider a battery of e.m.f. 3.0 V and of internal resistance 1.0 Ω. The **maximum current** that can be drawn from this battery is when its terminals are shorted-out. (The external resistance $R \approx 0$.) The maximum current is given by:

$$\text{maximum current} = \frac{E}{r} = \frac{3.0}{1.0} = 3.0\,\text{A}$$

The **terminal p.d.** of the battery depends on the resistance of the external resistor. For an external resistor of resistance 1.0 Ω, the terminal p.d. is 1.5 V – half of the e.m.f. The terminal p.d. approaches the value of the e.m.f. when the external resistance R is very much greater than the internal resistance of the battery. For example, a resistor of resistance 1000 Ω connected to the battery gives a terminal p.d. of 2.997 V. This is almost equal to the e.m.f. of the battery. The more current a battery supplies, the more its terminal p.d. will decrease. An example of this can be seen when a driver tries to start a car with the headlamps on. The starter motor requires a large current from the battery, the battery's terminal p.d. drops, and the headlamps dim.

QUESTION

6 A car battery has an e.m.f. of 12 V and an internal resistance of 0.04 Ω. The starter motor draws a current of 100 A.

a Calculate the terminal p.d. of the battery when the starter motor is in operation.

b Each headlamp is rated as '12 V, 36 W'. Calculate the resistance of a headlamp.

c To what value will the power output of each headlamp decrease when the starter motor is in operation? (Assume that the resistance of the headlamp remains constant.)

Potential dividers

How can we get an output of 3.0 V from a battery of e.m.f. 6.0 V? Sometimes we want to use only part of the e.m.f. of a supply. To do this, we use an arrangement of resistors called a **potential divider** circuit.

Figure 12.6 shows two potential divider circuits, each connected across a battery of e.m.f. 6.0 V and of negligible

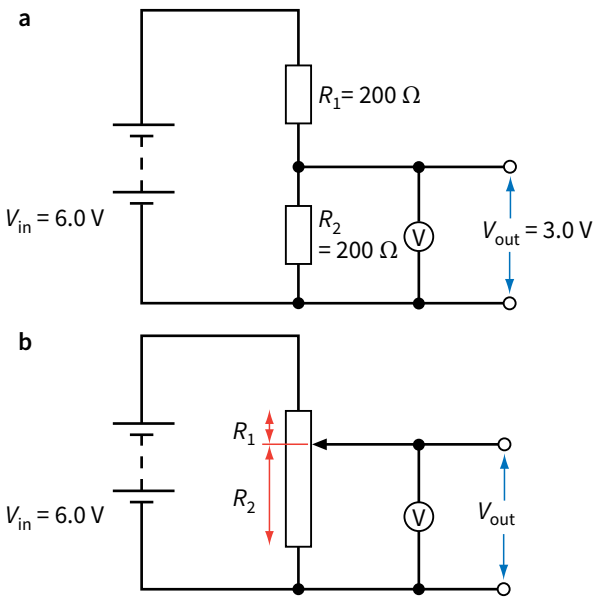


Figure 12.6 Two potential divider circuits.

internal resistance. The high-resistance voltmeter measures the voltage across the resistor of resistance R_2. We refer to this voltage as the output voltage, V_{out}, of the circuit. The first circuit, **a**, consists of two resistors of values R_1 and R_2. The voltage across the resistor of resistance R_2 is half of the 6.0 V of the battery. The second potential divider, **b**, is more useful. It consists of a single variable resistor. By moving the sliding contact, we can achieve any value of V_{out} between 0.0 V (slider at the bottom) and 6.0 V (slider at the top).

The output voltage V_{out} depends on the relative values of R_1 and R_2. You can calculate the value of V_{out} using the following **potential divider equation**:

$$V_{out} = \left(\frac{R_2}{R_1 + R_2}\right) \times V_{in}$$

In this equation, V_{in} is the total voltage across the two resistors.

QUESTION

7 Determine the range of V_{out} for the circuit in Figure 12.7 as the variable resistor R_2 is adjusted over its full range from 0 Ω to 40 Ω. (Assume the supply of e.m.f. 10 V has negligible internal resistance.)

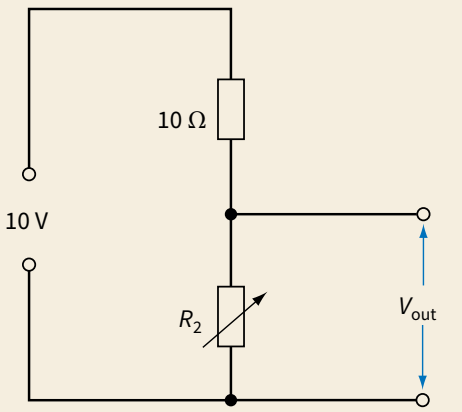


Figure 12.7 For Question 7.

Potentiometer circuits

A **potentiometer** is a device used for comparing potential differences. For example, it can be used to measure the e.m.f. of a cell, provided you already have a source whose e.m.f. is known accurately. As we will see, a potentiometer can be thought of as a type of potential divider circuit.

A potentiometer consists of a piece of resistance wire, usually 1 m in length, stretched horizontally between two points. In Figure 12.8, the ends of the wire are labelled A and B. A **driver cell** is connected across the length of wire. Suppose this cell has an e.m.f. E_o of 2.0 V. We can then say that point A is at a voltage of 2.0 V, B is at 0 V, and the midpoint of the wire is at 1.0 V. In other words, the voltage decreases steadily along the length of the wire.

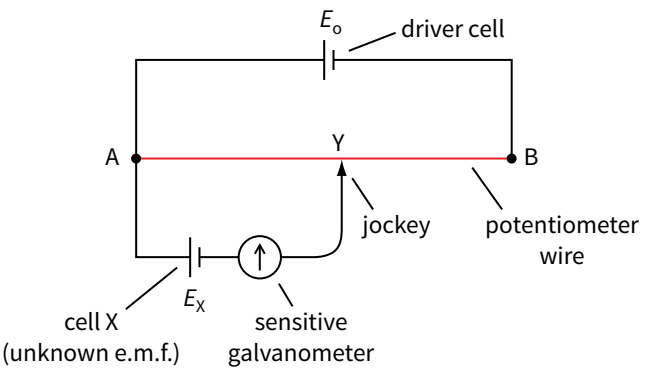


Figure 12.8 A potentiometer connected to measure the e.m.f. of cell X.

Now, suppose we wish to measure the e.m.f. E_X of cell X (this must have a value less than that of the driver cell). The positive terminal of cell X is connected to point A. (Note that both cells have their positive terminals connected to A.) A lead from the negative terminal is connected to a sensitive galvanometer (e.g. a microammeter), and a lead from the other terminal of the galvanometer ends with a metal **jockey**. This is a simple connecting device with a very sharp edge that allows very precise positioning on the wire.

If the jockey is touched onto the wire close to point A, the galvanometer needle will deflect in one direction. If the jockey is touched close to B, the galvanometer needle will deflect in the opposite direction. Clearly there must be some point Y along the wire which, when touched by the jockey, gives zero deflection – the needle moves neither to the left nor the right.

In finding this position, the jockey must be touched gently and briefly onto the wire; the deflection of the galvanometer shows whether the jockey is too far to the left or right. It is important not to slide the jockey along the potentiometer wire as this may scrape its surface, making it non-uniform so that the voltage does not vary uniformly along its length.

When the jockey is positioned at Y, the galvanometer gives zero deflection, showing that there is no current through it. This can only happen if the potential difference across the length of wire AY is equal to the e.m.f. of cell X. We can say that the potentiometer is balanced. If the balance point was exactly half-way along the wire, we would be able to say that the e.m.f. of X was half that of the driver cell.

To calculate the unknown e.m.f. E_X we measure the length AY. Then we have:

$$E_X = \frac{AY}{AB} \times E_o$$

where E_o is the e.m.f. of the driver cell.

The potentiometer can be thought of as a potential divider because the point of contact Y divides the resistance wire into two parts, equivalent to the two resistors of a potential divider.

Comparing e.m.f.s with a potentiometer

When a potentiometer is balanced, no current flows from the cell being investigated. This means that its terminal p.d. is equal to its e.m.f.; we do not have to worry about any 'lost volts'. This is a great advantage that a potentiometer has over a voltmeter, which must draw a small current in order to work.

However, there is a problem: the driver cell is supplying current to the potentiometer, and so the p.d. between A and B will be less than the e.m.f. of the driver cell (some volts are lost because of its internal resistance). To overcome this problem, we use the potentiometer to **compare** p.d.s. Suppose we have two cells whose e.m.f.s E_X and E_Y we want to compare. Each is connected in turn to the potentiometer, giving balance points at C and D – see Figure 12.9. (In the diagram, you can see immediately that E_Y must be greater than E_X because D is further to the right than C.)

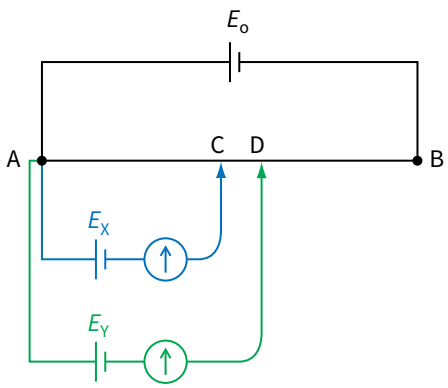


Figure 12.9 Comparing two e.m.f.s using a potentiometer.

The ratio of the e.m.f.s of the two cells will be equal to the ratio of the two lengths AC and AD:

$$\frac{E_X}{E_Y} = \frac{AC}{AD}$$

If one of the cells used has an accurately known e.m.f., the other can be calculated with the same degree of accuracy.

Comparing p.d.s

The same technique can be used to compare potential differences. For example, two resistors could be connected in series with a cell (Figure 12.10). The p.d. across one resistor is first connected to the potentiometer and the balance length found. This is repeated with the other resistor and the new balance point is found. The ratio of the lengths is the ratio of the p.d.s.

Since both resistors have the same current flowing through them, the ratio of the p.d.s is also the ratio of their resistances.

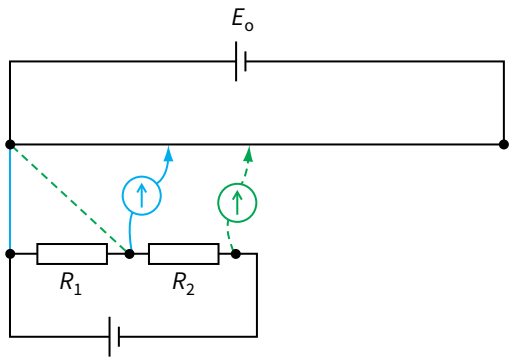


Figure 12.10 Comparing two potential differences using a potentiometer.

QUESTION

8 To make a potentiometer, a driver cell of e.m.f. 4.0 V is connected across a 1.00 m length of resistance wire.

a What is the potential difference across each 1 cm length of the wire? What length of wire has a p.d. of 1.0 V across it?

b A cell of unknown e.m.f. E is connected to the potentiometer and the balance point is found at a distance of 37.0 cm from the end of the wire to which the galvanometer is connected. Estimate the value of E. Explain why this can only be an estimate.

c A standard cell of e.m.f. 1.230 V gives a balance length of 31.2 cm. Use this value to obtain a more accurate value for E.

Summary

- A source of e.m.f., such as a battery, has an internal resistance. We can think of the source as having an internal resistance r in series with an e.m.f. E.
- The terminal p.d. of a source of e.m.f. is less than the e.m.f. because of 'lost volts' across the internal resistor:

  terminal p.d. = e.m.f. − 'lost volts'

  $$V = E - Ir$$

- A potential divider circuit consists of two or more resistors connected in series to a supply. The output voltage V_{out} across the resistor of resistance R_2 is given by:

  $$V_{out} = \left(\frac{R_2}{R_1 + R_2}\right) \times V_{in}$$

- A potentiometer can be used to compare potential differences.

End-of-chapter questions

1 A single cell of e.m.f. 1.5 V is connected across a 0.30 Ω resistor. The current in the circuit is 2.5 A.

a Calculate the terminal p.d. and explain why it is not equal to the e.m.f. of the cell. [2]

b Show that the internal resistance r of the cell is 0.30 Ω. [3]

c It is suggested that the power dissipated in the external resistor is a maximum when its resistance R is equal to the internal resistance r of the cell.

i Calculate the power dissipated when $R = r$. [1]

ii Show that the power dissipated when $R = 0.50\,\Omega$ and $R = 0.20\,\Omega$ is less than that dissipated when $R = r$, as the statement above suggests. [4]

2 A student is asked to compare the e.m.f.s of a standard cell and a test cell. He sets up the circuit shown in Figure 12.11 using the test cell.

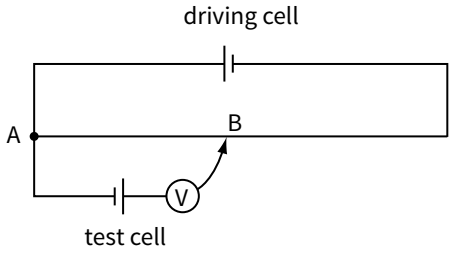


Figure 12.11 For End-of-chapter Question 2.

a **i** Explain why he is unable to find a balance point and state the change he must make in order to achieve balance. [2]

ii State how he would recognise the balance point. [1]

b He achieves balance when the distance AB is 22.5 cm. He repeats the experiment with a standard cell of e.m.f. of 1.434 V. The balance point using this cell is at 34.6 cm. Calculate the e.m.f. of the test cell. [2]

3 **a** Explain what is meant by the **internal resistance of a cell**. [2]

b When a cell is connected in series with a resistor of 2.00 Ω there is a current of 0.625 A. If a second resistor of 2.00 Ω is put in series with the first, the current falls to 0.341 A.

Calculate:

i the internal resistance of the cell [2]

ii the e.m.f. of the cell. [2]

c A car battery needs to supply a current of 200 A to turn over the starter motor. Explain why a battery made of a series of cells would not be suitable for a car battery. [2]

4 **a** State what is meant by the term **e.m.f. of a cell**. [2]

A student connects a high-resistance voltmeter across the terminals of a battery and observes a reading of 8.94 V. He then connects a 12 Ω resistor across the terminals and finds that the potential difference falls to 8.40 V.

b Explain why the measured voltage falls. [2]

c **i** Calculate the current in the circuit. [2]

ii Calculate the internal resistance of the cell. [2]

iii State any assumptions you made in your calculations. [1]

5 Figure 12.12 shows two circuits which could be used to act as a dimmer switch for a lamp.

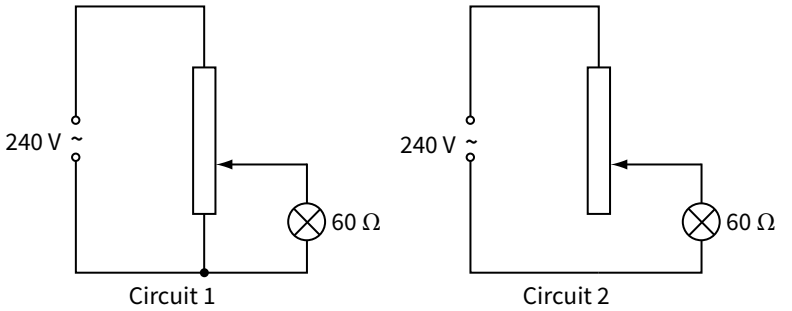


Figure 12.12 For End-of-chapter Question 5.

a Explain **one** advantage circuit 1 has over circuit 2. [2]

b i The lamp is rated at 60 W at 240 V. Calculate the resistance of the lamp filament at its normal operating temperature. [2]

ii State and explain how the resistance of the filament at room temperature would compare with the value calculated in **b i**. [2]

6 Figure 12.13 shows a potential divider. The battery has negligible internal resistance and the voltmeter has infinite resistance.

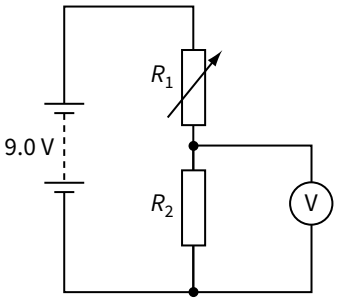


Figure 12.13 For End-of-chapter Question 6.

a State and explain how the reading on the voltmeter will change when the resistance of the variable resistor is increased. [2]

b Resistor R_2 has a resistance of 470 Ω. Calculate the value of the variable resistor when the reading on the voltmeter is 2.0 V. [2]

c The voltmeter is now replaced with one of resistance 2 kΩ. Calculate the reading on this voltmeter. [2]

7 Figure 12.14 shows a potentiometer circuit.

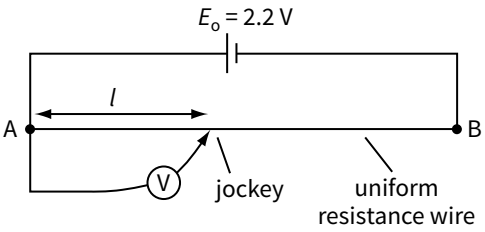


Figure 12.14 For End-of-chapter Question 7.

a i Sketch a graph of the reading on the voltmeter against I as the jockey is moved from point A to point B. [2]

ii State the readings on the voltmeter when the jockey is connected to A and when it is connected to B. (You may assume that the driver cell has negligible internal resistance.) [1]

iii Draw a circuit diagram to show how the potentiometer could be used to compare the e.m.f.s of two batteries. [3]

b When a pair of 4 Ω resistors are connected in series with a battery, there is a current of 0.60 A current through the battery. When the same two resistors are connected in parallel and then connected across the battery, there is a current of 1.50 A through it. Calculate the e.m.f. and the internal resistance of the battery. [4]

Chapter 13:
Waves

Learning outcomes

You should be able to:

- describe the motion of transverse and longitudinal waves
- describe waves in terms of their wavelength, amplitude, frequency, speed and intensity
- determine the frequency of sound waves using a cathode-ray oscilloscope
- state the wavelengths of the principal radiations of the electromagnetic spectrum

Vibrations making waves

The wind blowing across the surface of the sea produces waves. The surface of the water starts to move up and down, and these vibrations spread outwards – big waves may travel thousands of kilometres across the ocean before they break on a beach (Figure 13.1).



Figure 13.1 This photograph shows a wave breaking on the shore and dissipating the energy it has drawn from the wind in its journey across the ocean. The two scientists are 'storm chasers' who are recording the waves produced by a hurricane in the Gulf of Mexico.

Describing waves

When you pluck the string of a guitar, it vibrates. The vibrations create a wave in the air which we call sound. In fact, all vibrations produce **waves** of one type or another (Figure 13.2). Waves that move through a material (or a vacuum) are called **progressive waves**. A progressive wave transfers energy from one position to another.

At the seaside, a wave is what we see on the surface of the sea. The water moves around and a wave travels across the surface. In physics, we extend the idea of a wave to describe many other phenomena, including light, sound, etc. We do this by imagining an idealised wave, as shown in Figure 13.3 – you will never see such a perfect wave on the sea!



Figure 13.2 Radio telescopes detect radio waves from distant stars and galaxies; a rainbow is an effect caused by the reflection and refraction of light waves by water droplets in the atmosphere.

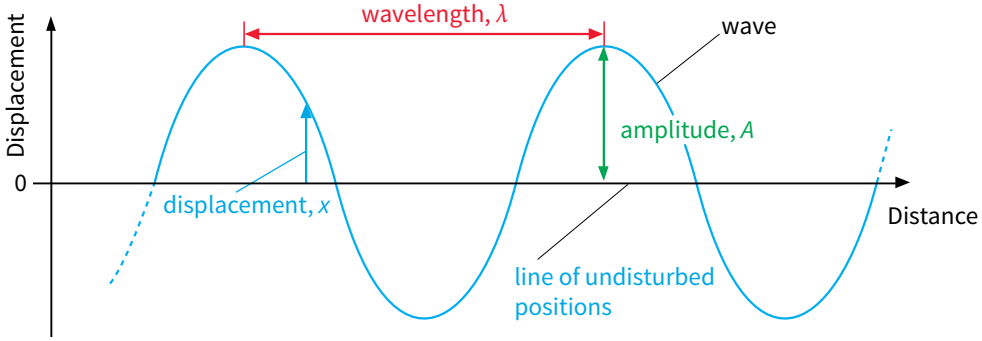


Figure 13.3 A displacement–distance graph illustrating the terms displacement, amplitude and wavelength.

Figure 13.3 or a similar graph of displacement against time illustrates the following important definitions about waves and wave motion:

- The distance of a point on the wave from its undisturbed position or equilibrium position is called the **displacement** x.
- The maximum displacement of any point on the wave from its undisturbed position is called the **amplitude** A. The amplitude of a wave on the sea is measured in units of distance, e.g. metres. The greater the amplitude of the wave, the louder the sound or the rougher the sea!
- The distance from any point on a wave to the next exactly similar point (e.g. crest to crest) is called the **wavelength** λ (the Greek letter lambda). The wavelength of a wave on the sea is measured in units of distance, e.g. metres.
- The time taken for one complete oscillation of a point in a wave is called the **period** T. It is the time taken for a point to move from one particular position and return to that same position, moving in the same direction. It is measured in units of time, e.g. seconds..
- The number of oscillations per unit time of a point in a wave is called its **frequency** f. For sound waves, the higher the frequency of a musical note, the higher is its pitch. Frequency is measured in hertz (Hz), where 1 Hz = one oscillation per second (1 kHz = 10^3 Hz and 1 MHz = 10^6 Hz). The frequency f of a wave is the reciprocal of the period T:

$$f = \frac{1}{T}$$

Waves are called **mechanical waves** if they need a substance (medium) through which to travel. Sound is one example of such a wave. Other cases are waves on strings, seismic waves and water waves (Figure 13.4).

Some properties of typical waves are given on page 183 in Table 13.1.



Figure 13.4 The impact of a droplet on the surface of a liquid creates a vibration, which in turn gives rise to waves on the surface.

QUESTION

1 Determine the wavelength and amplitude of each of the two waves shown in Figure 13.5.

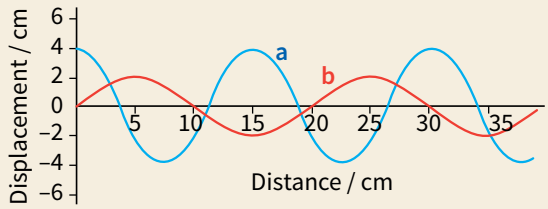


Figure 13.5 Two waves – for Question 1.

BOX 13.1: Measuring frequency

You can measure the frequency of sound waves using a cathode-ray oscilloscope (c.r.o.). Figure 13.6 shows how.

A microphone is connected to the input of the c.r.o. Sound waves are captured by the microphone and converted into a varying voltage which has the same frequency as the sound waves. This voltage is displayed on the c.r.o. screen.

It is best to think of a c.r.o. as a voltmeter which is capable of displaying a rapidly varying voltage. To do this, its spot moves across the screen at a steady speed, set by the time-base control. At the same time, the spot moves up and down according to the voltage of the input.

Hence the display on the screen is a graph of the varying voltage, with time on the (horizontal) x-axis. If we know the horizontal scale, we can determine the period and hence the frequency of the sound wave. Worked example **1** shows how to do this. (In Chapter 15 we will look at one method of measuring the wavelength of sound waves.)

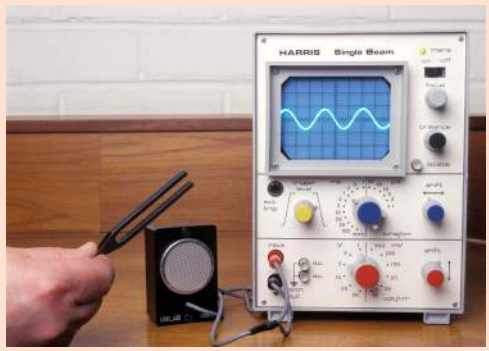


Figure 13.6 Measuring the frequency of sound waves from a tuning fork.

WORKED EXAMPLE

1 Figure 13.7 shows the trace on an oscilloscope screen when sound waves are detected by a microphone. The time-base is set at 1 ms div^{-1}. Determine the frequency of the sound waves.

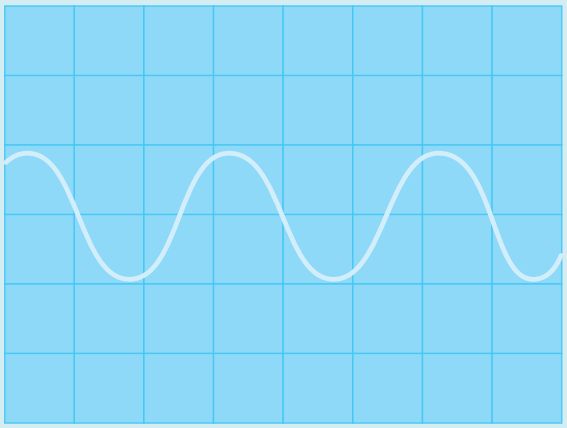


Figure 13.7 A c.r.o. trace – what is the frequency of these waves?

Step 1 Determine the period of the waves on the screen, in scale divisions. From Figure 13.7, you can see that one complete wave occupies three scale divisions (div).

period $T = 3.0$ div

Step 2 Determine the time interval represented by each scale division. The time-base control is set at 1 ms div^{-1}, so:

scale factor = 1 ms div^{-1}

Step 3 Convert the period in divisions to ms:

period $T = 3.0 \text{ div} \times 1 \text{ ms div}^{-1}$

$= 3.0 \text{ ms} = 3.0 \times 10^{-3} \text{ s}$

Hint: *Notice how div and div^{-1} cancel out.*

Step 4 Calculate the frequency from the period:

$$\text{frequency } f = \frac{1}{T} = \frac{1}{3.0 \times 10^{-3}} = 333 \text{ Hz}$$

So the wave frequency is approximately 330 Hz.

QUESTION

2 When a sound wave is displayed on a c.r.o. screen, two complete waves occupy five scale divisions. The calibrated time-base is set on 0.005 s per division. Determine the frequency of the waves.

Longitudinal and transverse waves

There are two distinct types of wave, **longitudinal** and **transverse**. Both can be demonstrated using a slinky spring lying along a bench.

Push the end of the spring back and forth; the segments of the spring become compressed and then stretched out, along the length of the spring. Wave pulses run along the spring. These are longitudinal waves.

Waggle the end of the slinky spring from side to side. The segments of the spring move from side to side as the wave travels along the spring. These are transverse waves.

So the distinction between longitudinal and transverse waves is as follows:

- In longitudinal waves, the particles of the medium vibrate **parallel** to the direction of the wave velocity.
- In transverse waves, the particles of the medium vibrate at **right angles** to the direction of the wave velocity.

Sound waves are an example of a longitudinal wave. Light and all other electromagnetic waves are transverse waves. Waves in water are quite complex. Particles of the water may move both up and down and from side to side as a water wave travels through the water. You can investigate water waves in a ripple tank. There is more about water waves in Table 13.1 (page 183) and in Chapter 14.

Representing waves

Figure 13.8 shows how we can represent longitudinal and transverse waves. The longitudinal wave shows how the material through which it is travelling is alternately compressed and expanded. This gives rise to high and low pressure regions, respectively.

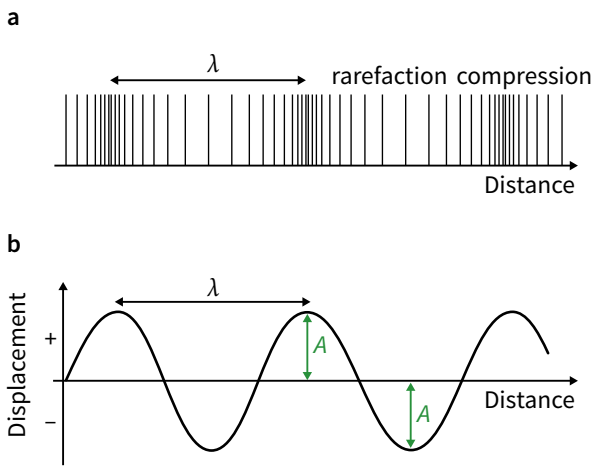


Figure 13.8 **a** Longitudinal waves and **b** transverse waves. A = amplitude, λ = wavelength.

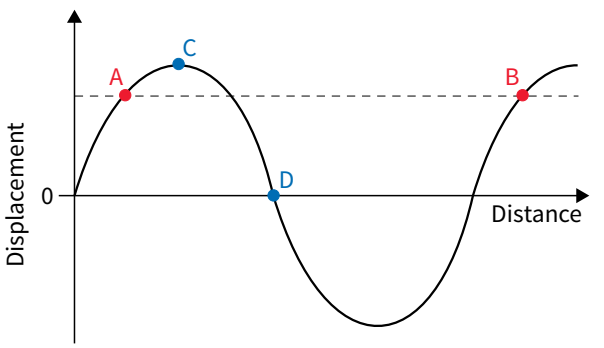
However, this is rather difficult to draw, so you will often see a longitudinal wave represented as if it were a sine wave. The displacement referred to in the graph is the displacement of the particles in the wave.

We can compare the **compressions** and **rarefactions** (or expansions) of the longitudinal wave with the peaks and troughs of the transverse wave.

Phase and phase difference

All points along a wave have the same pattern of vibration. However, different points do not necessarily vibrate in step with one another. As one point on a wave vibrates, the point next to it vibrates slightly out-of-step with it. We say that they vibrate out of phase with each other – there is a **phase difference** between them. This is the amount by which one oscillation leads or lags behind another.

Phase difference is measured in degrees. As you can see from Figure 13.9, two points A and B, with a separation of one whole wavelength λ, vibrate in phase with each other. The phase difference between these two points is 360°. (You can also say it is 0°.) The phase difference between any other two points between A and B can have any value between 0° and 360°. A complete cycle of the wave is thought of as 360°. In Chapter 14 we will see what it means to say that two waves are 'in phase' or 'out of phase' with one another.



Points A and B are vibrating; they have a phase difference of 360° or 0°. They are 'in phase'

Points C and D have a phase difference of 90°.

Figure 13.9 Different points along a wave have different phases.

QUESTION

3 Using axes of displacement and distance, sketch two waves A and B such that A has twice the wavelength and half the amplitude of B.

Wave energy

It is important to realise that, for both types of mechanical wave, the particles that make up the material through which the wave is travelling do not move along – they only oscillate about a fixed point. It is **energy** that is transmitted by the wave. Each particle vibrates; as it does so, it pushes its neighbour, transferring energy to it. Then that particle pushes its neighbour, which pushes its neighbour. In this way, energy is transmitted from one particle to the next, to the next, and so on down the line.

Intensity

The term **intensity** has a very precise meaning in physics. The intensity of a wave is defined as the rate of energy transmitted (i.e. power) per unit area at right angles to the wave velocity.

$$\text{intensity} = \frac{\text{power}}{\text{cross-sectional area}}$$

Intensity is measured in watts per square metre ($W\,m^{-2}$). For example, when the Sun is directly overhead, the intensity of its radiation is about $1.0\,kW\,m^{-2}$ (1 kilowatt per square metre). This means that energy arrives at the rate of about 1 kW ($1000\,J\,s^{-1}$) on each square metre of the surface of the Earth. At the top of the atmosphere, the intensity of sunlight is greater, about $1.37\,kW\,m^{-2}$.

QUESTION

4 A 100 W lamp emits electromagnetic radiation in all directions. Assuming the lamp to be a point source, calculate the intensity of the radiation:

a at a distance of 1.0 m from the lamp

b at a distance of 2.0 m from the lamp.

Hint: *Think of the area of a sphere at each of the two radii.*

Intensity and amplitude

The intensity of a wave generally decreases as it travels along. There are two reasons for this:

- The wave may 'spread out' (as in the example of light spreading out from a light bulb in Question 4).
- The wave may be absorbed or scattered (as when light passes through the Earth's atmosphere).

As a wave spreads out, its amplitude decreases. This suggests that the intensity I of a wave is related to its amplitude A. In fact, intensity is proportional to the square of the amplitude:

$$\text{intensity} \propto \text{amplitude}^2 \quad (I \propto A^2)$$

The relationship also implies that, for a particular wave:

$$\frac{\text{intensity}}{\text{amplitude}^2} = \text{constant}$$

So, if one wave has **twice** the amplitude of another, it has **four** times the intensity. This means that it is carrying energy at four times the rate.

QUESTION

5 Waves from a source have an amplitude of 5.0 cm and an intensity of 400 W m^{-2}.

a The amplitude of the waves is increased to 10.0 cm. What is their intensity now?

b The intensity of the waves is decreased to 100 W m^{-2}. What is their amplitude?

Wave speed

The speed with which energy is transmitted by a wave is known as the wave speed v. This is measured in m s^{-1}. The wave speed for sound in air at a pressure of 10^5 Pa and a temperature of 0 °C is about 340 m s^{-1}, while for light in a vacuum it is almost 300 000 000 m s^{-1}.

The wave equation

An important equation connecting the speed v of a wave with its frequency f and wavelength λ can be determined as follows. We can find the speed of the wave using:

$$\text{speed} = \frac{\text{distance}}{\text{time}}$$

But a wave will travel a distance of one whole wavelength in a time equal to one period T. So:

$$\text{wave speed} = \frac{\text{wavelength}}{\text{period}}$$

or

$$v = \frac{\lambda}{T}$$

$$v = (\frac{1}{T}) \times \lambda$$

However, $f = \frac{1}{T}$ and so:

wave speed = frequency × wavelength

$$v = f \times \lambda$$

A numerical example may help to make this clear. Imagine a wave of frequency 5 Hz and wavelength 3 m going past you. In 1 s, five complete wave cycles, each of length 3 m, go past. So the total length of the waves going past in 1 s is 15 m. The distance covered by the wave in one second is its speed, therefore the speed of the wave is 15 m s^{-1}.

Clearly, for a given speed of wave, the greater the wavelength, the smaller the frequency and vice versa. The speed of sound in air is constant (for a given temperature and pressure). The wavelength of sound can be made smaller by increasing the frequency of the source of sound.

Table 13.1 gives typical values of speed (v), frequency (f) and wavelength (λ) for some mechanical waves. You can check for yourself that $v = f\lambda$ is satisfied.

	Water waves in a ripple tank	Sound waves in air	Waves on a slinky spring
Speed / m s^{-1}	about 0.12	about 300	about 1
Frequency / Hz	about 6	20 to 20 000 (limits of human hearing)	about 2
Wavelength / m	about 0.2	15 to 0.015	about 0.5

Table 13.1 Speed (v), frequency (f) and wavelength (λ) data for some mechanical waves readily investigated in the laboratory.

WORKED EXAMPLE

2 Middle C on a piano tuned to concert pitch should have a frequency of 264 Hz (Figure 13.10). If the speed of sound is 330 m s^{-1}, calculate the wavelength of the sound produced when this key is played.

Step 1 We use the above equation in slightly rewritten form:

$$\text{wavelength} = \frac{\text{speed}}{\text{frequency}}$$

Step 2 Substituting the values for middle C we get:

$$\text{wavelength} = \frac{330}{264} = 1.25 \text{ m}$$

The human ear can detect sounds of frequencies between 20 Hz and 20 kHz, i.e. with wavelengths between 15 m and 15 mm.



Figure 13.10 Each string in a piano produces a different note.

QUESTIONS

6 Sound is a mechanical wave that can be transmitted through a solid. Calculate the frequency of sound of wavelength 0.25 m that travels through steel at a speed of 5060 m s^{-1}.

7 A cello string vibrates with a frequency of 64 Hz. Calculate the speed of the transverse waves on the string given that the wavelength is 140 cm.

8 An oscillator is used to send waves along a cord. Four complete wave cycles fit on a 20 cm length of the cord when the frequency of the oscillator is 30 Hz. For this wave, calculate:
   a its wavelength
   b its frequency
   c its speed.

9 Copy and complete Table 13.2. (You may assume that the speed of radio waves is 3.0 × 10^8 m s^{-1}.)

Station	Wavelength / m	Frequency / MHz
Radio A (FM)		97.6
Radio B (FM)		94.6
Radio B (LW)	1515	
Radio C (MW)	693	

Table 13.2 For Question 9.

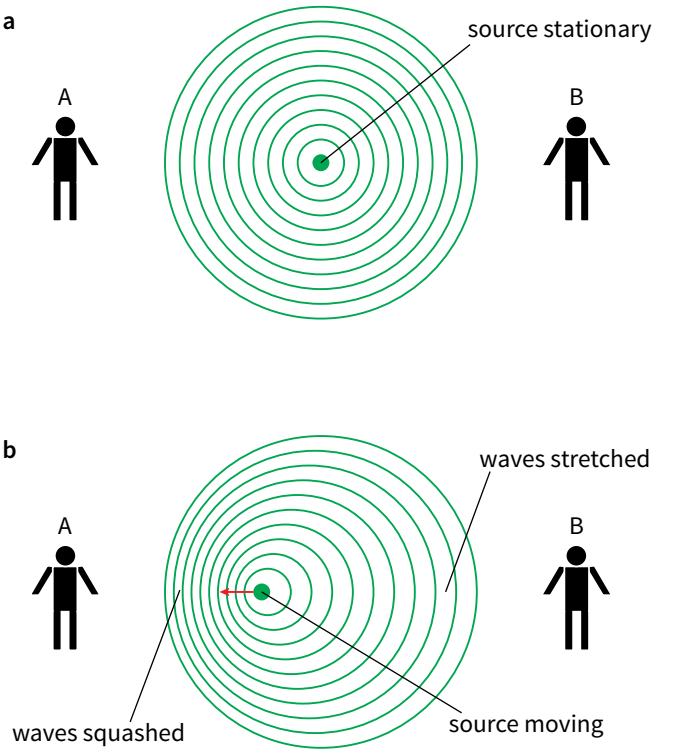


Figure 13.11 Sound waves, represented by wavefronts, emitted at constant frequency by **a** a stationary source, and **b** a source moving with speed v_s.

The Doppler effect

You may have noticed a change in pitch of the note heard when an emergency vehicle passes you while sounding its siren. The pitch is higher as it approaches you, and lower as it recedes into the distance. This is an example of the **Doppler effect**; you can hear the same thing if a train passes at speed while sounding its whistle.

Figure 13.11 shows why this change in frequency is observed. It shows a source of sound emitting waves with a constant frequency f_s, together with two observers A and B.

- If the source is stationary (Figure 13.11a), waves arrive at A and B at the same rate, and so both observers hear sounds of the same frequency f_s.
- If the source is moving towards A and away from B (Figure 13.11b), the situation is different. From the diagram you can see that the waves are squashed together in the direction of A and spread apart in the direction of B.

Observer A will observe waves whose wavelength is shortened. More waves per second arrive at A, and so A observes a sound of higher frequency than f_s. Similarly, the waves arriving at B have been stretched out and B will observe a frequency lower than f_s.

An equation for observed frequency

There are two different speeds involved in this situation. The source is moving with speed v_s. The sound waves travel through the air with speed v, which is unaffected by the speed of the source. (Remember, the speed of a wave depends only on the medium it is travelling through.)

The frequency and wavelength observed by an observer will change according to the speed v_s at which the source is moving. Figure 13.12 shows how we can calculate the observed wavelength λ_o and the observed frequency f_o.

The wave trains shown in Figure 13.12 represent the f_s waves emitted by the source in 1 s. Provided the source is stationary (Figure 13.12a), the length of this train is equal to the wave speed v since this is the distance the first wave travels away from the source in 1 s. The wavelength observed by the observer is simply $\lambda_o = \frac{v}{f_s}$.

The situation is different when the source is moving away from the observer (Figure 13.12b). In 1 s, the source moves a distance v_s. Now the train of f_s waves will have a length equal to $v + v_s$.

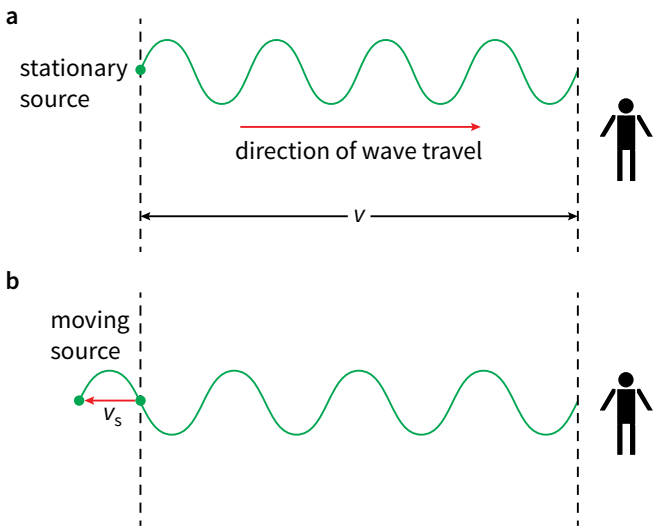


Figure 13.12 Sound waves, emitted at constant frequency by **a** a stationary source, and **b** a source moving with speed v_s away from the observer.

The observed wavelength is now given by $\lambda_o = \frac{(v + v_s)}{f_s}$. The observed frequency is given by:

$$f_o = \frac{v}{\lambda_o} = \frac{f_s \times v}{(v + v_s)}$$

This tells us how to calculate the observed frequency when the source is moving away from the observer. If the source is moving towards the observer, the train of f_s waves will be compressed into a shorter length equal to $v - v_s$, and the observed frequency will be given by:

$$f_o = \frac{v}{\lambda_o} = \frac{f_s \times v}{(v - v_s)}$$

We can combine these two equations to give a single equation for the Doppler shift in frequency due to a moving source:

$$\text{observed frequency } f_o = \frac{f_s \times v}{(v \pm v_s)}$$

where the plus sign applies to a receding source and the minus sign to an approaching source. Note these important points:

- The frequency f_s of the source is not affected by the movement of the source – it still emits f_s waves per second.
- The speed v of the waves as they travel through the air (or other medium) is also unaffected by the movement of the source.

Note that a Doppler effect can also be heard when an observer is moving relative to a stationary source, and also when both source and observer are moving. There is more about the Doppler effect and light later in this chapter.

WORKED EXAMPLE

3 A train with a whistle that emits a note of frequency 800 Hz is approaching an observer at a speed of 60 m s^{-1}. What frequency of note will the observer hear? (Speed of sound in air = 330 m s^{-1}.)

Step 1 Select the appropriate form of the Doppler equation. Here the source is approaching the observer so we choose the minus sign:

$$f_o = \frac{f_s \times v}{(v - v_s)}$$

Step 2 Substitute values from the question and solve:

$$f_o = \frac{800 \times 330}{(330 - 60)} = \frac{800 \times 330}{270}$$

$$= 978 \text{ Hz}$$

So the observer hears a note whose pitch is raised significantly, because the train is travelling at a speed which is an appreciable fraction of the speed of sound.

QUESTION

10 A plane's engine emits a note of constant frequency 120 Hz. It is flying away from an observer at a speed of 80 m s^{-1}. Determine:

a the observed wavelength of the sound received by the observer

b its observed frequency.

(Speed of sound in air = 330 m s^{-1}.)

Electromagnetic waves

You should be familiar with the idea that light is a region of the **electromagnetic spectrum**. It is not immediately obvious that light has any connection at all with electricity, magnetism and waves. These topics had been the subject of study by physicists for centuries before the connections between them became apparent.

An electric current always gives rise to a **magnetic field** (this is known as electromagnetism). A magnetic field is created by any **moving** charged particles such as electrons. Similarly, a changing magnetic field will induce a current in a nearby conductor. These observations led to the unification of the theories of electricity and magnetism by Michael Faraday in the mid-19th century. A vast technology based on the theories of electromagnetism developed rapidly, and continues to expand today (Figure 13.13).

Faraday's studies were extended by James Clerk Maxwell. He produced mathematical equations that predicted that a changing electric or magnetic field would give rise to waves travelling through space. When he calculated the speed of these waves, it turned out to be the known speed of light. He concluded that light is a wave, known as an **electromagnetic wave**, that can travel through space (including a vacuum) as a disturbance of electric and magnetic fields.

Faraday had unified electricity and magnetism; now Maxwell had unified electromagnetism and light. In the 20th century, Abdus Salam (Figure 13.14) managed to unify electromagnetic forces with the weak nuclear force, responsible for radioactive decay. Physicists continue to strive to unify the big ideas of physics; you may occasionally hear talk of **a theory of everything**. This would not truly explain **everything**, but it would explain all known forces, as well as the existence of the various fundamental particles of matter.



Figure 13.13 These telecommunications masts are situated 4500 metres above sea level in Ecuador. They transmit microwaves, a form of electromagnetic radiation, across the mountain range of the Andes.



Figure 13.14 Abdus Salam, the Pakistani physicist, won the 1979 Nobel Prize for Physics for his work on unification of the fundamental forces.

Electromagnetic radiation

By the end of the 19th century, several types of electromagnetic wave had been discovered:

- radio waves – these were discovered by Heinrich Hertz when he was investigating electrical sparks
- infrared and ultraviolet waves – these lie beyond either end of the visible spectrum
- X-rays – these were discovered by Wilhelm Röntgen and were produced when a beam of electrons collided with a metal target such as tungsten
- γ-rays – these were discovered by Henri Becquerel when he was investigating radioactive substances.

We now regard all of these types of radiation as parts of the same electromagnetic spectrum, and we know that they can be produced in a variety of different ways.

The speed of light

James Clerk Maxwell showed that the speed c of electromagnetic radiation in a vacuum (free space) was independent of the frequency of the waves. In other words, all types of electromagnetic wave travel at the same speed in a vacuum. In the SI system of units, c has the value:

$$c = 299\,792\,458\ \text{m s}^{-1}$$

The approximate value for the speed of light in a vacuum (often used in calculations) is $3.0 \times 10^8\ \text{m s}^{-1}$.

The wavelength λ and frequency f of the radiation are related by the equation:

$$c = f\lambda$$

When light travels from a vacuum into a material medium such as glass, its speed **decreases** but its frequency **remains the same**, and so we conclude that its wavelength must decrease. We often think of different forms of electromagnetic radiation as being characterised by their different wavelengths, but it is better to think of their different frequencies as being their fundamental characteristic, since their wavelengths depend on the medium through which they are travelling.

Light waves show the Doppler effect in the same way that sound waves do. So, for example, if an astronomer looks at the light from a distant star which is receding from Earth at speed v_s, its wavelength will be increased and its frequency will be decreased. The change in wavelength $\Delta\lambda$ is simply given by $\Delta\lambda/\lambda = v_s/c$.

Since longer wavelengths are towards the red end of the visible spectrum, the light from the star will look redder than if it were stationary. This is the origin of the 'red shift' which allows astronomers to determine the speed at which stars and galaxies are moving away from us, and which first provided evidence that the Universe is expanding.

The observed frequency of light from a moving source can be calculated using the same equation as for sound, $f_{obs} = \frac{f \times c}{(c \pm v_s)}$ but there is an important condition. The speed of the source v_s must be small compared to the speed of light c. For speeds approaching c, the equation must be altered to take account of the theory of relativity.

QUESTIONS

11 Red light of wavelength 700 nm in a vacuum travels into glass, where its speed decreases to $2.0 \times 10^8\,m\,s^{-1}$. Determine:

a the frequency of the light in a vacuum

b its frequency and wavelength in the glass.

12 An astronomer observes light from a distant star. A particular line in its spectrum has a wavelength of 550 nm. When measured in the laboratory, the same spectral line has a wavelength of 535 nm. Determine:

a the change in wavelength of the spectral line

b the speed of the star

c the direction of movement of the star (towards or away from the observer).

(Speed of light in free space = $3.0 \times 10^8\,m\,s^{-1}$)

Orders of magnitude

Table 13.3 shows the approximate ranges of wavelengths in a vacuum of the principal bands which make up the electromagnetic spectrum. This information is shown as a diagram in Figure 13.15.

Here are some points to note:

- There are no clear divisions between the different ranges or bands in the spectrum. The divisions shown in Table 13.3 are somewhat arbitrary.
- The naming of subdivisions is also arbitrary. For example, microwaves are sometimes regarded as a subdivision of radio waves.
- The ranges of X-rays and γ-rays overlap. The distinction is that X-rays are produced when electrons decelerate rapidly or when they hit a target metal at high speeds. γ-rays are produced by nuclear reactions such as radioactive decay. There is no difference whatsoever in the radiation between an X-ray and a γ-ray of wavelength, say, 10^{-11} m.

Radiation	Wavelength range / m
radio waves	$>10^6$ to 10^{-1}
microwaves	10^{-1} to 10^{-3}
infrared	10^{-3} to 7×10^{-7}
visible	7×10^{-7} (red) to 4×10^{-7} (violet)
ultraviolet	4×10^{-7} to 10^{-8}
X-rays	10^{-8} to 10^{-13}
γ-rays	10^{-10} to 10^{-16}

Table 13.3 Wavelengths (in a vacuum) of the electromagnetic spectrum.

QUESTIONS

13 Copy Table 13.3. Add a third column showing the range of frequencies of each type of radiation.

14 Study Table 13.3 and answer the questions.

a Which type of radiation has the narrowest range of wavelengths?

b Which has the second narrowest range?

c What is the range of wavelengths of microwaves, in millimetres?

d What is the range of wavelengths of visible light, in nanometres?

e What is the frequency range of visible light?

15 For each of the following wavelengths measured in a vacuum, state the type of electromagnetic radiation to which it corresponds.

a 1 km **c** 5000 nm **e** 50 nm

b 3 cm **d** 500 nm **f** 10^{-12} m

16 For each of the following frequencies, state the type of electromagnetic radiation to which it corresponds.

a 200 kHz **c** 5×10^{14} Hz

b 100 MHz **d** 10^{18} Hz

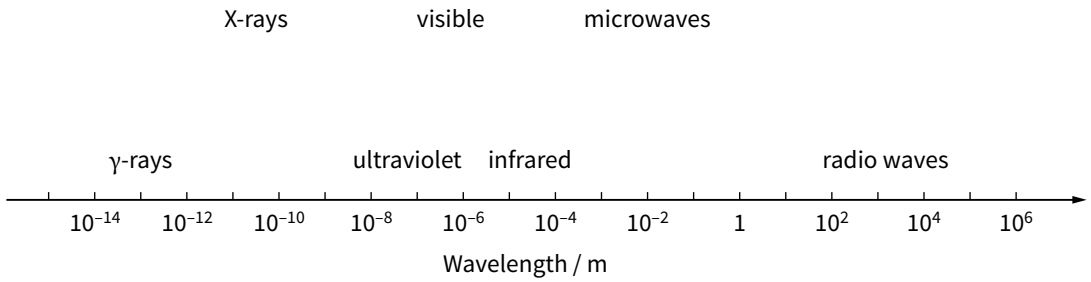


Figure 13.15 Wavelengths of the electromagnetic spectrum. The boundaries between some regions are fuzzy.

The nature of electromagnetic waves

An electromagnetic wave is a disturbance in the electric and magnetic fields in space. Figure 13.16 shows how we can represent such a wave. In this diagram, the wave is travelling from left to right.

The electric field is shown oscillating in the vertical plane. The magnetic field is shown oscillating in the horizontal plane. These are arbitrary choices; the point is that the two fields vary at right angles to each other, and also at right angles to the direction in which the wave is travelling. This shows that electromagnetic waves are transverse waves.

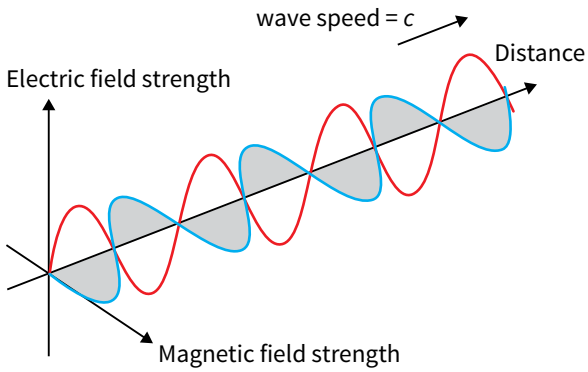


Figure 13.16 An electromagnetic wave is a periodic variation in electric and magnetic fields.

Summary

- Mechanical waves are produced by vibrating objects.
- A progressive wave carries energy from one place to another.
- Two points on a wave separated by a distance of one wavelength have a phase difference of 0° or 360°.
- There are two types of wave – longitudinal and transverse. Longitudinal waves have vibrations parallel to the direction in which the wave travels, whereas transverse waves have vibrations at right angles to the direction in which the wave travels.
- The frequency f of a wave is related to its period T by the equation:
  $$f = \frac{1}{T}$$
- The frequency of a sound wave can be measured using a calibrated cathode-ray oscilloscope.
- The speed of all waves is given by the wave equation:
  wave speed = frequency × wavelength
  $$v = f\lambda$$
- The Doppler effect is the change in an observed wave frequency when a source moves with speed v_s. The observed frequency is given by:
  $$f_o = \frac{f_s \times v}{(v \pm v_s)}$$
- The intensity of a wave is defined as the wave power transmitted per unit area at right angles to the wave velocity. Hence:
  $$\text{intensity} = \frac{\text{power}}{\text{cross-sectional area}}$$
  Intensity has units of W m^{-2}.
- The intensity I of a wave is proportional to the square of the amplitude A ($I \propto A^2$).
- All electromagnetic waves travel at the same speed of $3.0 \times 10^8\ \text{m s}^{-1}$ in a vacuum, but have different wavelengths and frequencies.
- The regions of the electromagnetic spectrum in order of increasing wavelength are: γ-rays, X-rays, ultraviolet, visible, infrared, microwaves and radio waves.

End-of-chapter questions

1 Figure 13.17 shows the screen of an oscilloscope. The time-base of the oscilloscope is set at 500 μs div^{-1}. Calculate the time period of the signal and hence its frequency. [3]

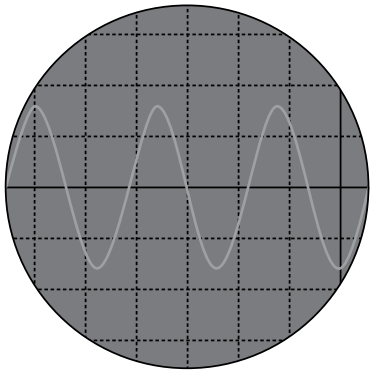


Figure 13.17 For End-of-chapter Question 1.

2 **a** State **two** main properties of electromagnetic waves. [2]

b State **one** major difference between microwaves and radio waves. [1]

c **i** Estimate the wavelength in metres of X-rays. [1]

ii Use your answer to **i** to determine the frequency of the X-rays. [1]

3 A student is sitting on the beach, observing a power boat moving at speed on the sea. The boat has a siren emitting a constant sound of frequency 420 Hz.

The boat moves around in a circular path with a speed of 25 m s^{-1}. The student notices that the pitch of the siren changes with a regular pattern.

a Explain why the pitch of the siren changes, as observed by the student. [1]

b Determine the maximum and minimum frequencies that the student will hear. [4]

c At which point in the boat's motion will the student hear the most high-pitched note? [1]
(Speed of sound in air = 330 m s^{-1})

4 Figure 13.18 shows some air particles as a sound wave passes.



Figure 13.18 For End-of-chapter Question 4.

a On a copy of the diagram, mark:

i a region of the wave which shows a compression – label it C [1]

ii a region of the wave which shows a rarefaction – label it R. [1]

b Describe how the particle labelled P moves as the wave passes. [2]

c The sound wave has a frequency of 240 Hz. Explain, in terms of the movement of an individual particle, what this means. [2]

d The wave speed of the sound is 320 m s^{-1}. Calculate the wavelength of the wave. [2]

5 The diagram represents a wave on the surface of water at one instant of time. The speed of the wave is 7.0 cm s^{-1}. Calculate the amplitude, wavelength and time period for one oscillation.

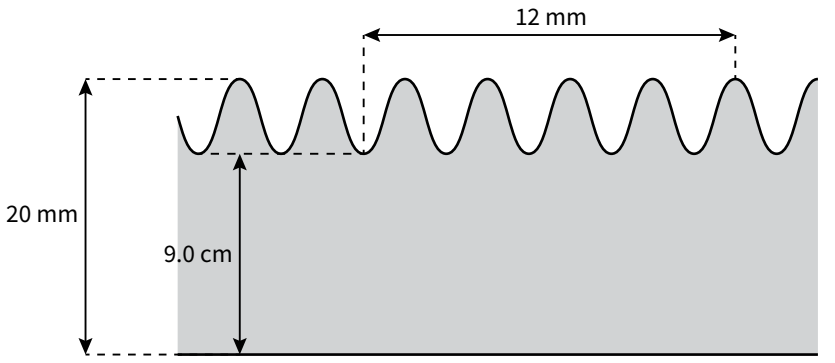


Figure 13.19 For End-of-chapter Question 5.

6 **a** Light is referred to as a type of electromagnetic wave. Explain what is meant by the term **electromagnetic wave**. [2]

b **i** Two stars emit radiation with the same power. Star A is twice as far from the Earth as star B. Explain how the intensities of the radiations compare to an observer on the Earth. [2]

ii State how the amplitudes of the two signals received by the observer compare. [2]

c The main signal from star A is in the ultraviolet region of the spectrum with a wavelength of 7.5×10^{8} m. Calculate the wavelength of this radiation. [2]

7 Figure 13.20 shows a loudspeaker producing a sound and a microphone connected to a cathode-ray oscilloscope.

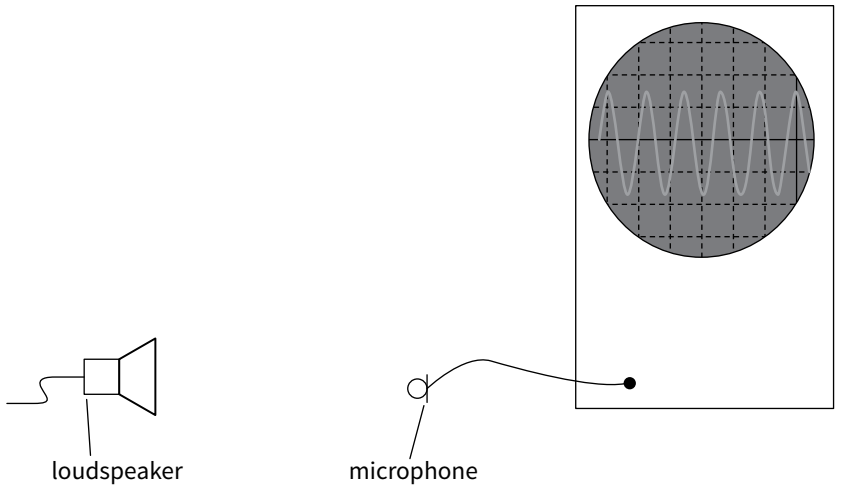


Figure 13.20 For End-of-chapter Question 7.

a Sound is described as a longitudinal wave. Explain how the trace on the oscilloscope relates to the movements of the air particles in the sound wave. [3]

b The time-base on the oscilloscope is set at 5 ms div^{-1}. Calculate the frequency of the wave. [2]

c The wavelength of the wave is found to be 1.98 m. Calculate the speed of the wave. [2]

8 Figure 13.21 shows the traces of two sound waves displayed on an oscilloscope screen.

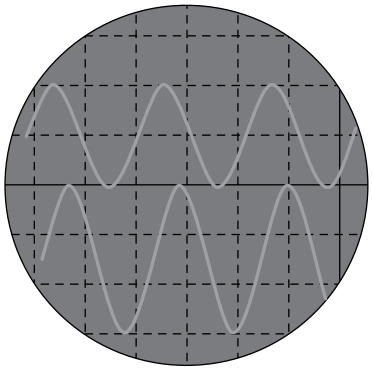


Figure 13.21 For End-of-chapter Question 8.

a State and explain how the wavelengths of the two waves compare. [2]

b Explain the meaning of the term **phase difference**. Illustrate your answer by estimating the phase difference between the two waves shown on the screen. [3]

c Calculate the ratio of the intensities of the two waves. [2]

9 **a** State what is meant by the **Doppler effect**. [2]

b A star is moving away from the Earth at a speed of $6.4 \times 10^5\,\text{m s}^{-1}$. Light of frequency $6.500 \times 10^{14}\,\text{Hz}$ emitted by the star is received on Earth. Calculate the change in frequency caused by the Doppler effect. Give your answer to 2 significant figures. [3]

c When looking at distant stars and galaxies, the further a star or galaxy is from Earth, the more the light emitted from the star is shifted towards the red end of the spectrum.

i State and explain whether the stars and galaxies are moving towards or away from the Earth. [2]

ii Explain what the greater shift in frequency for stars that are further away implies. [1]

10 The Doppler effect can be used to measure the speed of blood. Ultrasound, which is sound of high frequency, is passed from a transmitter into the body, where it reflects off particles in the blood. The shift in frequency is measured by a stationary detector, placed outside the body and close to the transmitter.

In one patient, particles in the blood are moving at a speed of $30\,\text{cm s}^{-1}$ in a direction directly away from the transmitter. The speed of ultrasound in the body is $1500\,\text{cm s}^{-1}$.

This situation is partly modelled by considering the particles to be emitting sound of frequency 4.000 MHz as they move away from the detector. This sound passes to the detector outside the body and the frequency measured by the detector is not 4.000 MHz.

a **i** State whether the frequency received by the stationary detector is higher or lower than the frequency emitted by the moving particles. [1]

ii Explain your answer to **i**. [3]

b Calculate the difference between the frequency emitted by the moving particles and the frequency measured by the detector. [3]

c Suggest why there is also a frequency difference between the sound received by the particles and the sound emitted by the transmitter. [1]

Chapter 14: Superposition of waves

Learning outcomes

You should be able to:

- explain and use the principle of superposition of waves
- explain experiments that show diffraction and interference
- solve problems involving two-slit and multiple-slit interference

Combining waves

Light travels as waves and can produce beautiful, natural effects such as the iridescent colours of a butterfly's wing (Figure 14.1). However, these colours do not come from pigments in the wing. Instead, they arise when light waves, scattered from different points on the wing, meet in your eye and combine to produce the colours that we see.



Figure 14.1 The iridescent colours on a butterfly's wing demonstrate the beauty of nature. In this chapter we will study the effect known as interference, which leads to the production of these glorious colours.

The principle of superposition of waves

In Chapter 13, we studied the production of waves and the difference between longitudinal and transverse waves. In this chapter we are going to consider what happens when two or more waves meet at a point in space and combine together (Figure 14.2).

So what happens when two waves arrive together at the same place? We can answer this from our everyday experience. What happens when the beams of light waves from two torches cross over? They pass straight through one another. Similarly, sound waves pass through one another, apparently without affecting each other. This is very different from the behaviour of **particles**. Two bullets meeting in mid-air would ricochet off one another in a very un-wave-like way. If we look carefully at how two sets of waves interact when they meet, we find some surprising results.

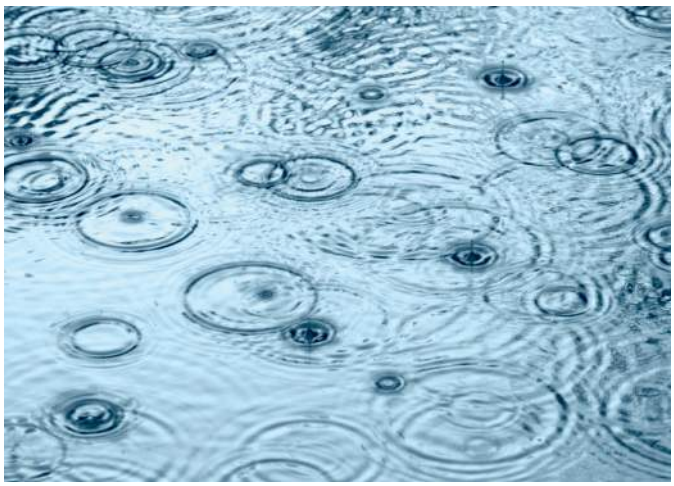


Figure 14.2 Here we see ripples produced when drops of water fall into a swimming pool. The ripples overlap to produce a complex pattern of crests and troughs.

When two waves meet they combine, with the displacements of the two waves adding together. Figure 14.3 shows the displacement–distance graphs for two sinusoidal waves (blue and green) of different wavelengths. It also shows the resultant wave (red), which comes from combining these two. How do we find this resultant displacement shown in red?

Consider position A. Here the displacement of both waves is zero, and so the resultant must also be zero.

At position B, both waves have positive displacement. The resultant displacement is found by adding these together.

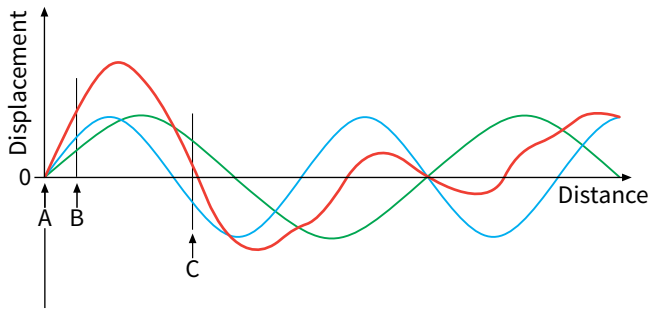


Figure 14.3 Adding two waves by the principle of superposition – the red line is the resultant wave.

At position C, the displacement of one wave is positive while the other is negative. The resultant displacement lies between the two displacements. In fact, the resultant displacement is the algebraic sum of the displacements of waves A and B; that is, their sum, taking account of their signs (positive or negative).

We can work our way along the distance axis in this way, calculating the resultant of the two waves by algebraically adding them up at intervals. Notice that, for these two waves, the resultant wave is a rather complex wave with dips and bumps along its length.

The idea that we can find the resultant of two waves which meet at a point simply by adding up the displacements at each point is called the **principle of superposition** of waves. This principle can be applied to more than two waves and also to all types of waves. A statement of the principle of superposition is shown below:

> When two or more waves meet at a point, the resultant displacement is the algebraic sum of the displacements of the individual waves.

QUESTION

1 On graph paper, draw two 'triangular' waves like those shown in Figure 14.4. (These are easier to work with than sinusoidal waves.) One should have wavelength 8 cm and amplitude 2 cm; the other wavelength 16 cm and amplitude 3 cm. Use the principle of superposition of waves to determine the resultant displacement at suitable points along the waves, and draw the complete resultant wave.

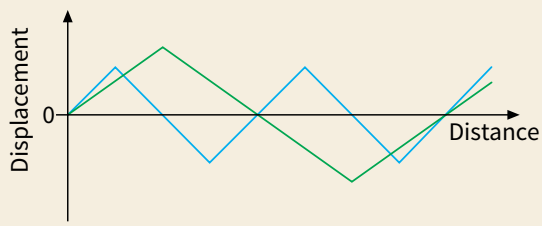


Figure 14.4 Two triangular waves – for Question 1.

Diffraction of waves

You should be aware that all waves (such as sound and light) can be reflected and refracted. Another wave phenomenon that applies to all waves is that they can be diffracted. **Diffraction** is the spreading of a wave as it passes through a gap or around an edge. It is easy to observe and investigate diffraction effects using water waves, as shown in Box 14.1.

Diffraction of sound and light

Diffraction effects are greatest when waves pass through a gap with a width roughly equal to their wavelength. This is useful in explaining why we can observe diffraction readily for some waves, but not for others. For example, sound waves in the audible range have wavelengths from a few millimetres to a few metres. Thus we might expect to observe diffraction effects for sound in our environment. Sounds, for example, diffract as they pass through doorways. The width of a doorway is comparable to the wavelength of a sound and so a noise in one room spreads out into the next room.

Visible light has much shorter wavelengths (about 5×10^{-7} m). It is not diffracted noticeably by doorways because the width of the gap is a million times larger than the wavelength of light. However, we can observe diffraction of light by passing it through a very narrow slit or a small hole. When laser light is directed onto a slit whose width is comparable to the wavelength of the incident light, it spreads out into the space beyond to form a smear on the screen (Figure 14.5). An adjustable slit allows you to see the effect of gradually narrowing the gap.

You can see the effects of diffraction for yourself by making a narrow slit with your two thumbs and looking through the slit at a distant light source (Figure 14.8). By gently pressing your thumbs together to narrow the gap between them, you can see the effect of narrowing the slit.

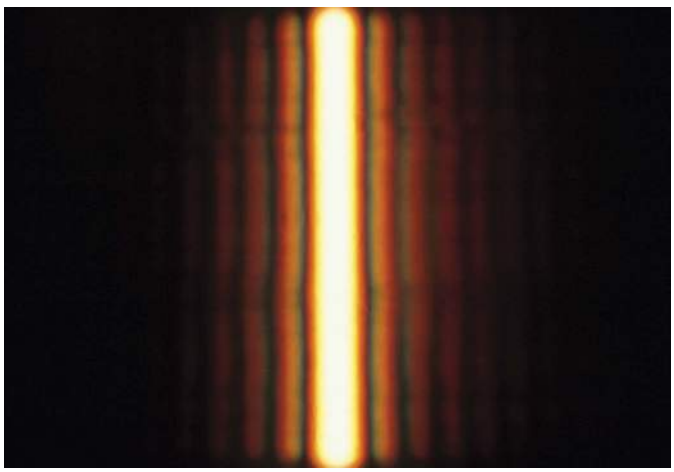


Figure 14.5 Light is diffracted as it passes through a slit.

BOX 14.1: Observing diffraction in a ripple tank

A ripple tank can be used to show diffraction. Plane waves are generated using a vibrating bar, and **move towards** a gap in a barrier (Figure 14.6). Where the ripples strike the barrier, they are reflected back. Where they arrive at the gap, however, they pass through and spread out into the space beyond. It is this spreading out of waves as they travel through a gap (or past the edge of a barrier) that is called diffraction.

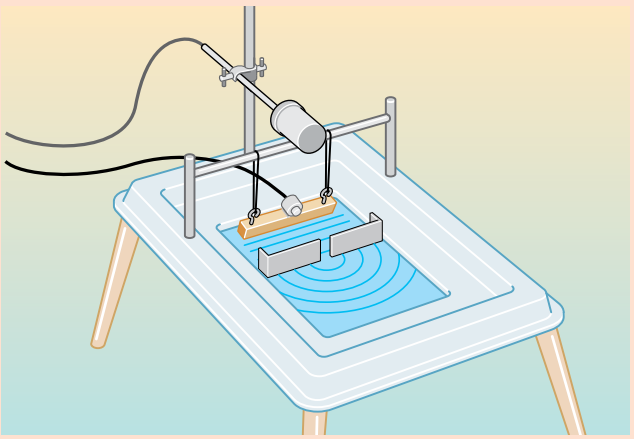


Figure 14.6 Ripples, initially straight, spread out into the space beyond the gap in the barrier.

The extent to which ripples are diffracted depends on the width of the gap. This is illustrated in Figure 14.6. The lines in this diagram show the wavefronts. It is as if we are looking down on the ripples from above, and drawing lines to represent the tops of the ripples at some instant in time. The separation between adjacent wavefronts is equal to the wavelength λ of the ripples.

When the waves encounter a gap in a barrier, the amount of diffraction depends on the width of the gap. There is hardly any noticeable diffraction when the gap is very much larger than the wavelength. As the gap becomes narrower, the diffraction effect becomes more pronounced. It is greatest when the width of the gap is roughly equal to the wavelength of the ripples.

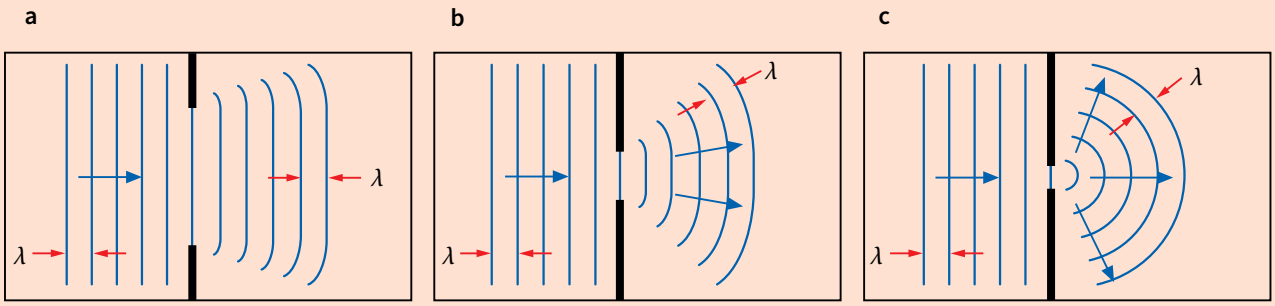


Figure 14.7 The extent to which ripples spread out depends on the relationship between their wavelength and the width of the gap. In **a**, the width of the gap is very much greater than the wavelength and there is hardly any noticeable diffraction. In **b**, the width of the gap is greater than the wavelength and there is limited diffraction. In **c**, the gap width is approximately equal to the wavelength and the diffraction effect is greatest.

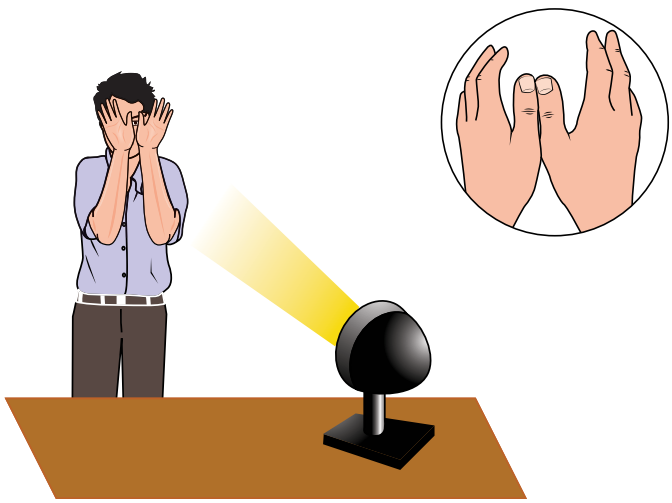


Figure 14.8 You can see the effects of diffraction by looking through a narrow slit. What happens when you make the slit narrower? What happens to the amount of diffraction when you put different coloured filters in front of the lamp? What does this tell you about the wavelengths of the different colours?

Diffraction of radio and microwaves

Radio waves can have wavelengths of the order of a kilometre. These waves are easily diffracted by gaps in the hills and by the tall buildings around our towns and cities. Microwaves, used by the mobile phone network, have wavelengths of about 10 cm. These waves are not easily diffracted (because their wavelengths are much smaller than the dimensions of the gaps) and mostly travel through space in straight lines.

Cars need external radio aerials because radio waves have wavelengths longer than the size of the windows, so they cannot diffract into the car. If you try listening to a radio in a train without an external aerial, you will find that FM signals can be picked up weakly (their wavelength is about 3 m), but AM signals, with longer wavelengths, cannot get in at all.

QUESTION

2 A microwave oven (Figure 14.9) uses microwaves with a wavelength of 12.5 cm. The front door of the oven is made of glass with a metal grid inside; the gaps in the grid are a few millimetres across. Explain how this design allows us to see the food inside the oven, while the microwaves are not allowed to escape into the kitchen (where they might cook us).



Figure 14.9 A microwave oven has a metal grid in the door to keep microwaves in and let light out.

Explaining diffraction

Diffraction is a wave effect that can be explained by the principle of superposition. We have to think about what happens when a plane ripple reaches a gap in a barrier (Figure 14.10). Each point on the surface of the water in the gap is moving up and down. Each of these moving points can be thought of as a source of new ripples spreading out into the space beyond the barrier. Now we have a lot of new ripples, and we can use the principle of superposition to find their resultant effect. Without trying to calculate the effect of an infinite number of ripples, we can say that in some directions the ripples add together while in other directions they cancel out.

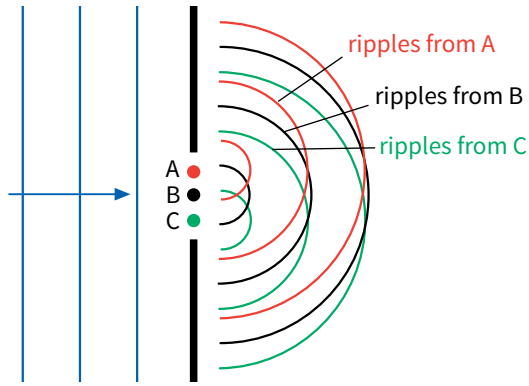


Figure 14.10 Ripples from all points across the gap contribute to the pattern in the space beyond.

Interference

Adding waves of different wavelengths and amplitudes results in complex waves. We can find some interesting effects if we consider what happens when two waves of the same wavelength overlap at a point. Again, we will use the principle of superposition to explain what we observe.

BOX 14.2: Observing interference

Interference of sound waves

A simple experiment shows what happens when two sets of sound waves meet. Two loudspeakers are connected to a single signal generator (Figure 14.11). They each produce sound waves of the same wavelength. Walk around in the space in front of the loudspeakers; you will hear the resultant effect.

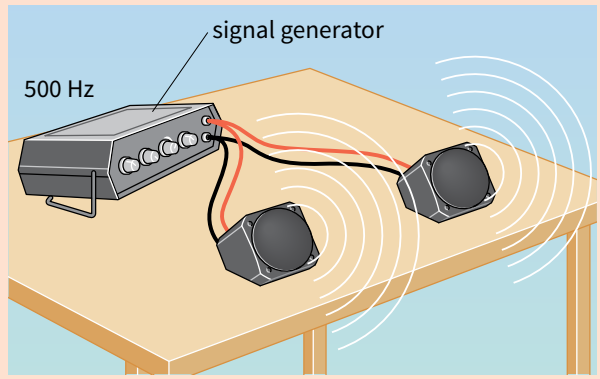


Figure 14.11 The sound waves from two loudspeakers combine to give an interference pattern. This experiment is best done outside so that reflections of sounds (or echoes) do not affect the results.

BOX 14.2: Observing interference (continued)

A naive view might be that we would hear a sound twice as loud as that from a single loudspeaker. However, this is not what we hear. At some points, the sound is **louder** than for a single speaker. At other points, the sound is much **quieter**. The space around the two loudspeakers consists of a series of loud and quiet regions. We are observing the phenomenon known as **interference**.

Interference in a ripple tank

The two dippers in the ripple tank (Figure 14.12) should be positioned so that they are just touching the surface of the water. When the bar vibrates, each dipper acts as a source of circular ripples spreading outwards. Where these sets of ripples overlap, we observe an interference pattern. Another way to observe interference in a ripple tank is to use plane waves passing through two gaps in a barrier. The water waves are diffracted at the two gaps and then interfere beyond the gaps. Figure 14.13 shows the interference pattern produced by two vibrating sources in a ripple tank.

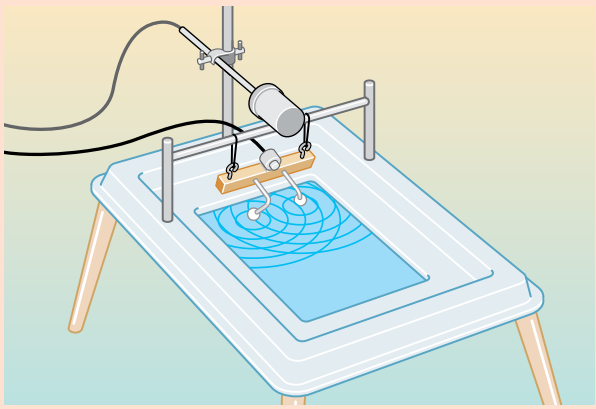


Figure 14.12 A ripple tank can be used to show how two sets of circular ripples combine.



Figure 14.13 Ripples from two point sources produce an interference pattern.

Explaining interference

Figure 14.14 shows how interference arises. The loudspeakers in Figure 14.11 (Box 14.2) are emitting waves that are in phase because both are connected to the same signal generator. At each point in front of the loudspeakers, waves are arriving from the two loudspeakers. At some points, the two waves arrive in phase (in step) with one another and with equal amplitude (Figure 14.14a). The principle of superposition predicts that the resultant wave has twice the amplitude of a single wave. We hear a louder sound.

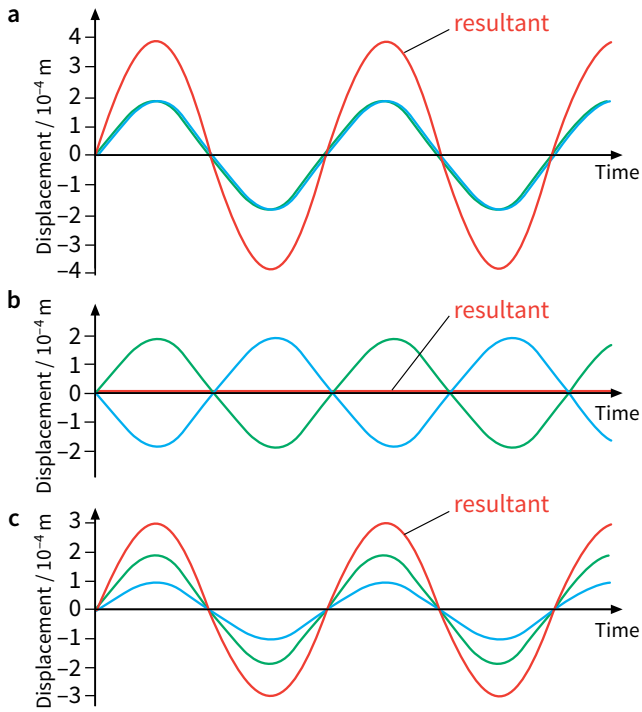


Figure 14.14 Adding waves by the principle of superposition. Blue and green waves of the same amplitude may give **a** constructive or **b** destructive interference, according to the phase difference between them. **c** Waves of different amplitudes can also interfere constructively.

At other points, something different happens. The two waves arrive completely out of phase or in antiphase (phase difference is 180°) with one another (Figure 14.14b). There is a cancelling out, and the resultant wave has zero amplitude. At this point, we would expect silence. At other points again, the waves are neither perfectly out of step nor perfectly in step, and the resultant wave has amplitude less than that at the loudest point.

Where two waves arrive at a point in phase with one another so that they add up, we call this effect **constructive interference**. Where they cancel out, the effect is known as **destructive interference**. Where two waves have different

amplitudes but are in phase (Figure 14.14c), constructive interference results in a wave whose amplitude is the sum of the two individual amplitudes.

QUESTION

3 Explain why the two loudspeakers must produce sounds of precisely the same frequency if we are to hear the effects of interference.

How can we explain the interference pattern observed in a ripple tank (Box 14.2)? Look at Figure 14.15 and compare it to Figure 14.13. Figure 14.15 shows two sets of waves setting out from their sources. At a position such as A, ripples from the two sources arrive in phase with one another, and constructive interference occurs. At B, the two sets of ripples arrive out of phase, and there is destructive interference. Although waves are arriving at B, the surface of the water remains approximately flat.

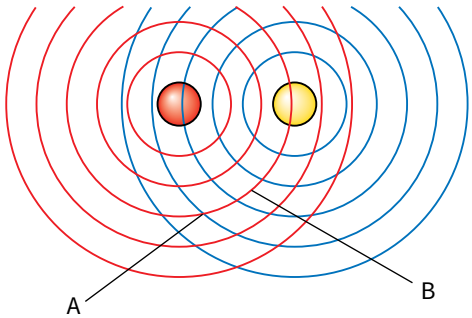


Figure 14.15 The result of interference depends on the path difference between the two waves.

Whether the waves combine constructively or destructively at a point depends on the path difference of the waves from the two sources. The path difference is defined as the extra distance travelled by one of the waves compared with the other.

At point A in Figure 14.15, the waves from the red source have travelled 3 whole wavelengths. The waves from the yellow source have travelled 4 whole wavelengths. The path difference between the two sets of waves is 1 wavelength. A path difference of 1 wavelength is equivalent to a phase difference of zero. This means that they are in phase, so they interfere constructively.

Now think about destructive interference. At point B, the waves from the red source have travelled 3 wavelengths; the waves from the yellow source have travelled 2.5 wavelengths. The path difference between the two sets of waves is 0.5 wavelengths, which is equivalent to a phase difference of 180°. The waves interfere destructively because they are in antiphase. In general, the conditions for constructive interference and destructive interference are outlined below. These conditions apply to **all** waves (water waves, light, microwaves, radio waves, sound, etc.) that show interference effects. In the equations below, n stands for any integer (any whole number, including zero).

- For **constructive interference** the path difference is a whole number of wavelengths:
  path difference = 0, λ, 2λ, 3λ, etc.
  or path difference = $n\lambda$
- For **destructive interference** the path difference is an odd number of half wavelengths:
  path difference = $\frac{1}{2}\lambda$, $1\frac{1}{2}\lambda$, $2\frac{1}{2}\lambda$, etc.
  or path difference = $(n + \frac{1}{2})\lambda$

BOX 14.3: Interference of radiation

Interference of light

Here is one way to show the interference effects produced by light. A simple arrangement involves directing the light from a laser through two slits (Figure 14.16). The slits are two clear lines on a black slide, separated by a fraction of a millimetre. Where the light falls on the screen, a series of equally spaced dots of light are seen (see Figure 14.21 on page 201). These bright dots are referred to as interference 'fringes', and they are regions where

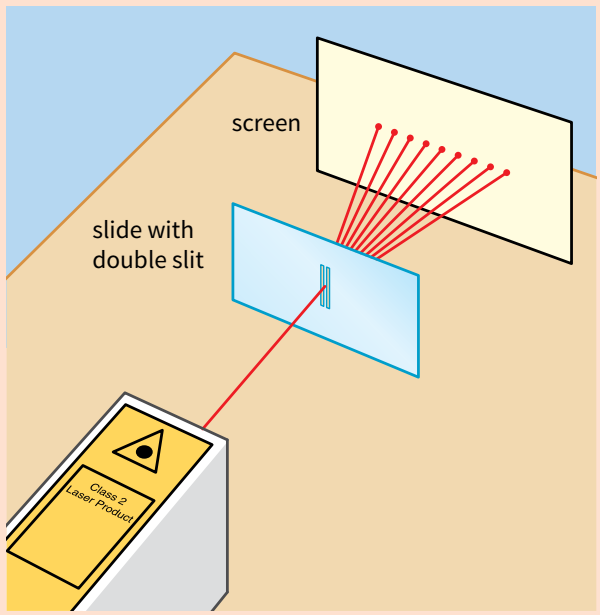


Figure 14.16 Light beams from the two slits interfere in the space beyond.

BOX 14.3: Interference of radiation (continued)

light waves from the two slits are arriving in phase with each other, i.e. there is constructive interference. The dark regions in between are the result of destructive interference.

Safety note

If you carry out experiments using a laser, you should **follow correct safety procedures**. In particular, you should **wear eye protection** and **avoid allowing the beam to enter your eye directly**.

These bright and dark fringes are the equivalent of the loud and quiet regions that you detected if you investigated the interference pattern of sounds from the two loudspeakers described in Box 14.2. Bright fringes correspond to loud sound, dark fringes to quiet sound or silence.

You can check that light is indeed reaching the screen from both slits as follows. Mark a point on the screen where there is a dark fringe. Now carefully cover up one of the slits so that light from the laser is only passing through one slit. You should find that the pattern of interference fringes disappears. Instead, a broad band of light appears across the screen. This broad band of light is the diffraction pattern produced by a single slit. The point that was dark is now light. Cover up the other slit instead, and you will see the same effect. You have now shown that light is arriving at the screen from both slits, but at some points (the dark fringes) the two beams of light cancel each other out.

You can achieve similar results with a bright light bulb rather than a laser, but a laser is much more convenient because the light is concentrated into a narrow, more intense beam. This famous experiment is called the Young double-slit experiment (see page 200), although Thomas Young had no laser available to him when he first carried it out in 1801.

Interference of microwaves

Using 2.8 cm wavelength microwave equipment (Figure 14.17), you can observe an interference pattern. The microwave transmitter is directed towards the double gap in a metal barrier. The microwaves are diffracted at the two gaps so that they spread out into the region beyond, where they can be detected using the probe receiver. By moving the probe around, it is possible to detect regions of high intensity (constructive interference) and low intensity (destructive interference). The probe may be connected to a meter, or to an audio amplifier and loudspeaker to give an audible output.

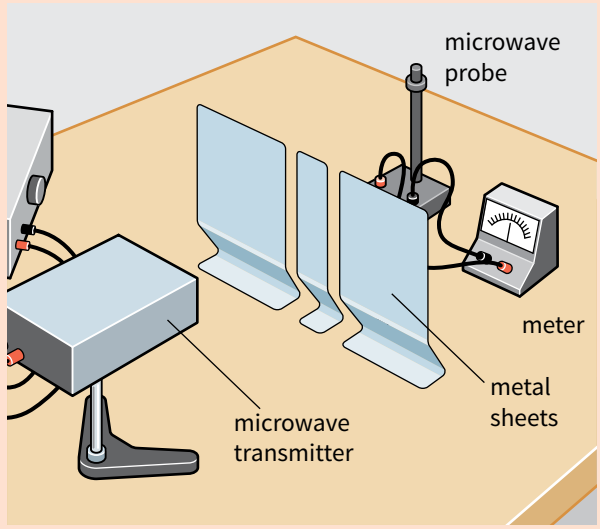


Figure 14.17 Microwaves can also be used to show interference effects.

QUESTION

4 Look at the experimental arrangement shown in Figure 14.17. Suppose that the microwave probe is placed at a point of low intensity in the interference pattern. What do you predict will happen if one of the gaps in the barrier is now blocked?

Coherence

We are surrounded by many types of wave – light, infrared radiation, radio waves, sound, and so on. There are waves coming at us from all directions. So why do we not observe interference patterns all the time? Why do we need specialised equipment in a laboratory to observe these effects?

In fact, we can see interference of light occurring in everyday life. For example, you may have noticed haloes of light around street lamps or the Moon on a foggy night.

You may have noticed light and dark bands of light if you look through fabric at a bright source of light. These are interference effects.

We usually need specially arranged conditions to produce interference effects that we can measure. Think about the demonstration with two loudspeakers. If they were connected to different signal generators with slightly different frequencies, the sound waves might start off in phase with one another, but they would soon go out of phase (Figure 14.18). We would hear loud, then soft, then loud again. The interference pattern would keep shifting around the room.

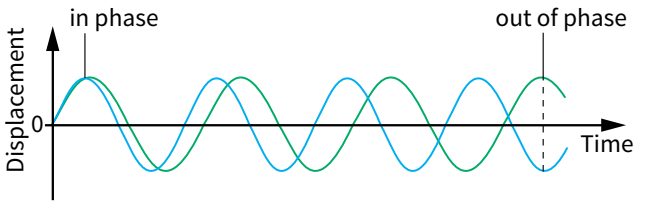


Figure 14.18 Waves of slightly different wavelengths (and therefore frequencies) move in and out of phase with one another.

By connecting the two loudspeakers to the **same** signal generator, we can be sure that the sound waves that they produce are constantly in phase with one another. We say that they act as two **coherent** sources of sound waves (coherent means sticking together). Coherent sources emit waves that have a **constant phase difference**. Note that the two waves can only have a constant phase difference if their frequency is the same and remains constant.

Now think about the laser experiment. Could we have used two lasers producing exactly the same frequency and hence wavelength of light? Figure 14.19a represents the light from a laser. We can think of it as being made up of many separate bursts of light. We cannot guarantee that these bursts from two lasers will always be in phase with one another.

This problem is overcome by using a single laser and dividing its light using the two slits (Figure 14.19b). The slits act as two coherent sources of light. They are constantly in phase with one another (or there is a constant phase difference between them).

If they were not coherent sources, the interference pattern would be constantly changing, far too fast for our eyes to detect. We would simply see a uniform band of light, without any definite bright and dark regions. From this you should be able to see that, in order to observe interference, we need two coherent sources of waves.

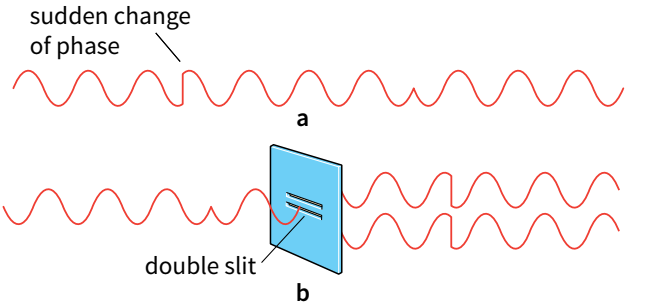


Figure 14.19 Waves must be coherent if they are to produce a clear interference pattern.

QUESTION

5 Draw sketches of displacement against time to illustrate the following:

- **a** two waves having the same amplitude and in phase with one another
- **b** two waves having the same amplitude and with a phase difference of 90°
- **c** two waves initially in phase but with slightly different wavelengths.

Use your sketches to explain why two coherent sources of waves are needed to observe interference.

The Young double-slit experiment

Now we will take a close look at a famous experiment which Thomas Young performed in 1801. He used this experiment to show the wave nature of light. A beam of light is shone on a pair of parallel slits placed at right angles to the beam. Light diffracts and spreads outwards from each slit into the space beyond; the light from the two slits overlaps on a screen. An interference pattern of light and dark bands called 'fringes' is formed on the screen.

Explaining the experiment

In order to observe interference, we need two sets of waves. The sources of the waves must be coherent – the phase difference between the waves emitted at the sources must remain constant. This also means that the waves must have the same wavelength. Today, this is readily achieved by passing a single beam of laser light through the two slits. A laser produces intense coherent light. As the light passes through the slits, it is diffracted so that it spreads out into the space beyond (Figure 14.20). Now we have two overlapping

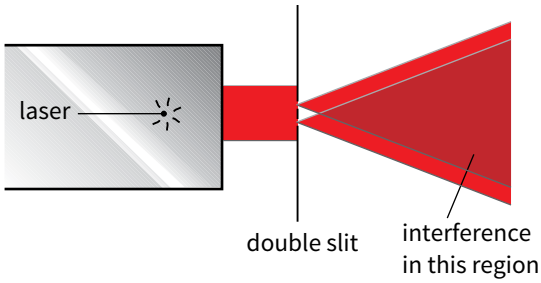


Figure 14.20 Interference occurs where diffracted beams from the two slits overlap.

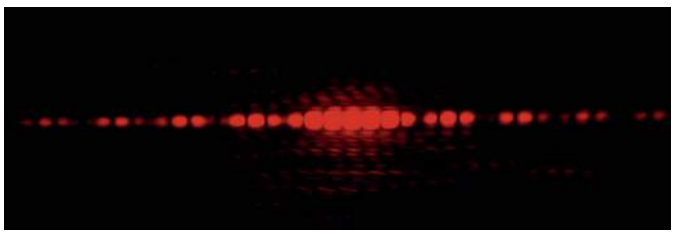


Figure 14.21 Interference fringes obtained using a laser and a double slit.

sets of waves, and the pattern of fringes on the screen shows us the result of their interference (Figure 14.21).

How does this pattern arise? We will consider three points on the screen (Figure 14.22), and work out what we would expect to observe at each.

Point A

This point is directly opposite the midpoint of the slits. Two rays of light arrive at A, one from slit 1 and the other from slit 2. Point A is equidistant from the two slits, and so the two rays of light have travelled the same distance. The path difference between the two rays of light is zero. If we assume that they were in phase (in step) with each other when they left the slits, then they will be in phase when they arrive at A. Hence they will interfere constructively, and we will observe a bright fringe at A.

Point B

This point is slightly to the side of point A, and is the midpoint of the first dark fringe. Again, two rays of light arrive at B, one from each slit. The light from slit 1 has to travel slightly further than the light from slit 2, and so the two rays are no longer in step. Since point B is at the midpoint of the dark fringe, the two rays must be in antiphase (phase difference of 180°). The path difference between the two rays of light must be half a wavelength and so the two rays interfere destructively.

Point C

This point is the midpoint of the next bright fringe, with AB = BC. Again, ray 1 has travelled further than ray 2; this

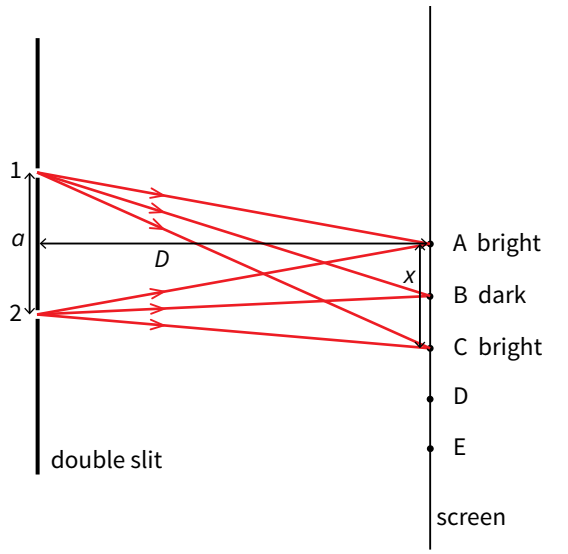


Figure 14.22 Rays from the two slits travel different distances to reach the screen.

time, it has travelled an extra distance equal to a whole wavelength λ. The path difference between the rays of light is now a whole wavelength. The two rays are in phase at the screen. They interfere constructively and we see a bright fringe.

The complete interference pattern (Figure 14.21) can be explained in this way.

QUESTION

6 Consider points D and E on the screen, where BC = CD = DE. State and explain what you would expect to observe at D and E.

Determining wavelength λ

The double-slit experiment can be used to determine the wavelength λ of light. The following three quantities have to be measured:

- **Slit separation *a*** – This is the distance between the centres of the slits, which is the distance between slits 1 and 2 in Figure 14.22.
- **Fringe separation *x*** – This is the distance between the centres of adjacent bright (or dark) fringes, which is the distance AC in Figure 14.22.
- **Slit-to-screen distance *D*** – This is the distance from the midpoint of the slits to the central fringe on the screen.

Once these three quantities have been measured, the wavelength λ of the light can be found using the relationship:

$$\lambda = \frac{ax}{D}$$

WORKED EXAMPLE

1 In a double-slit experiment using light from a helium–neon laser, a student obtained the following results:

width of 10 fringes $10x = 1.5\,\text{cm}$

separation of slits $a = 1.0\,\text{mm}$

slit-to-screen distance $D = 2.40\,\text{m}$

Determine the wavelength of the light.

Step 1 Work out the fringe separation:

$$\text{fringe separation } x = \frac{1.5 \times 10^{-2}}{10} = 1.5 \times 10^{-3}\,\text{m}$$

Step 2 Substitute the values of a, x and D in the expression for wavelength λ:

$$\lambda = \frac{ax}{D}$$

Therefore:

$$\lambda = \frac{1.0 \times 10^{-3} \times 1.5 \times 10^{-3}}{2.40} = 6.3 \times 10^{-7}\,\text{m}$$

Hint: *Don't forget to convert all the distances into metres.*

So the wavelength is $6.3 \times 10^{-7}\,\text{m}$ or 630 nm.

QUESTION

7 If the student in Worked example **1** moved the screen to a distance of 4.8 m from the slits, what would the fringe separation become?

BOX 14.4: Using Young's slits to measure λ

The Young double-slit experiment can be used to determine the wavelength of light λ. Here we look at a number of practical features of the experiment and consider how the uncertainty in the value of λ can be reduced.

One way to carry out the double-slit experiment is shown in Figure 14.23. Here, a white light source is used, rather than a laser. A monochromatic filter allows only one wavelength of light to pass through. A single slit diffracts the light. This diffracted light arrives in phase at the double slit, which ensures that the two parts of the double slit behave as coherent sources of light. The double slit is placed a centimetre or two beyond the single slit, and the fringes are observed on a screen a metre or so away. The experiment has to be carried out in a darkened room, as the intensity of the light is low and the fringes are hard to see.

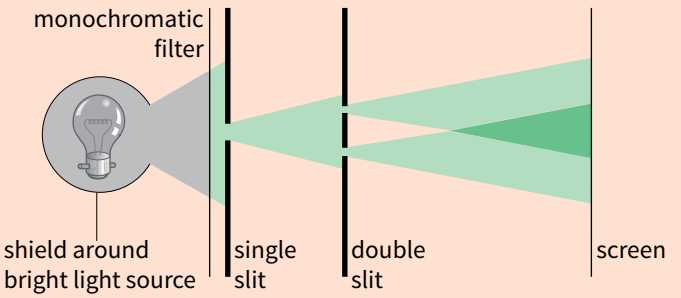


Figure 14.23 To observe interference fringes with white light, you must use a single slit before the double slit.

There are three important factors involved in the way the equipment is set up:

- All slits are a fraction of a millimetre in width. Since the wavelength of light is less than a micrometre ($10^{-6}\,\text{m}$), this gives a small amount of diffraction in the space beyond. If the slits were narrower, the intensity of the light would be too low for visible fringes to be achieved.
- The double slits are about a millimetre apart. If they were much further apart, the fringes would be too close together to be distinguishable.
- The screen is about a metre from the slits. The fringes produced are clearly separated without being too dim.

Measuring *a*, *x* and *D*

Measuring slit separation a: a travelling microscope is suitable for measuring a. It is difficult to judge the position of the centre of a slit. If the slits are the same width, the separation of their left-hand edges is the same as the separation of their centres.

BOX 14.4: Using Young's slits to measure λ (continued)

Measuring fringe width x: it is best to measure across several fringes (say, ten) and then to calculate the average separation later. A metre rule or travelling microscope can be used.

Measuring the slit-to-screen distance D: this can be measured using a metre rule or a tape measure.

Reducing percentage errors

Why use a laser rather than white light? With a laser, the light beam is more concentrated, and the initial single slit is not necessary. The greater intensity of the beam means that the screen can be further from the slits, so that the fringes are further apart; this reduces the percentage error in measurements of x and D, and hence λ can be determined more accurately.

A laser has a second advantage. The light from a laser is monochromatic; that is, it consists of a single wavelength. This makes the fringes very clear, and they are present in large numbers across the screen. With white light, a range of wavelengths is present. Different wavelengths form fringes at different points across the screen, smearing them out so that they are not as clear.

Using white light with no filter results in a central fringe which is white (because all wavelengths are in phase here), but the other fringes show coloured effects, as the different wavelengths interfere constructively at different points. In addition, only a few fringes are visible in the interference pattern.

QUESTIONS

8 Use $\lambda = \frac{ax}{D}$ to explain the following observations:

- **a** With the slits closer together, the fringes are further apart.
- **b** Interference fringes for blue light are closer together than for red light.
- **c** In an experiment to measure the wavelength of light, it is desirable to have the screen as far from the slits as possible.

9 Yellow sodium light of wavelength 589 nm is used in the Young double-slit experiment. The slit separation is 0.20 mm, and the screen is placed 1.20 m from the slits. Calculate the separation of neighbouring fringes formed on the screen.

10 In a double-slit experiment, filters were placed in front of a white light source to investigate the effect of changing the wavelength of the light. At first, a red filter was used instead (λ = 600 nm) and the fringe separation was found to be 2.40 mm. A blue filter was then used instead (λ = 450 nm). Determine the fringe separation with the blue filter.

Diffraction gratings

A transmission diffraction grating is similar to the slide used in the double-slit experiment, but with many more slits than just two. It consists of a large number of equally spaced lines ruled on a glass or plastic slide. Each line is capable of diffracting the incident light. There may be as many as 10 000 lines per centimetre. When light is shone through this grating, a pattern of interference fringes is seen.

A reflection diffraction grating consists of lines made on a reflecting surface so that light is both reflected and diffracted by the grating. The shiny surface of a compact disc (CD) or DVD is an everyday example of a reflection diffraction grating. Hold a CD in your hand so that you are looking at the reflection of light from a lamp. You will observe coloured bands (Figure 14.24). A CD has



Figure 14.24 A CD acts as a reflection diffraction grating. White light is reflected and diffracted at its surface, producing a display of spectral colours.

thousands of equally spaced lines of microscopic pits on its surface; these carry the digital information. It is the diffraction from these lines that produces the coloured bands of light from the surface of the CD.

Observing diffraction with a transmission grating

In Figure 14.25, monochromatic light from a laser is incident normally on a transmission diffraction grating. In the space beyond, interference fringes are formed. These can be observed on a screen, as with the double slit. However, it is usual to measure the angle θ at which they are formed, rather than measuring their separation. With double slits, the fringes are equally spaced and the angles are very small. With a diffraction grating, the angles are much greater and the fringes are not equally spaced.

The fringes are also referred to as **maxima**. The central fringe is called the zeroth-order maximum, the next fringe is the first-order maximum, and so on. The pattern is symmetrical, so there are two first-order maxima, two second-order maxima, and so on.

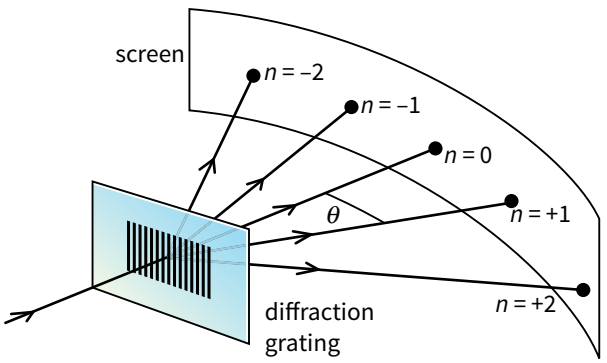


Figure 14.25 The diffracted beams form a symmetrical pattern on either side of the undiffracted central beam.

Explaining the experiment

The principle is the same as for the double-slit experiment, but here we have light passing through many slits. As it passes through each slit, it diffracts into the space beyond. So now we have many overlapping beams of light, and these interfere with one another. It is difficult to achieve constructive interference with many beams, because they all have to be in phase with one another.

There is a bright fringe, the zeroth-order maximum, in the straight-through direction ($\theta = 0$) because all of the rays here are travelling parallel to one another and in phase, so the interference is constructive (Figure 14.26a).

Imagine if you could look through the grating at the source of light. Your eye would be focused on the light

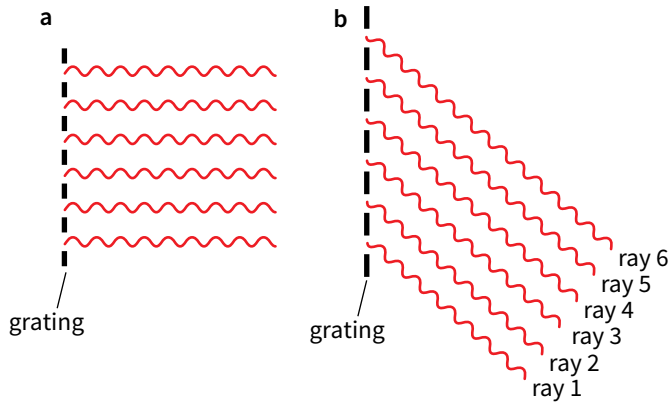


Figure 14.26 **a** Waves from each slit are in phase in the straight-through direction. **b** In the direction of the first-order maximum, the waves are in phase, but each one has travelled one wavelength further than the one below it.

source, which is far away. All the rays with $\theta = 0$ come together at the back of your eye, where an image is formed. It is here that interference occurs.

The first-order maximum forms as follows. Rays of light emerge from all of the slits; to form a bright fringe, all the rays must be in phase. In the direction of the first-order maximum, ray 1 has travelled the smallest distance (Figure 14.26b). Ray 2 has travelled an extra distance equal to one whole wavelength and is therefore in phase with ray 1. The path difference between ray 1 and ray 2 is equal to one wavelength λ. Ray 3 has travelled two extra wavelengths and is in phase with rays 1 and 2. In fact, the rays from all of the slits are in step in this direction, and a bright fringe results.

QUESTION

11 Explain how the second-order maximum arises. Use the term **path difference** in your explanation.

Determining wavelength λ with a grating

By measuring the angles at which the maxima occur, we can determine the wavelength of the incident light. The wavelength λ of the monochromatic light is related to the angle θ by:

$$d \sin\theta = n\lambda$$

where d is the distance between adjacent lines of the grating and n is known as the **order** of the maximum; n can only have integer values 0, 1, 2, 3, etc. The distance d is known as the grating element or grating spacing. This is illustrated in Worked example 2.

WORKED EXAMPLE

2 Monochromatic light is incident normally on a diffraction grating having 3000 lines per centimetre. The angular separation of the zeroth- and first-order maxima is found to be 10°. Calculate the wavelength of the incident light.

Step 1 Calculate the slit separation (grating spacing) d. Since there are 3000 slits per centimetre, their separation must be:

$$d = \frac{1\,\text{cm}}{3000} = 3.33 \times 10^{-4}\,\text{cm} = 3.33 \times 10^{-6}\,\text{m}$$

Step 2 Rearrange the equation $d\sin\theta = n\lambda$ and substitute values:

$\theta = 10.0°$, $n = 1$

$$\lambda = \frac{d\sin\theta}{n} = \frac{3.36 \times 10^{-6} \times \sin 10°}{1}$$

$$\lambda = 5.8 \times 10^{-7}\,\text{m} = 580\,\text{nm}$$

QUESTIONS

12 **a** For the case described in Worked example 2, at what angle would you expect to find the second-order maximum ($n = 2$)?

b Repeat the calculation of θ for $n = 3$, 4, etc. What is the limit to this calculation? How many maxima will there be altogether in this interference pattern?

13 Consider the equation $d\sin\theta = n\lambda$. How will the diffraction pattern change if:

a the wavelength of the light is increased?

b the diffraction grating is changed for one with more lines per centimetre (slits that are more closely spaced)?

QUESTION

14 A student is trying to make an accurate measurement of the wavelength of green light from a mercury lamp ($\lambda = 546\,\text{nm}$). Using a double slit of separation 0.50 mm, he finds he can see ten clear fringes on a screen at a distance of 0.80 m from the slits. The student can measure their overall width to within ±1 mm. He then tries an alternative experiment using a diffraction grating that has 3000 lines per centimetre. The angle between the two second-order maxima can be measured to within ±0.1°.

a What will be the width of the ten fringes that he can measure in the first experiment?

b What will be the angle of the second-order maximum in the second experiment?

c Suggest which experiment you think will give the more accurate measurement of λ.

Diffracting white light

A diffraction grating can be used to split white light up into its constituent colours (wavelengths). This splitting of light is known as **dispersion**, shown in Figure 14.27. A beam of white light is shone onto the grating. A zeroth-order, white maximum is observed at $\theta = 0°$, because all waves of each wavelength are in phase in this direction.

On either side, a series of spectra appear, with violet closest to the centre and red furthest away. We can see why different wavelengths have their maxima at different angles if we rearrange the equation $d\sin\theta = n\lambda$ to give:

$$\sin\theta = \frac{n\lambda}{d}$$

BOX 14.5: Diffraction gratings versus double slits

It is worth comparing the use of a diffraction grating to determine wavelength with the Young two-slit experiment.

- With a diffraction grating, the maxima are very **sharp**.
- With a diffraction grating, the maxima are also very **bright**. This is because rather than contributions from only two slits, there are contributions from a thousand or more slits.
- With two slits, there may be a large inaccuracy in the measurement of the slit separation a. The fringes are close together, so their separation may also be measured imprecisely.
- With a diffraction grating, there are many slits per centimetre, so d can be measured precisely. Because the maxima are widely separated, the angle θ can also be measured to a high degree of precision. So an experiment with a diffraction grating can be expected to give measurements of wavelength to a much higher degree of precision than a simple double-slit arrangement.

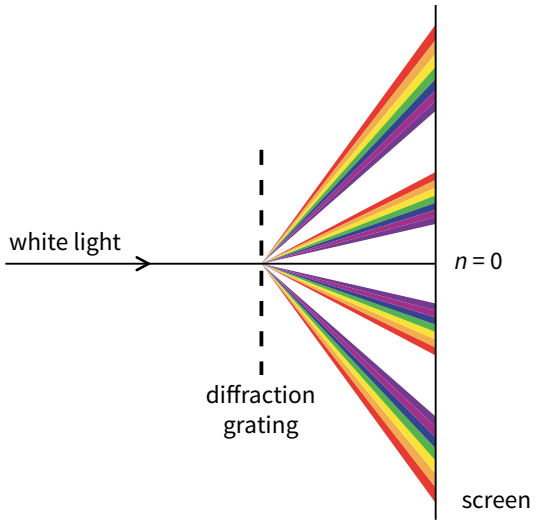


Figure 14.27 A diffraction grating is a simple way of separating white light into its constituent wavelengths.

From this it follows that the greater the wavelength λ, the greater the value of $\sin\theta$ and hence the greater the angle θ. Red light is at the long wavelength end of the visible spectrum, and so it appears at the greatest angle.

QUESTION

15 White light is incident normally on a diffraction grating with a slit separation d of 2.00×10^{-6} m.

a Calculate the angle **between** the red and violet ends of the first-order spectrum. The visible spectrum has wavelengths between 400 nm and 700 nm.

b Explain why the second- and third-order spectra overlap.

Summary

- The principle of superposition states that when two or more waves meet at a point, the resultant displacement is the algebraic sum of the displacements of the individual waves.
- When waves pass through a slit, they may be diffracted so that they spread out into the space beyond. The diffraction effect is greatest when the wavelength of the waves is similar to the width of the gap.
- Interference is the superposition of waves from two coherent sources.
- Two sources are coherent when they emit waves that have a **constant phase difference.** (This can only happen if the waves have the same frequency or wavelength.)
- For **constructive interference** the path difference is a whole number of wavelengths:

  path difference = $0, \lambda, 2\lambda, 3\lambda$, etc. or

  path difference = $n\lambda$
- For **destructive interference** the path difference is an odd number of half wavelengths:

  path difference = $\frac{1}{2}\lambda, 1\frac{1}{2}\lambda, 2\frac{1}{2}\lambda$, etc. or

  path difference = $(n + \frac{1}{2})\lambda$
- When light passes through a double slit, it is diffracted and an interference pattern of equally spaced light and dark fringes is observed. This can be used to determine the wavelength of light using the equation:

  $$\lambda = \frac{ax}{D}$$

  This equation can be used for all waves, including sound and microwaves.
- A diffraction grating diffracts light at its many slits or lines. The diffracted light interferes in the space beyond the grating. The equation for a diffraction grating is:

  $d \sin\theta = n\lambda$

End-of-chapter questions

1 **a** Copy the waves shown in Figure 14.28 onto a sheet of graph paper and use the principle of superposition to show the resultant wave. [2]

b How does the wavelength of the resulting wave compare with that of the component waves? [1]

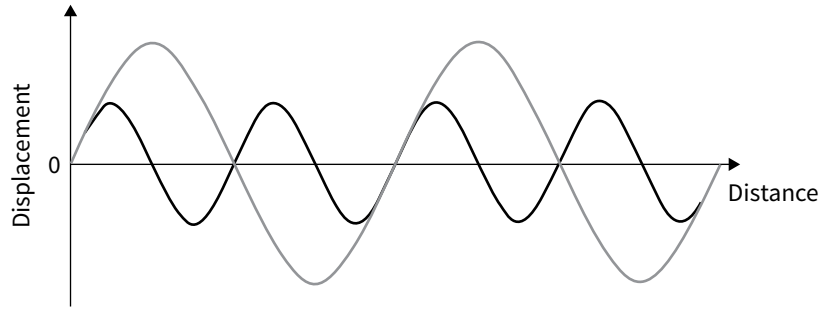


Figure 14.28 For End-of-chapter Question 1.

2 Figure 14.29 shows a ripple tank being used to demonstrate diffraction of water waves.

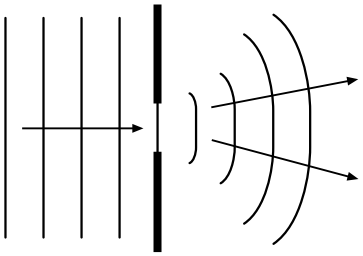


Figure 14.29 For End-of-chapter Question 2.

Suggest how the diffracted wave shape will change if:

a the wavelength of the incident wave is increased [1]

b the wavelength of the incident wave is decreased. [1]

3 Explain why, in remote mountainous regions, such as the Hindu Kush, radio signals from terrestrial transmitters can be received, but television reception can only be received from satellite transmissions. [2]

4 Damita and Jamal are organising a disco. Damita suggests that feeding the sound from the music centre to a second loudspeaker will increase the loudness of the music. Jamal says it won't work as there will be places where the sound will be very loud, due to constructive interference, and places where it will be much quieter, due to destructive interference. State who is correct and explain your reasoning. [2]

5 The constant frequency signal from a signal generator is fed to two loudspeakers placed 1.5 m apart. A girl, who is 8 metres away from the speakers, walks across in a line parallel to the line between the speakers. She finds that there is a distance of 1.2 m between successive spots where the sound is very quiet. Calculate the wavelength of the sound. [2]

6 Two signal generators feed signals with slightly different frequencies to two separate loudspeakers. Suggest why a sound of continuously rising and falling loudness is heard. [3]

7 A hydrogen discharge lamp produces a spectral line of wavelength 656 nm. Calculate the angles at which a diffraction grating, with 5000 lines cm^{-1}, would produce the first and second maxima for this light. [5]

8 **a** Explain what is meant by the term **superposition**. [2]

b In a Young double-slit experiment using yellow light of wavelength 590 nm from a sodium discharge tube, a student sets up a screen 1.8 m from the double slit. He measures the distance between 12 fringes as 16.8 mm. Calculate the separation of the slits. [3]

c State the effect of:

i using slits of narrower width (but the same separation) [2]

ii using slits with a smaller separation but of the same width [2]

9 **a** A laser light is described as producing light which is both highly coherent and highly monochromatic. Explain what is meant by the terms **coherent** and **monochromatic**. [2]

b Figure 14.30 shows the setup used to analyse the spectrum of a sodium discharge lamp using a diffraction grating with 5000 lines cm^{-1}. Figure 14.31 shows the spectral lines observed in the developed photographic film.

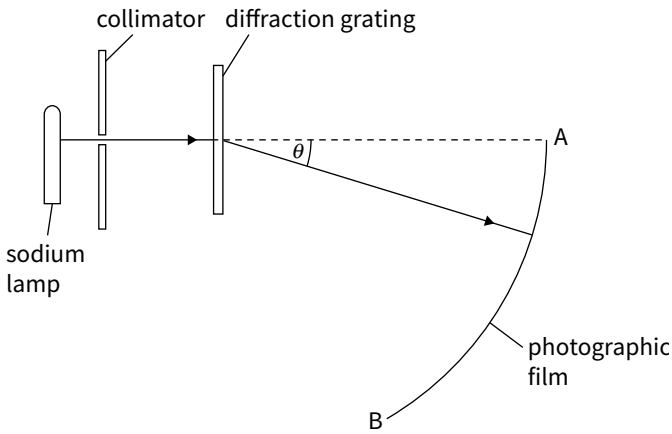


Figure 14.30 Experimental setup for End-of-chapter Question 9.

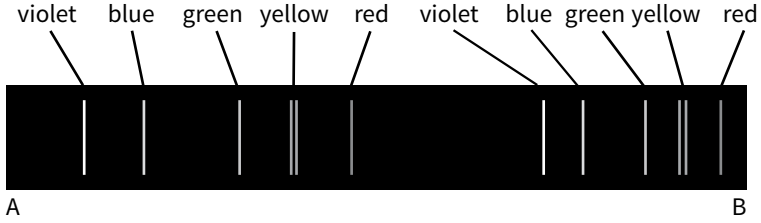


Figure 14.31 Spectral lines produced.

i Explain why two spectra are observed. [2]

ii Study the two spectra and describe **two** differences between them. [2]

iii The green maximum near end A is at an angle θ of 19.5°. Calculate the wavelength of the green light. [3]

iv Calculate the angle produced by the second green line. [2]

10 **a** Explain what is meant by the term **destructive interference**. [2]

b A student sets up an experiment to investigate the interference pattern formed by microwaves of wavelength 1.5 cm. The apparatus is set up as in Figure 14.17 on page 199. The distance between the centres of the two slits is 12.5 cm. The detector is centrally placed 1.2 m from the metal plates where it detects a maximum. The student moves the detector 450 cm across the bench parallel to the plates. Calculate how many maxima the detector will be moved through. [3]

c Calculate the frequency of these microwaves. [2]

11 **a** Explain what is meant by the **diffraction** of a wave. [2]

b Figure 14.32 shows waves, in a ripple tank, spreading out from two slits.

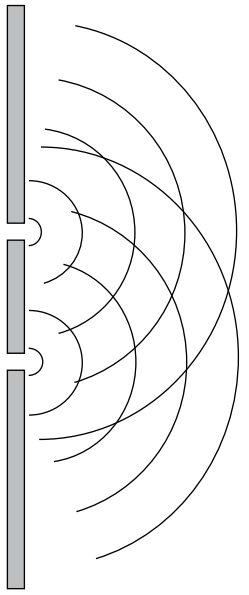


Figure 14.32 For End-of-chapter Question 11.

On a copy of Figure 14.32, draw:

i a line showing points along the central maximum – label this line **0** [1]

ii a line showing the points along first maximum – label this line **1** [1]

iii a line showing points along one of the first minima – label this line **min**. [1]

c The centres of the slits are 12 cm apart. At a distance of 60 cm from the barrier, the first maxima are 18 cm either side of the central maximum. Calculate the wavelength of the waves. You may assume that the formula developed for light waves is applicable in this example. [3]

Chapter 15: Stationary waves

Learning outcomes

You should be able to:

- demonstrate and explain the formation of stationary waves
- identify nodes and antinodes on a stationary wave
- determine the wavelength of sound waves using stationary waves

The bridge that broke

Figure 15.1a shows the Normandy Bridge under construction in France. When designing bridges, engineers must take into account the possibility of the wind causing a build-up of stationary waves, which may lead the bridge to oscillate violently. Famously, this happened in October 1940 to the Tacoma Narrows Bridge in Washington State, USA. High winds caused the bridge to vibrate with increasing amplitude until it fell apart (Figure 15.1b).

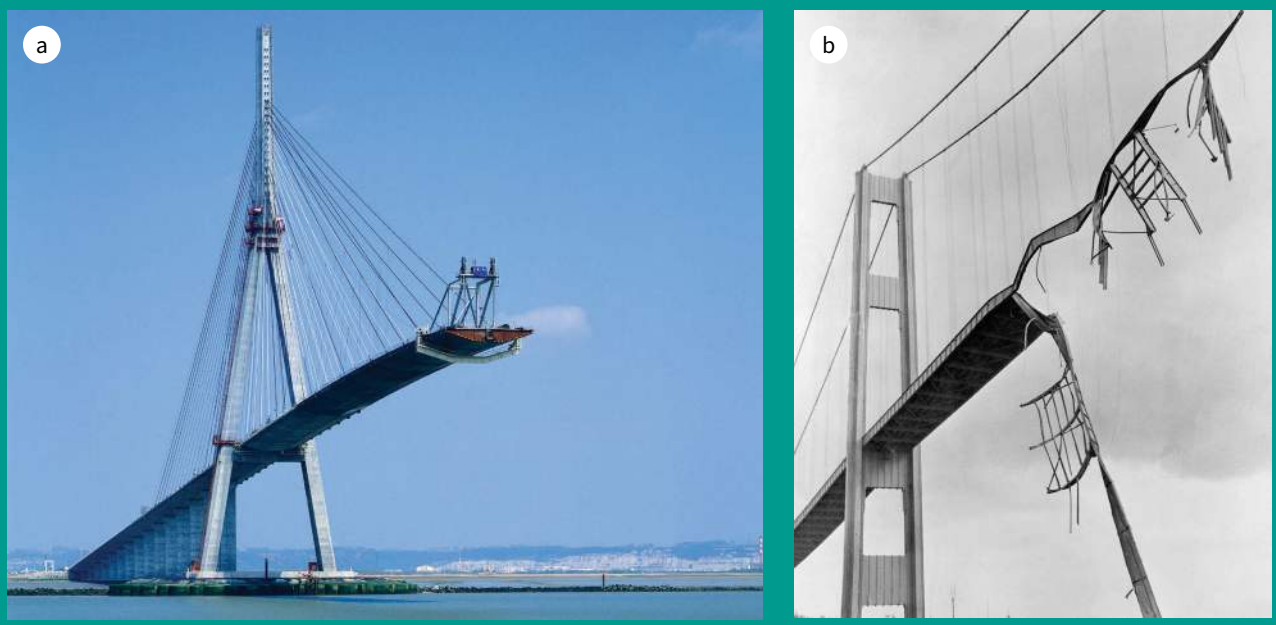


Figure 15.1 **a** A suspension bridge under construction. **b** One that failed – the Tacoma Narrows Bridge.

From moving to stationary

The waves we have considered so far in Chapters 13 and 14 have been **progressive waves**; they start from a source and travel outwards, transferring energy from one place to another. A second important class of waves is **stationary waves** (standing waves). These can be observed as follows. Use a long spring or a slinky spring. A long rope or piece of rubber tubing will also do. Lay it on the floor and fix one end firmly. Move the other end from side to side so that transverse waves travel along the length of the spring and reflect off the fixed end (Figure 15.2). If you adjust the frequency of the shaking, you should be able to achieve a stable pattern like one of those shown in Figure 15.3. Alter the frequency in order to achieve one of the other patterns.

You should notice that you have to move the end of the spring with just the right frequency to get one of these interesting patterns. The pattern disappears when the frequency of the shaking of the free end of the spring is slightly increased or decreased.

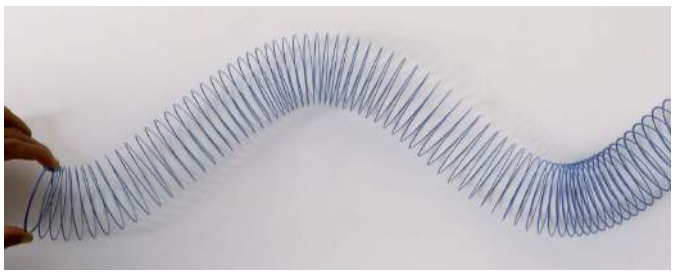


Figure 15.2 A slinky spring is used to generate a stationary wave pattern.

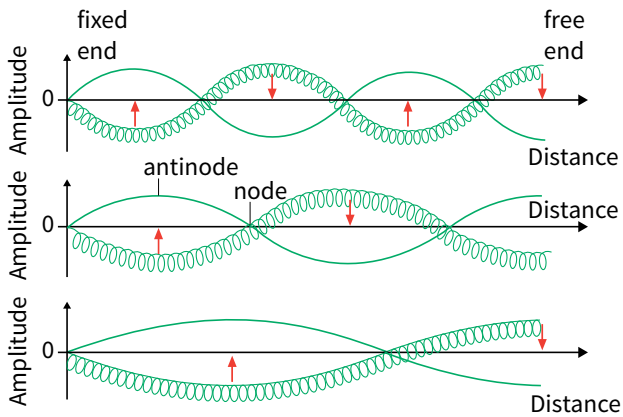


Figure 15.3 Different stationary wave patterns are possible, depending on the frequency of vibration.

Nodes and antinodes

What you have observed is a stationary wave on the long spring. There are points along the spring that remain (almost) motionless while points on either side are oscillating with the greatest amplitude. The points that do not move are called the **nodes** and the points where the spring oscillates with maximum amplitude are called the **antinodes**. At the same time, it is clear that the wave profile is not travelling along the length of the spring. Hence we call it a stationary wave or a standing wave.

We normally represent a stationary wave by drawing the shape of the spring in its two extreme positions (Figure 15.4). The spring appears as a series of loops, separated by nodes. In this diagram, point A is moving downwards. At the same time, point B in the next loop is moving upwards. The phase difference between points A and B is 180°. Hence the sections of spring in adjacent loops are always moving in antiphase; they are half a cycle out of phase with one another.

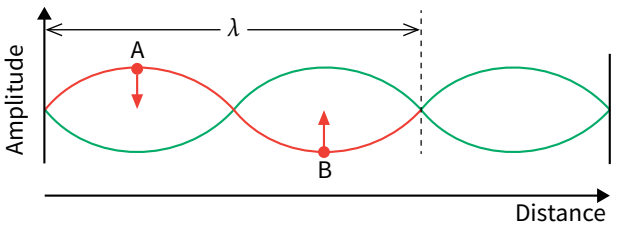


Figure 15.4 The fixed ends of a long spring must be nodes in the stationary wave pattern.

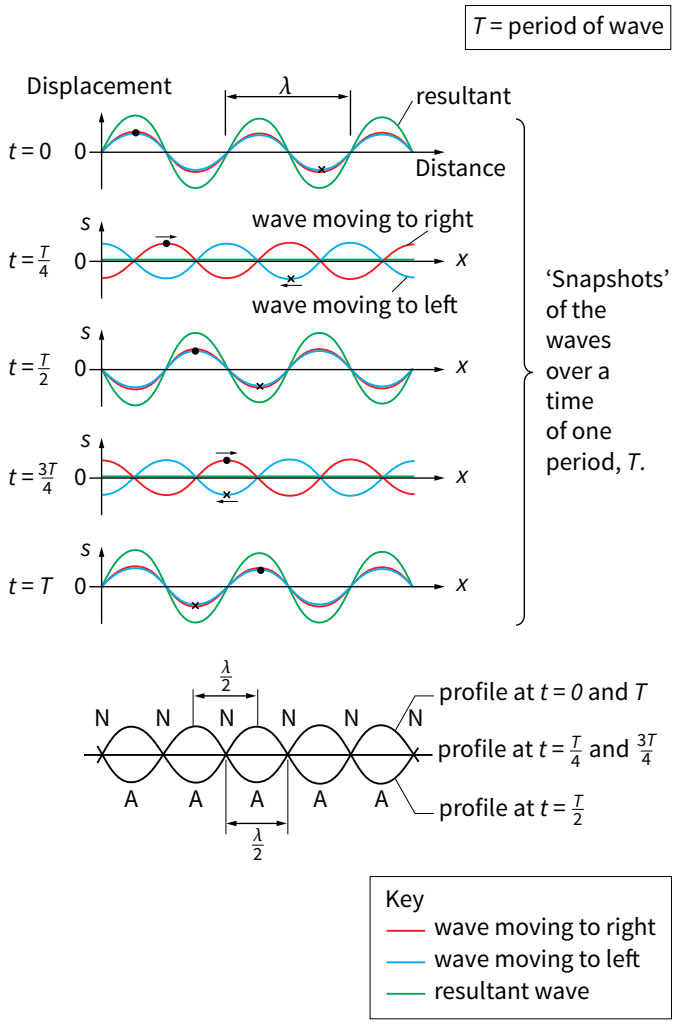


Figure 15.5 The blue-coloured wave is moving to the left and the red-coloured wave to the right. The **principle of superposition** of waves is used to determine the resultant displacement. The profile of the long spring is shown in green.

Formation of stationary waves

Imagine a string stretched between two fixed points, for example a guitar string. Pulling the middle of the string and then releasing it produces a stationary wave. There is a node at each of the fixed ends and an antinode in the middle. Releasing the string produces two progressive waves travelling in opposite directions. These are reflected at the fixed ends. The reflected waves combine to produce the stationary wave.

Figure 15.3 shows how a stationary wave can be set up using a long spring. A stationary wave is formed whenever two progressive waves of the same amplitude and wavelength, travelling in **opposite** directions, superpose. Figure 15.5 uses a displacement–distance graph (s–x) to illustrate the formation of a stationary wave along a long spring (or a stretched length of string):

- At time $t = 0$, the progressive waves travelling to the left and right are in phase. The waves combine **constructively**, giving an amplitude twice that of each wave.
- After a time equal to one-quarter of a period ($t = \frac{T}{4}$), each wave has travelled a distance of one quarter of a wavelength to the left or right. Consequently, the two waves are in antiphase (phase difference = 180°). The waves combine **destructively**, giving zero displacement.
- After a time equal to one-half of a period ($t = \frac{T}{2}$), the two waves are back in phase again. They once again combine **constructively**.
- After a time equal to three-quarters of a period ($t = \frac{3T}{4}$), the waves are in antiphase again. They combine **destructively**, with the resultant wave showing zero displacement.
- After a time equal to one whole period ($t = T$), the waves combine **constructively**. The profile of the spring is as it was at $t = 0$.

This cycle repeats itself, with the long spring showing nodes and antinodes along its length. The separation between adjacent nodes or antinodes tells us about the progressive waves that produce the stationary wave.

A closer inspection of the graphs in Figure 15.5 shows that the separation between adjacent nodes or antinodes is related to the wavelength λ of the progressive wave. The important conclusions are:

separation between two adjacent nodes (or between two adjacent antinodes) $= \frac{\lambda}{2}$

separation between adjacent node and antinode $= \frac{\lambda}{4}$

The wavelength λ of **any** progressive wave can be determined from the separation between neighbouring nodes or antinodes of the resulting standing wave pattern. (This separation is $= \frac{\lambda}{2}$.) This can then be used to determine either the speed v of the progressive wave or its frequency f by using the wave equation:

$v = f\lambda$

It is worth noting that a stationary wave does not travel and therefore has no speed. It does not transfer energy between two points like a progressive wave. Table 15.1 shows some of the key features of a progressive wave and its stationary wave.

	Progressive wave	Stationary wave
wavelength	λ	λ
frequency	f	f
speed	v	zero

Table 15.1 A summary of progressive and stationary waves.

QUESTION

1 A stationary (standing) wave is set up on a vibrating spring. Adjacent nodes are separated by 25 cm. Determine:

a the wavelength of the stationary wave

b the distance from a node to an adjacent antinode.

BOX 15.1: Observing stationary waves

Here we look at experimental arrangements for observing stationary waves, for mechanical waves on strings, microwaves, and sound waves in air columns.

Stretched strings – Melde's experiment

A string is attached at one end to a vibration generator, driven by a signal generator (Figure 15.6). The other end hangs over a pulley and weights maintain the tension in the string. When the signal generator is switched on, the string vibrates with small amplitude. Larger amplitude stationary waves can be produced by adjusting the frequency.

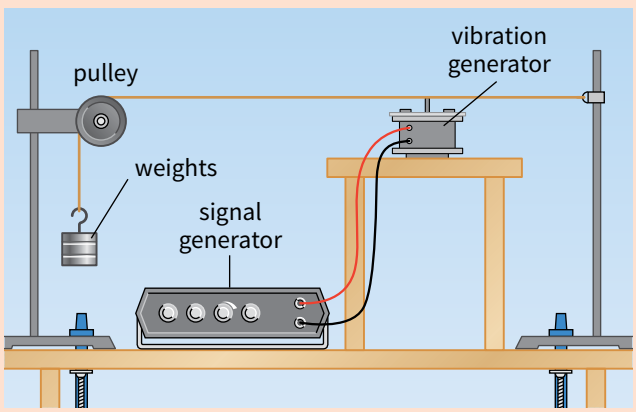


Figure 15.6 Melde's experiment for investigating stationary waves on a string.

The pulley end of the string cannot vibrate; this is a node. Similarly, the end attached to the vibrator can only move a small amount, and this is also a node. As the frequency is increased, it is possible to observe one loop (one antinode), two loops, three loops and more. Figure 15.7 shows a vibrating string where the frequency of the vibrator has been set to produce two loops.

A flashing stroboscope is useful to reveal the motion of the string at these frequencies, which look blurred to the eye. The frequency of vibration is set so that there are two loops along the string; the frequency of the stroboscope is set so that it almost matches that of the vibrations. Now we can see the string moving 'in slow motion', and it is easy to see the opposite movements of the two adjacent loops.

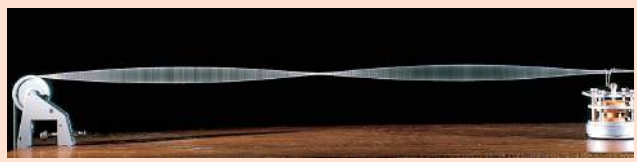


Figure 15.7 When a stationary wave is established, one half of the string moves upwards as the other half moves downwards. In this photograph, the string is moving too fast to observe the effect.

BOX 15.1: Observing stationary waves (continued)

This experiment is known as **Melde's experiment**, and it can be extended to investigate the effect of changing the length of the string, the tension in the string and the thickness of the string.

Microwaves

Start by directing the microwave transmitter at a metal plate, which reflects the microwaves back towards the source (Figure 15.8). Move the probe receiver around in the space between the transmitter and the reflector and you will observe positions of high and low intensity. This is because a stationary wave is set up between the transmitter and the sheet; the positions of high and low intensity are the antinodes and nodes respectively.

If the probe is moved along the direct line from the transmitter to the plate, the wavelength of the microwaves can be determined from the distance between the nodes. Knowing that microwaves travel at the speed of light c ($3.0 \times 10^8\,\text{m s}^{-1}$), we can then determine their frequency f using the wave equation:

$c = f\lambda$

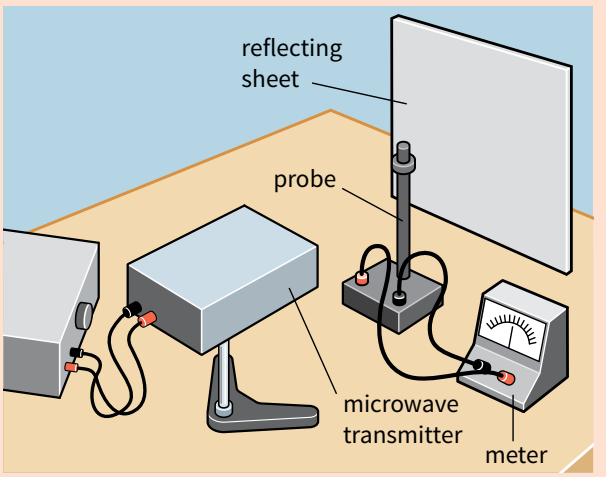


Figure 15.8 A stationary wave is created when microwaves are reflected from the metal sheet.

An air column closed at one end

A glass tube (open at both ends) is clamped so that one end dips into a cylinder of water. By adjusting its height in the clamp, you can change the length of the column of air in the tube (Figure 15.9). When you hold a vibrating tuning fork above the open end, the air column may be forced to vibrate, and the note of the tuning fork sounds much louder. This is an example of a phenomenon called **resonance**. The experiment described here is known as the **resonance tube**.

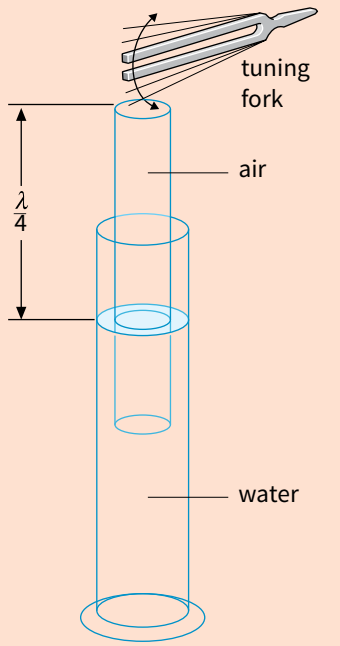


Figure 15.9 A stationary wave is created in the air in the tube when the length of the air column is adjusted to the correct length.

For resonance to occur, the length of the air column must be just right. The air at the bottom of the tube is unable to vibrate, so this point must be a node. The air at the open end of the tube can vibrate most freely, so this is an antinode. Hence the length of the air column must be one-quarter of a wavelength (Figure 15.10a). (Alternatively, the length of the air column could be set to equal three-quarters of a wavelength – see Figure 15.10b.)

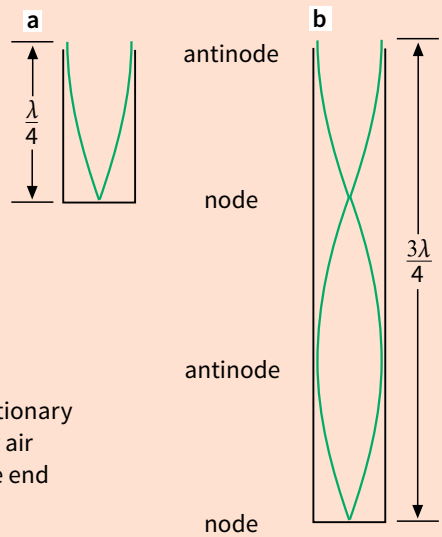


Figure 15.10 Stationary wave patterns for air in a tube with one end closed.

BOX 15.1: Observing stationary waves (continued)

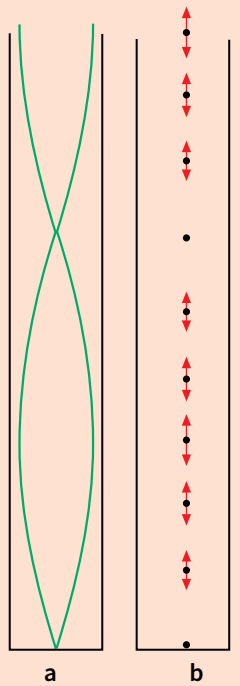


Figure 15.11 **a** The standard representation of a standing sound wave may suggest that it is a transverse wave. **b** A sound wave is really a longitudinal wave, so that the particles vibrate as shown.

Take care! The representation of standing sound waves can be misleading. Remember that a sound wave is a longitudinal wave, but the diagram we draw is more like a transverse wave. Figure 15.11a shows how we normally represent a standing sound wave, while Figure 15.11b shows the direction of vibration of the particles along the wave.

Open-ended air columns

The air in a tube which is open at both ends will vibrate in a similar way to that in a closed column. Take an open-ended tube and blow gently across the top. You should hear a note whose pitch depends on the length of the tube. Now cover the bottom of the tube with the palm of your hand and repeat the process. The pitch of the note now produced will be about an octave higher than the previous note, which means that the frequency is approximately twice the original frequency.

It is rather surprising that a standing wave can be set up in an open column of air in this way. What is going on? Figure 15.12 compares the situation for open and closed tubes. An open-ended tube has two open ends, so there must be an antinode at each end. There is a node at the midpoint.

For a tube of length l you can see that in the closed tube the standing wave formed is one-quarter of a wavelength, so the wavelength is $4l$, whereas in the open tube it is half a wavelength, giving a wavelength of $2l$. Closing one end of the tube thus halves the wavelength of the note and so the frequency doubles.

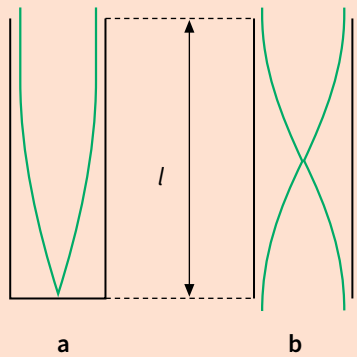


Figure 15.12 Standing wave patterns for sound waves in **a** a closed tube, and **b** an open tube.

QUESTIONS

2 Look at the stationary (standing) wave on the string in Figure 15.7 on page 213. The length of the vibrating section of the string is 60 cm.

a Determine the wavelength of the stationary wave and the separation of the two neighbouring antinodes.

The frequency of vibration is increased until a stationary wave with three antinodes appears on the string.

b Sketch a stationary wave pattern to illustrate the appearance of the string.

c What is the wavelength of this stationary wave?

3 **a** Draw a stationary wave pattern for the microwave experiment in Box 15.1. Clearly show whether there is a node or an antinode at the reflecting sheet.

b The separation of two adjacent points of high intensity is found to be 14 mm. Calculate the wavelength and frequency of the microwaves.

4 Explain how two sets of identical but oppositely travelling waves are established in the microwave and air column experiments described in Box 15.1.

Stationary waves and musical instruments

The production of different notes by musical instruments often depends on the creation of stationary waves (Figure 15.13). For a stringed instrument such as a guitar, the two ends of a string are fixed, so nodes must be established at these points. When the string is plucked half-way along its length, it vibrates with an antinode at its midpoint. This is known as the **fundamental mode of vibration** of the string. The fundamental frequency is the **minimum frequency** of a standing wave for a given system or arrangement.



Figure 15.13 When a guitar string is plucked, the vibrations of the strings continue for some time afterwards. Here you can clearly see a node close to the end of each string.

Similarly, the air column inside a wind instrument is caused to vibrate by blowing, and the note that is heard depends on a stationary wave being established. By changing the length of the air column, as in a trombone, the note can be changed. Alternatively, holes can be uncovered so that the air can vibrate more freely, giving a different pattern of nodes and antinodes.

In practice, the sounds that are produced are made up of several different stationary waves having different patterns of nodes and antinodes. For example, a guitar string may vibrate with two antinodes along its length. This gives a note having twice the frequency of the fundamental, and is described as a **harmonic** of the fundamental. The musician's skill is in stimulating the string or air column to produce a desired mixture of frequencies.

The frequency of a harmonic is always a multiple of the fundamental frequency. The diagrams show some of the modes of vibration of a fixed length of string (Figure 15.14) and an air column in a tube of a given length that is closed at one end (Figure 15.15).

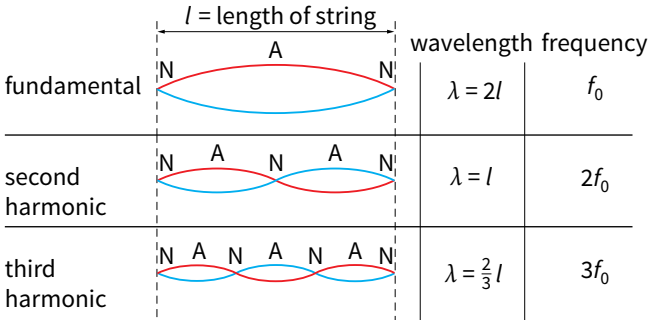


Figure 15.14 Some of the possible stationary waves for a fixed string of length l. The frequency of the harmonics is a multiple of the fundamental frequency f_0.

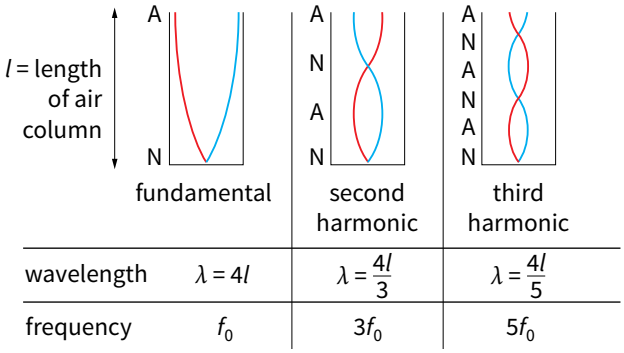


Figure 15.15 Some of the possible stationary waves for an air column, closed at one end. The frequency of each harmonic is an odd multiple of the fundamental frequency f_0.

Determining the wavelength and speed of sound

Since we know that adjacent nodes (or antinodes) of a stationary wave are separated by half a wavelength, we can use this fact to determine the wavelength λ of a progressive wave. If we also know the frequency f of the waves, we can find their speed v using the wave equation $v = f\lambda$.

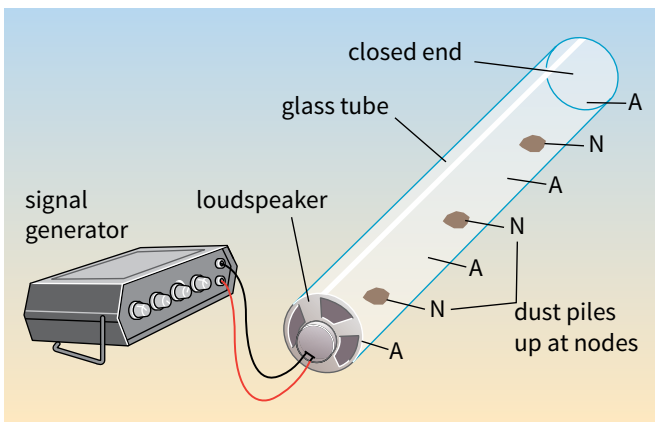


Figure 15.16 Kundt's dust tube can be used to determine the speed of sound.

One approach uses Kundt's dust tube (Figure 15.16). A loudspeaker sends sound waves along the inside of a tube. The sound is reflected at the closed end. When a stationary wave is established, the dust (fine powder) at the antinodes vibrates violently. It tends to accumulate at the nodes, where the movement of the air is zero. Hence the positions of the nodes and antinodes can be clearly seen.

BOX 15.2: Using stationary sound waves to determine λ and v

This method is shown in Figure 15.17; it is the same arrangement as used for microwaves (Box 15.1 above). The loudspeaker produces sound waves, and these are reflected from the vertical board. The microphone detects the stationary sound wave in the space between the speaker and the board, and its output is displayed on the oscilloscope. It is simplest to turn off the time-base of the oscilloscope, so that the spot no longer moves across the screen. The spot moves up and down the screen, and the height of the vertical trace gives a measure of the intensity of the sound.

By moving the microphone along the line between the speaker and the board, it is easy to detect nodes and antinodes. For maximum accuracy, we do not measure the separation of adjacent nodes; it is better to measure the distance across several nodes.

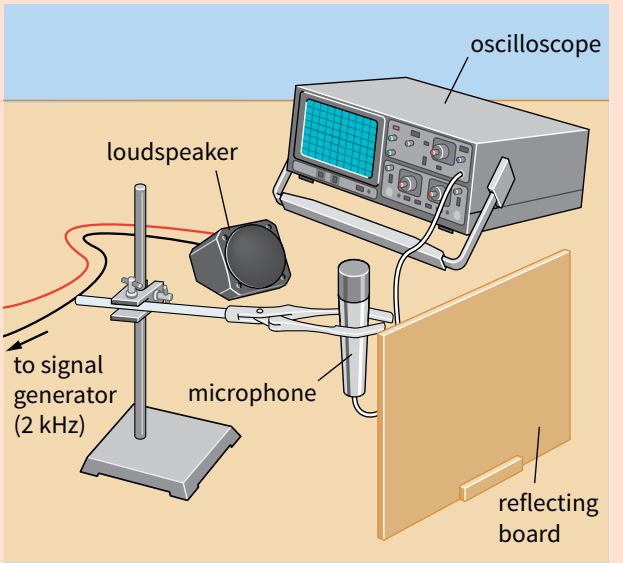


Figure 15.17 A stationary sound wave is established between the loudspeaker and the board.

QUESTIONS

5 **a** For the arrangement shown in Figure 15.17, suggest why it is easier to determine accurately the position of a node rather than an antinode.

b Explain why it is better to measure the distance across several nodes.

6 For sound waves of frequency 2500 Hz, it is found that two nodes are separated by 20 cm, with three antinodes between them.

a Determine the wavelength of these sound waves.

b Use the wave equation $v = f\lambda$ to determine the speed of sound in air.

BOX 15.3: Reducing and eliminating errors

The resonance tube experiment (Figure 15.9 on page 214) can be used to determine the wavelength and speed of sound with a high degree of accuracy. However, to do this, it is necessary to take account of a systematic error in the experiment.

Look at the representation of the stationary waves in the tubes shown in Figure 15.10 on page 214. In each case, the antinode at the top of the tube is shown extending slightly beyond the open end of the tube. This is because experiment shows that the air slightly beyond the end of the tube vibrates as part of the stationary wave. This is shown more clearly in Figure 15.18.

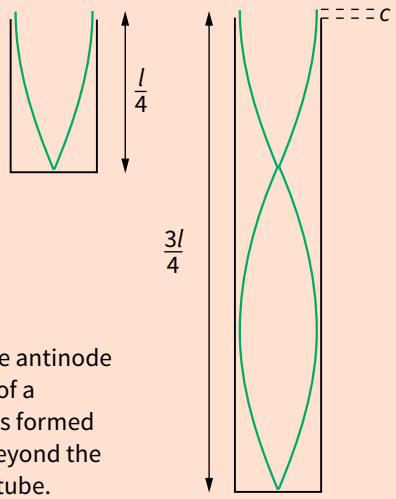


Figure 15.18 The antinode at the open end of a resonance tube is formed at a distance c beyond the open end of the tube.

BOX 15.3: Reducing and eliminating errors (continued)

The antinode is at a distance c beyond the end of the tube, where c is called the **end-correction**. Unfortunately, we do not know the value of c. It cannot be measured directly. However, we can write:

for the shorter tube, $\frac{\lambda}{4} = l_1 + c$

for the longer tube, $\frac{3\lambda}{4} = l_2 + c$

Subtracting the first equation from the second equation gives:

$$\frac{3\lambda}{4} - \frac{\lambda}{4} = (l_2 + c) - (l_1 + c)$$

Simplifying gives:

$$\frac{\lambda}{2} = l_2 - l_1$$

Hence:

$$\lambda = 2(l_2 - l_1)$$

So, although we do not know the value of c, we can make two measurements (l_1 and l_2) and obtain an accurate value of λ. (You may be able to see from Figure 15.18 that the difference in lengths of the two tubes is indeed equal to half a wavelength.)

The end-correction c is an example of a **systematic error**. When we measure the length l of the tube, we are measuring a length which is consistently less than the quantity we really need to know ($l + c$). However, by understanding how the systematic error affects the results, we have been able to remove it from our measurements.

You will find more detailed discussion of systematic errors in Chapter P1: Practical skills at AS level.

QUESTION

7 In a resonance tube experiment, resonance is obtained for sound waves of frequency 630 Hz when the length of the air column is 12.6 cm and again when it is 38.8 cm. Determine:

- **a** the wavelength of the sound waves causing resonance
- **b** the end-correction for this tube
- **c** the speed of sound in air.

Summary

- Stationary waves are formed when two identical waves travelling in opposite directions meet and superpose. This usually happens when one wave is a reflection of the other.
- A stationary wave has a characteristic pattern of nodes and antinodes.
- A node is a point where the amplitude is always zero.
- An antinode is a point of maximum amplitude.
- Adjacent nodes (or adjacent antinodes) are separated by a distance equal to half a wavelength.
- We can use the wave equation $v = f\lambda$ to determine the speed v or the frequency f of a progressive wave. The wavelength λ is found using the nodes or antinodes of the stationary wave pattern.

End-of-chapter questions

1 Figure 15.19 shows a stationary wave on a string.

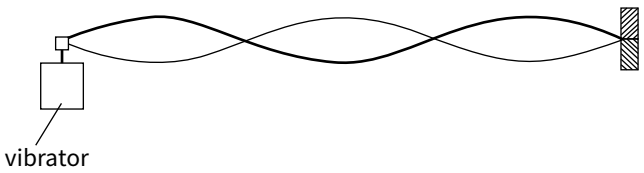


Figure 15.19 For End-of-chapter Question 1.

a On a copy of Figure 15.19, label one **node** (N) and one **antinode** (A). [1]

b Mark on your diagram the wavelength of the standing wave and label it λ. [1]

c The frequency of the vibrator is doubled. Describe the changes in the standing wave pattern. [1]

2 A tuning fork which produces a note of 256 Hz is placed above a tube which is nearly filled with water. The water level is lowered until resonance is first heard.

a Explain what is meant by the term **resonance**. [1]

b The length of the column of air above the water when resonance is first heard is 31.2 cm. Calculate the speed of the sound wave. [2]

3 **a** State **two** similarities and **two** differences between progressive waves and stationary waves. [4]

b Figure 15.20 shows an experiment to measure the speed of a sound in a string. The frequency of the vibrator is adjusted until the standing wave shown in Figure 15.20 is formed.

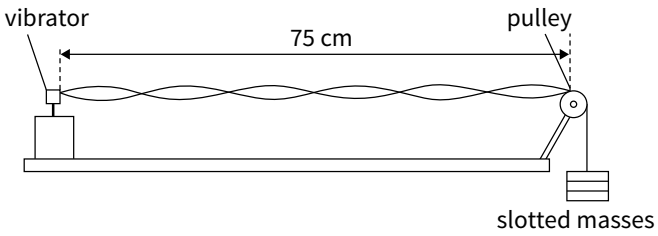


Figure 15.20 For End-of-chapter Question 3.

i On a copy of the diagram, mark a node (label it N) and an antinode (label it A). [2]

ii The frequency of the vibrator is 120 Hz. Calculate the speed at which a progressive wave would travel along the string. [3]

c The experiment is now repeated with the load on the string halved. In order to get a similar standing wave the frequency has to be decreased to 30 Hz. Explain, in terms of the speed of the wave in the string, why the frequency must be adjusted. [2]

4 Figure 15.21 shows a standing wave, of frequency 400 Hz, produced by a loudspeaker in a closed tube.

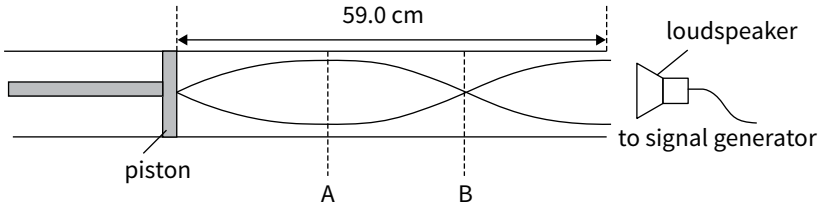


Figure 15.21 For End-of-chapter Question 4.

a Describe the movement of the air particles at:

i A [2]

ii B. [1]

b The piston is slowly moved outwards. The next resonant peak is heard when the length of the column of air is 99.4 cm.

Calculate:

i the speed of the sound [4]

ii the end correction which must be applied when using this tube. [1]

5 **a** Explain what is meant by:

i a **coherent** source of waves. [2]

ii **phase difference**. [2]

b A student, experimenting with microwaves, sets up the arrangement shown in Figure 15.22.

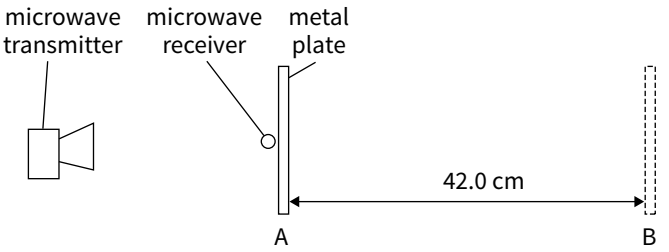


Figure 15.22 For End-of-chapter Question 5.

With the metal plate at position A there is a very small signal. He slowly moves the plate back, leaving the receiver in the same position. As he does so he finds that the intensity initially rises until it becomes a maximum, then falls back to a minimum. This cycle repeats a total of five times until the plate reaches position B, where once again there is a minimum.

i Explain why a series of maxima and minima are heard. [2]

ii Calculate the frequency of the microwaves. [5]

c Explain why there was a minimum when the plate was at position A, next to the detector. [2]

6 Figure 15.23 shows an experiment to measure the speed of sound in air.

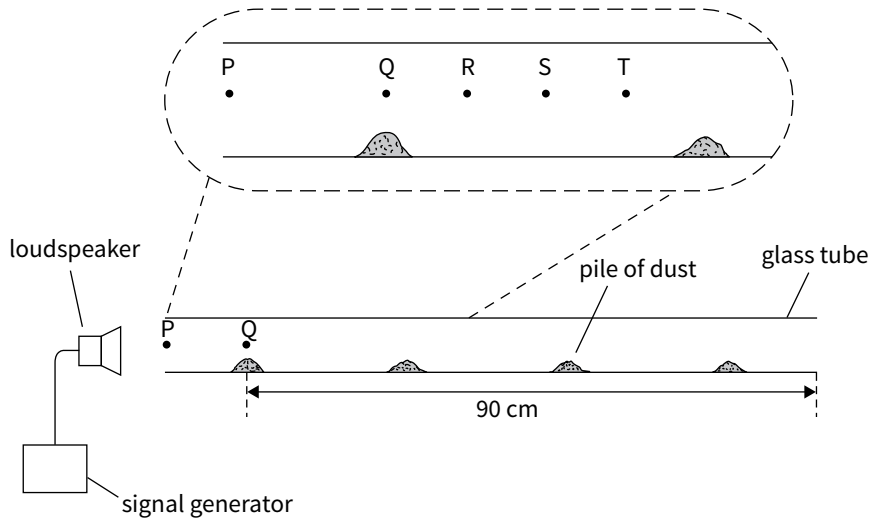


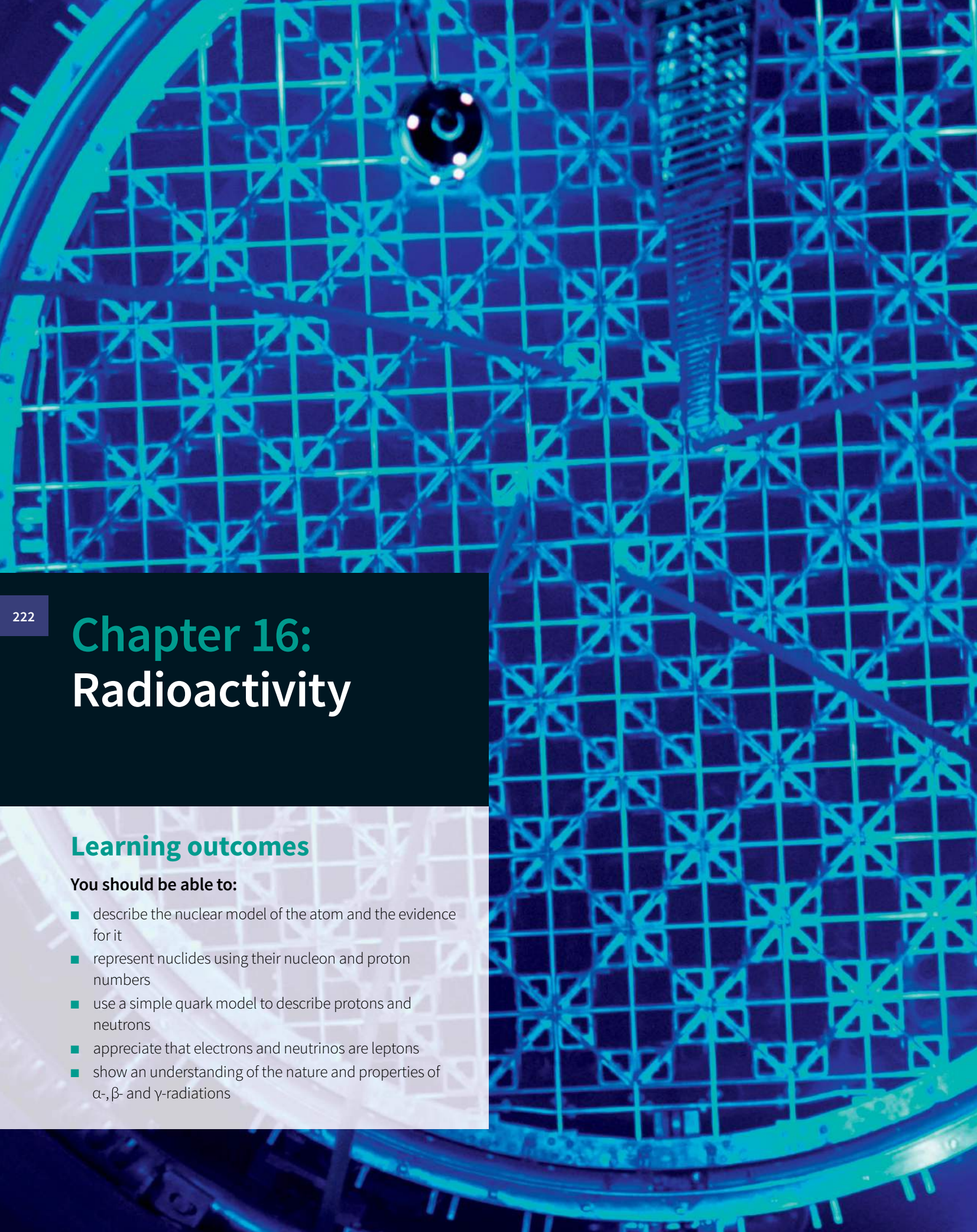
Figure 15.23 For End-of-chapter Question 6.

A small amount of dust is scattered along the tube. The loudspeaker is switched on. When the frequency is set at 512 Hz the dust collects in small piles as shown in Figure 15.23.

a Determine the wavelength of the sound wave and calculate the speed of sound in the air in the tube. [3]

b On a copy of the diagram, show the movement of the air particles at positions P, Q, R, S and T. [3]

c Mark two points on your diagram where the movements of the air particles are 180° out of phase with each other – label them A and B. [1]



Chapter 16: Radioactivity

Learning outcomes

You should be able to:

- describe the nuclear model of the atom and the evidence for it
- represent nuclides using their nucleon and proton numbers
- use a simple quark model to describe protons and neutrons
- appreciate that electrons and neutrinos are leptons
- show an understanding of the nature and properties of α-, β- and γ-radiations

Radioactivity at work

Radioactive substances have many uses, for example in engineering and medicine. They must be handled with great care to ensure that no-one becomes contaminated and so exposed to the radiation which comes from these substances (Figure 16.1). In this chapter we will look at the nature of radioactive substances and the different types of radiation they produce.



Figure 16.1 A worker at a nuclear power station is checked for any radioactive material on his body.

Looking inside the atom

The idea that matter is composed of very small particles called atoms was first suggested by the Greeks some 2000 years ago. However, it was not until the middle of the 19th century that any ideas about the **inside** of the atom were proposed.

It was the English scientist J.J. Thomson who suggested that the atom is a neutral particle made of a positive charge with lumps of negative charge (electrons) in it. He could not determine the charge and the mass of the negative particles separately, but it was clear that a new particle, probably much smaller than the hydrogen atom, had been discovered. Since atoms are neutral and physicists had discovered a negatively charged part of an atom, it meant that there were both positive and negative charges in an atom. We now call this the **plum pudding model** of the atom (positive pudding with negative plums!).

Other experiments show that the electron has a mass of approximately 9.11×10^{-31} kg (m_e) and a charge of -1.60×10^{-19} C ($-e$). Today we use the idea of the electron to explain all sorts of phenomena, including electrostatics, current electricity and electronics.

Alpha-particle scattering and the nucleus

Early in the 20th century, many physicists were investigating the recently discovered phenomenon of radioactivity, the process whereby unstable nuclei emit radiation. One kind of radiation they found consisted of what they called α-particles (alpha-particles).

These α-particles were known to be smaller than atoms, and had relatively high kinetic energies. Hence they were useful in experiments designed to discover the composition of atoms.



Figure 16.2 Ernest Rutherford (on the right) in the Cavendish Laboratory, Cambridge, England. He had a loud voice that could disturb sensitive apparatus and so the notice was a joke aimed at him.

In 1906, while experimenting with the passage of α-particles through a thin mica sheet, Ernest Rutherford (Figure 16.2) noticed that most of the α-particles passed straight through. This suggested to him that there might be a large amount of empty space in the atom, and by 1909 he had developed what we now call the **nuclear model of the atom**.

In 1911 Rutherford carried out a further series of experiments with Hans Geiger and Ernest Marsden at the University of Manchester using gold foil in place of the mica. They directed parallel beams of α-particles at a piece of gold foil only 10^{-6} m thick. Most of the α-particles went straight through. Some were deflected slightly, but about 1 in 20 000 were deflected through an angle of more than 90°, so that they appeared to bounce back off the foil. This helped to confirm Rutherford in his thinking about the atom – that it was mostly empty space, with most of the mass and all of the positive charge concentrated in a tiny region at the centre. This central **nucleus** only affected the α-particles when they came close to it.

Later, Rutherford wrote: 'It was quite the most incredible event that has happened to me in my life. It was almost as incredible as if you fired a 15-inch shell at a piece of tissue paper and it came back and hit you.' In fact, he was not quite as surprised as this suggests, because the results confirmed ideas he had used in designing the experiment.

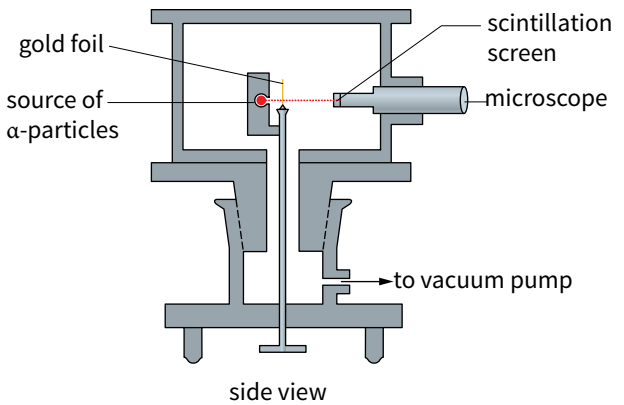


Figure 16.3 The apparatus used for the α-scattering experiment. The microscope can be moved round to detect scattered radiation at different angles.

Figure 16.3 shows the apparatus used in the α-scattering experiment. Notice the following points:

- The α-particle source was encased in metal with a small aperture, allowing a fine beam of α-particles to emerge.
- Air in the apparatus was pumped out to leave a vacuum; α-radiation is absorbed by a few centimetres of air.
- One reason for choosing gold was that it can be made into a very thin sheet or foil. Rutherford's foil was only a few hundreds of atoms thick.
- The α-particles were detected when they struck a solid 'scintillating' material. Each α-particle gave a tiny flash of light and these were counted by the experimenters (Geiger and Marsden).
- The detector could be moved round to detect α-particles scattered through different angles.

Geiger and Marsden had the difficult task of observing and counting the tiny flashes of light produced by individual α-particles striking the scintillation screen. They had to spend several minutes in the darkened laboratory to allow the pupils of their eyes to become dilated so that they could see the faint flashes. Each experimenter could only stare into the detector for about a minute before the strain was too much and they had to change places.

Explaining α-scattering

How can we explain the back-scattering of α-particles by the gold atoms?

If the atom was as Thomson pictured it, with negatively charged electrons scattered through a 'pudding' of positive charge, an individual α-particle would pass through it like a bullet, hardly being deflected at all. This is because the α-particles are more massive than electrons – they might push an electron out of the atom, but their own path would be scarcely affected.

On the other hand, if the mass and positive charge of the atom were concentrated at one point in the atom, as Rutherford suggested, an α-particle striking this part would be striking something more massive than itself and with a greater charge. A head-on collision would send the α-particle backwards.

The paths of an α-particle near a nucleus are shown in Figure 16.5. Rutherford reasoned that the large deflection of the α-particle must be due to a very small charged nucleus. From his experiments he calculated that the diameter of the gold nucleus was about 10^{-14} m. It has since been shown that the very large deflection of the α-particle is due to the electrostatic repulsion between the positive charge of the α-particle and the positive charge of the nucleus of the atom. The closer the path of the α-particle gets to the nucleus, the greater will be this repulsion. An α-particle making a 'head-on' collision with a nucleus is back-scattered through 180°. The α-particle and nucleus both experience an equal but opposite repulsive electrostatic force F. This force has a much greater effect on the motion of the α-particle than on the massive nucleus of gold.

BOX 16.1: An analogy for Rutherford scattering

A very simple analogy (or model) of the experiment is shown in Figure 16.4. When you roll a ball-bearing down a slope towards the 'cymbal', it may be deflected, but even if it is rolled directly at the cymbal's centre, it does not come back – it rolls over the centre and carries on to the other side. However, using the 'tin hat' shape, with a much narrower but higher central bulge, any ball-bearings rolled close to the centre will be markedly deflected, and those rolled directly towards it will come straight back.

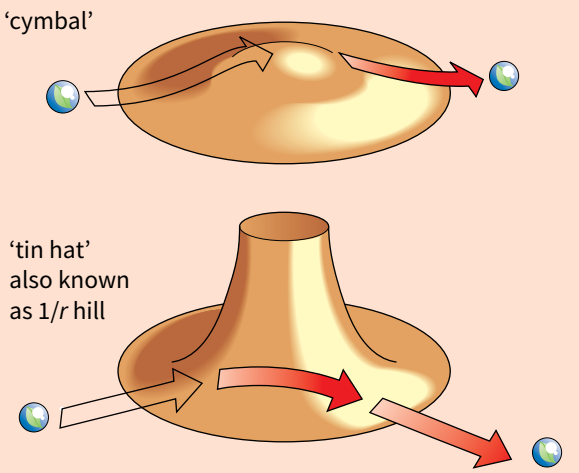


Figure 16.4 An analogy for Rutherford's experiment.

The shape of the cymbal represents the shape of the electric field of an atom in the 'plum pudding' model: low central intensity and spread out. The 'tin hat' represents the shape of the electric field for the nuclear model: high central intensity and concentrated.

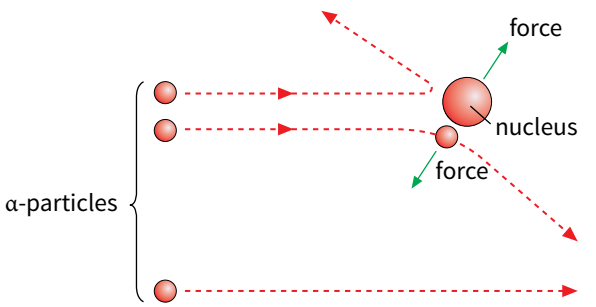


Figure 16.5 Possible paths of an α-particle near a nucleus. The nucleus and the α-particle both experience electrostatic repulsion.

From the α-particle scattering experiment, Rutherford deduced the following.

- An α-particle is deviated due to the repulsive force between the α-particle and the positive charge in the atom.
- Most α-particles have little or no deviation – so most of an atom is empty space.
- A very few α-particles are deviated more than 90° – so most of the mass of an atom is concentrated in a small space (the nucleus) and most of the atom is empty space.

QUESTIONS

1 Rutherford's scattering experiments were done in an evacuated container. Explain why this is necessary.

2 In Rutherford's experiment, α-particles were directed at a thin gold foil. A small fraction of the α-particles were back-scattered through 180°.

Describe and explain how the fraction back-scattered changes if each of the following changes are (separately) made.

a A thicker foil is used.

b Faster α-particles are used.

c A silver foil is used – a silver nucleus has less positive charge than a gold nucleus.

A simple model of the atom

After Rutherford had presented his findings, the nuclear model of the atom gained rapid acceptance. This was partly because it helped chemists to explain the phenomenon of chemical bonding (the way in which atoms bond together to form molecules). Subsequently, the proton was discovered. It had a positive charge, equal and opposite to that of the electron. However, its mass was too small to account for the entire mass of the atom and it was not until the early 1930s that this puzzle was solved by the discovery of the neutron, an uncharged particle with a similar mass to that of the proton. This suggests a model for the atom like the one shown in Figure 16.6:

- Protons and neutrons make up the nucleus of the atom.
- The electrons move around the nucleus in a cloud, some closer to and some further from the centre of the nucleus.

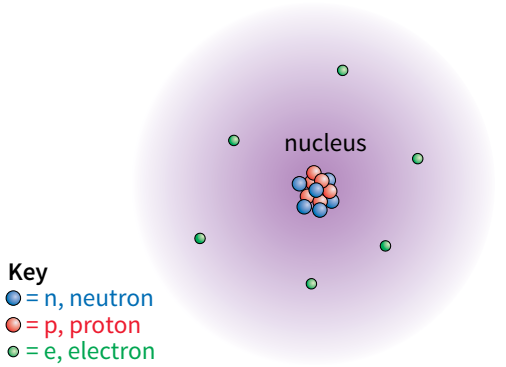


Figure 16.6 A simple model of the atom. If the nucleus were drawn to scale, it would be invisible (and the electrons are even smaller!).

From this model it looks as though all matter, including ourselves, is mostly empty space. For example, if we scaled up the hydrogen atom so that the nucleus was the size of a 1 cm diameter marble, the orbiting electron would be a grain of sand some 800 m away!

The scale of things

It is useful to have an idea of the approximate sizes of typical particles:

- radius of proton ~ radius of neutron ~ 10^{-15} m
- radius of nucleus ~ 10^{-15} m to 10^{-14} m
- radius of atom ~ 10^{-10} m
- size of molecule ~ 10^{-10} m to 10^{-6} m.

(Some molecules, such as large protein molecules, are very large indeed – compared to an atom!)

The radii of nuclear particles are often quoted in femtometres (fm), where 1 fm = 10^{-15} m.

Nuclear density

We can picture a proton as a small, positively charged sphere. Knowing its mass and radius, we can calculate its density:

mass of proton $m_p = 1.67 \times 10^{-27}$ kg

radius of proton $r = 0.80$ fm $= 0.80 \times 10^{-15}$ m

(In fact, the radius of the proton is not very accurately known; it is probably between 0.80×10^{-15} m and 0.86×10^{-15} m.)

$$\text{volume of proton} = \tfrac{4}{3}\pi r^3 = \tfrac{4}{3}\pi \times (0.80 \times 10^{-15})^3$$
$$= 2.14 \times 10^{-45}\,\text{m}^3 \approx 2.1 \times 10^{-45}\,\text{m}^3$$

$$\text{density} = \frac{\text{mass}}{\text{volume}}$$

$$\text{density} = \frac{1.67 \times 10^{-27}}{2.14 \times 10^{-45}} \approx 7.8 \times 10^{17}\,\text{kg m}^{-3}$$

So the proton has a density of roughly 10^{18} kg m^{-3}. This is also the density of a neutron, and of an atomic nucleus, because nuclei are made of protons and neutrons held closely together.

Compare the density of nuclear material with that of water whose density is 1000 kg m^{-3} – the nucleus is 10^{15} times as dense. Nuclear matter the size of a tiny grain of sand would have a mass of about a million tonnes! This is a consequence of the fact that the nucleus occupies only a tiny fraction of the volume of an atom. The remainder is occupied by the cloud of orbiting electrons whose mass makes up less than one-thousandth of the atomic mass.

QUESTION

3 Gold has a density of 19 700 kg m^{-3}.A mass of 193 g of gold contains 6.02×10^{23} atoms. Use this information to estimate the volume of a gold atom, and hence its radius. State any assumptions you make.

Nucleons and electrons

We will start this section with a summary of the particles mentioned so far (Table 16.1). All nuclei, except the lightest form of hydrogen, contain protons and neutrons, and each nucleus is described by the number of protons and neutrons that it contains.

- Protons and neutrons in a nucleus are collectively called **nucleons**. For example, in a nucleus of gold, there are 79 protons and 118 neutrons, giving a total of 197 nucleons altogether.
- The total number of nucleons in a nucleus is called the **nucleon number** (or mass number) A.
- The nucleon number is equal to the sum of the number of neutrons in the nucleus, the **neutron number** N, and the number of protons, the **proton number** (or atomic number) Z, i.e.
  $A = N + Z$

Particle	Relative mass (proton = 1)[a]	Charge[b]
proton (p)	1	$+e$
neutron (n)	1	0
electron (e)	0.0005	$-e$
alpha-particle (α)	4	$+2e$

[a] The numbers given for the masses are approximate.
[b] $e = 1.60 \times 10^{-19}$ C.

Table 16.1 Summary of the particles that we have met so far in this chapter. The α-particle is in fact a helium nucleus (with two protons and two neutrons).

The nucleus of any atom can be represented by the symbol for the element along with the nucleon number and proton number, as shown below:

$^{\text{nucleon number}}_{\text{proton number}}\text{element symbol}\quad {}^{A}_{Z}\text{X}$

oxygen ${}^{16}_{8}\text{O}$ gold ${}^{197}_{79}\text{Au}$ uranium ${}^{238}_{92}\text{U}$

A specific combination of protons and neutrons in a nucleus is called a **nuclide**.

The proton and nucleon numbers of some common nuclides are shown in Table 16.2.

QUESTIONS

4 Table 16.2 shows the proton and nucleon numbers of several nuclei. Determine the number of neutrons in the nuclei of the following elements shown in the table:

a nitrogen
b bromine
c silver
d gold
e mercury.

5 State the charge of each of the following in terms of the elementary charge *e*:

a proton
b neutron
c nucleus
d molecule
e α-particle.

Isotopes

Although atoms of the same element may be identical chemically, their nuclei may be slightly different. The number of protons in the nucleus of an atom determines what element it is: helium always has 2 protons, carbon 6 protons, oxygen 8 protons, neon 10 protons, radium 88 protons, uranium 92 protons, and so on.

However, the number of neutrons in the nuclei for a given element can vary. Take neon as an example. Three different naturally occurring forms of neon are:

${}^{20}_{10}\text{Ne}\qquad {}^{21}_{10}\text{Ne}\qquad {}^{22}_{10}\text{Ne}$

The first has 10 neutrons in the nucleus, the second 11 neutrons and the third 12 neutrons. These three types of neon nuclei are called **isotopes** of neon. Each isotope has the same number of protons (for neon this is 10) but a different number of neutrons. The word 'isotope' comes from the Greek *isotopos* (same place), because all isotopes of the same element have the same place in the Periodic Table of elements.

Isotopes are nuclei of the same element with different numbers of neutrons but the same number of protons.

Element	Nucleon number *A*	Proton number *Z*	Element	Nucleon number *A*	Proton number *Z*
hydrogen	1	1	bromine	79	35
helium	4	2	silver	107	47
lithium	7	3	tin	120	50
beryllium	9	4	iodine	130	53
boron	11	5	caesium	133	55
carbon	12	6	barium	138	56
nitrogen	14	7	tungsten	184	74
oxygen	16	8	platinum	195	78
neon	20	10	gold	197	79
sodium	23	11	mercury	202	80
magnesium	24	12	lead	206	82
aluminium	27	13	bismuth	209	83
chlorine	35	17	radium	226	88
calcium	40	20	uranium	238	92
iron	56	26	plutonium	239	94
nickel	58	28	americium	241	95

Table 16.2 Proton and nucleon numbers of some nuclides.

Any atom is electrically neutral (it has no net positive or negative charge), so the number of electrons surrounding the nucleus must equal the number of protons in the nucleus of the atom. If an atom gains or loses an electron, it is no longer electrically neutral and is called an **ion**.

For an atom, the number of protons (and hence the number of electrons) determines the chemical properties of the atom. The number of protons and the number of neutrons determine the nuclear properties. It is important to realise that, since the number of protons, and therefore the number of electrons, in isotopes of the same element are identical, they will all have the same chemical properties but very different nuclear properties.

Hydrogen has three important isotopes, $^{1}_{1}H$, $^{2}_{1}H$ (deuterium) and $^{3}_{1}H$ (tritium) (Figure 16.7). $^{1}_{1}H$ and deuterium occur naturally, but tritium has to be made. Deuterium and tritium form the fuel of many fusion research reactors. Hydrogen is the most abundant element in the Universe (Figure 16.8), because it consists of just one proton and one electron, which is the simplest structure possible for an atom.

The different numbers of neutrons in the isotopes of an element means that the isotopes will have different relative atomic masses. There are differences too in some of their physical properties, such as density and boiling point. For example, heavy water, which is water containing deuterium, has a boiling point of 104 °C under normal atmospheric pressure.

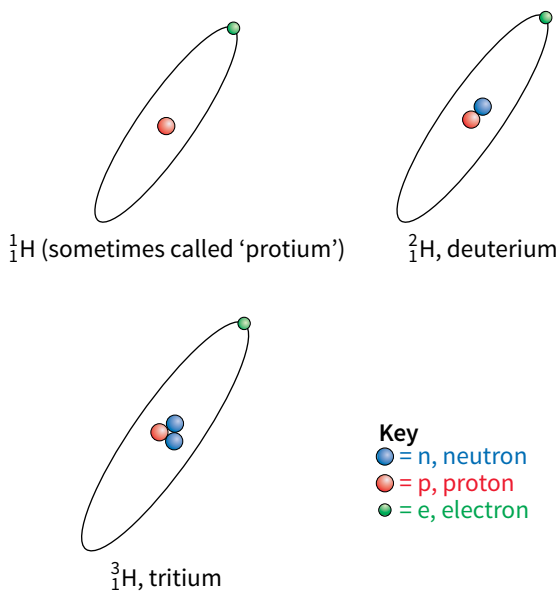


Figure 16.7 The isotopes of hydrogen.



Figure 16.8 The Horsehead Nebula in Orion. The large coloured regions are expanses of dust and gas, mostly hydrogen, that are ionised by nearby stars so that they emit light. The dark 'horse head' is where the areas of gas and dust remain in atomic form and block out the light from behind.

Table 16.3 gives details of some other commonly occurring isotopes.

Element	Nucleon number *A*	Proton number *Z*	Neutron number *N*
hydrogen	1	1	0
	2	1	1
carbon	12	6	6
	14	6	8
oxygen	16	8	8
	18	8	10
neon	20	10	10
	21	10	11
potassium	39	19	20
	40	19	21
strontium	88	38	50
	90	38	52
caesium	135	55	80
	137	55	82
lead	206	82	124
	208	82	126
radium	226	88	138
	228	88	140
uranium	235	92	143
	238	92	146

Table 16.3 Some commonly occurring isotopes.

QUESTIONS

6 Uranium has atomic number 92. Two of its common isotopes have nucleon numbers 235 and 238. Determine the number of neutrons for these isotopes.

7 There are seven naturally occurring isotopes of mercury, with nucleon numbers (and relative abundances) of 196 (0.2%), 198 (10%), 199 (16.8%), 200 (23.1%), 201 (13.2%), 202 (29.8%) and 204 (6.9%).

a Determine the proton and neutron numbers for each isotope.

b Determine the average relative atomic mass (equivalent to the 'average nucleon number') of naturally occurring mercury.

8 Group the following imaginary elements A–H into isotopes and name them using the Periodic Table in the back of the book (Appendix 3).

	A	B	C	D	E	F	G	H
Proton number	20	23	21	22	20	22	22	23
Nucleon number	44	50	46	46	46	48	50	51

Forces in the nucleus

As you know from earlier in this chapter, there are two kinds of particle in the nucleus of an atom: protons, which each carry positive charge $+e$; and neutrons, which are uncharged. It is therefore quite surprising that the nucleus holds together at all. You would expect the electrostatic repulsions from all those positively charged protons to blow it apart. The fact that this does not happen is very good evidence for the existence of an attractive force between the nucleons. This is called the **strong nuclear force**. It only acts over very short distances (10^{-14} m), and it is what holds the nucleus together.

Diluting the protons

In small nuclei the strong nuclear force from all the nucleons reaches most of the others in the nucleus, but as we go on adding protons and neutrons the balance becomes much finer. The longer-range electrostatic force affects the whole nucleus, but the short-range strong nuclear force of any particular nucleon only affects those nucleons around it – the rest of the nucleus is unaffected. In a large nucleus the nucleons are not held together so tightly, and this can make the nucleus unstable. The more protons there are in a nucleus, the greater the electric forces between them, and we need a few extra neutrons to help 'keep the protons apart'. This is why heavy nuclei have more neutrons than protons.

The proton and neutron numbers for some common nuclides are shown in Table 16.3. You can see that for light elements these two numbers are the same, but they become very different for heavy elements. Adding more neutrons helps to keep the nucleus stable, but when the number of protons is greater than 83, adding more neutrons is not enough. Elements with a proton number greater than 83 are all unstable – they undergo radioactive decay.

Most atoms that make up our world have stable nuclei; that is, they do not change as time goes by, which is quite fortunate really! However, some are less stable and give out radiation. Whether or not an atom is unstable depends on the numbers of protons and neutrons in its nucleus. Hydrogen-1 (1p), helium-4 (2p, 2n), carbon-12 (6p, 6n) and oxygen-16 (8p, 8n) are all stable – but add or subtract neutrons and the situation changes.

For example, add a neutron to helium-4 and you get helium-5, a very unstable nucleus – it undergoes radioactive emission. (There is much more about radioactive decay later in this chapter.)

QUESTION

9 State which of the following forces act between protons and neutrons in a nucleus.

a gravitational

b electrostatic

c strong nuclear.

Fundamental particles?

Chemistry is very complicated because there are literally billions of different molecules that can exist. The discovery of the Periodic Table simplified things because it suggested that there were roughly 92 different elements whose atoms could be arranged to make these various molecules. The idea that atoms are made up of just three types of particle (protons, neutrons and electrons) seemed to simplify things still more, and scientists were very happy with it because it seemed to provide a very simple explanation of a complex world. Protons, neutrons and electrons were thought of as fundamental particles, which could not be subdivided further.

However, in the middle decades of the 20th century, physicists discovered many other particles that did not fit this pattern. They gave them names such as pions, kaons, muons, etc., using up most of the letters of the Greek alphabet.

These new particles were found in two ways:

- by looking at cosmic rays, which are particles that arrive at the Earth from outer space
- by looking at the particles produced by high-energy collisions in particle accelerators (Figure 16.9).

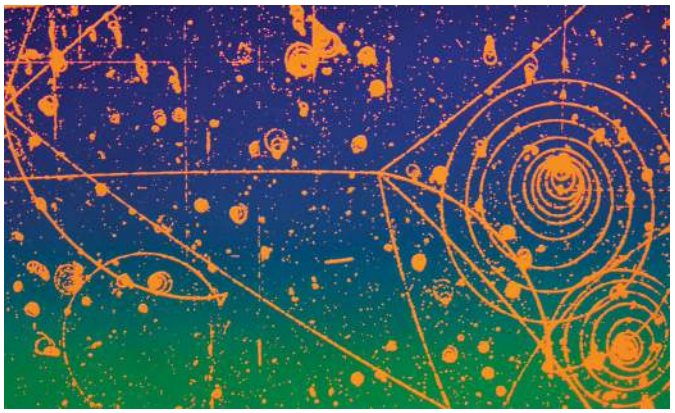


Figure 16.9 Particle tracks in a bubble chamber detector. A particle has entered from the left and then struck another particle just to the right of the centre. Four new particles fly out from the point of impact.

The discovery of new particles with masses different from those of protons, neutrons and electrons suggested that these were not fundamental particles. Various attempts were made to tidy up this very confusing picture.

In principle, we can never know for certain whether a particle such as the electron is truly fundamental; the possibility will always remain that a physicist will discover some deeper underlying structure.

Families of particles

Today, sub-atomic particles are divided into two families:

- **Hadrons** such as protons and neutrons. These are all particles that are affected by the strong nuclear force.
- **Leptons** such as electrons. These are particles that are unaffected by the strong nuclear force.

The word 'hadron' comes from a Greek word meaning 'bulky', while 'lepton' means 'light' (in mass). It is certainly true that protons and neutrons are bulky compared to electrons.

At the Large Hadron Collider (Figure 16.10) at the CERN laboratory in Geneva, physicists are experimenting with hadrons in the hope of finding answers to some fundamental questions about this family of particles. In 2013, they announced the discovery of the Higgs boson, a particle which was predicted 50 years earlier and which is required to explain why matter has mass.



Figure 16.10 Particle accelerators have become bigger and bigger as scientists have sought to look further and further into the fundamental nature of matter. This is one of the particle detectors of the Large Hadron Collider (LHC), as it was about to be installed. The entire collider is 27 km in circumference.

Inside hadrons

To sort out the complicated picture of the hadron family of particles, Murray Gell-Mann in 1964 proposed a new model. He suggested that they were made up of just a few different particles, which he called **quarks**.

Figure 16.11 shows icons used to represent three quarks, together with the corresponding antiquarks. These are called the up (u), down (d) and strange (s) quarks. Gell-Mann's idea was that there are two types of hadron: baryons, made up of three quarks, and mesons, made up of two quarks. In either case, the quarks are held together by the strong nuclear force. For example:

- A proton is made up of two up quarks and a down quark; proton = (uud).
- A neutron is made up of one up quark and two down quarks; neutron = (udd).
- A pi^+ meson is made up of an up quark and a down antiquark; pi^+ meson = $(u\bar{d})$.
- A phi meson is made up of a strange quark and an antistrange quark; phi meson = $(s\bar{s})$.

Antiquarks are shown with a 'bar' on top of the letter for the quark. Antiquarks are needed to account for the existence of antimatter. This is matter that is made of antiparticles; when a particle meets its antiparticle, they annihilate each other, leaving only photons of energy.

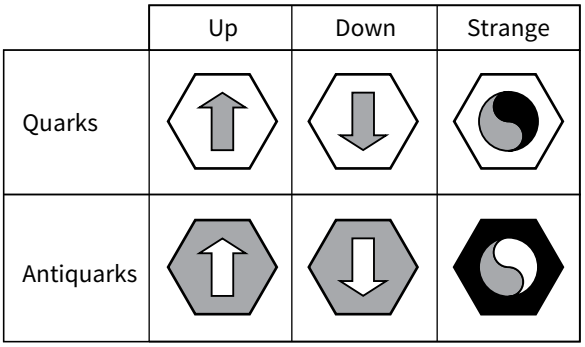


Figure 16.11 Icons representing three 'flavours' of quark, up, down and strange, and their antiquarks.

Discovering radioactivity

The French physicist Henri Becquerel (Figure 16.12) is credited with the discovery of radioactivity in 1896. He had been looking at the properties of uranium compounds when he noticed that they affected photographic film – he realised that they were giving out radiation all the time and he performed several ingenious experiments to shed light on the phenomenon.

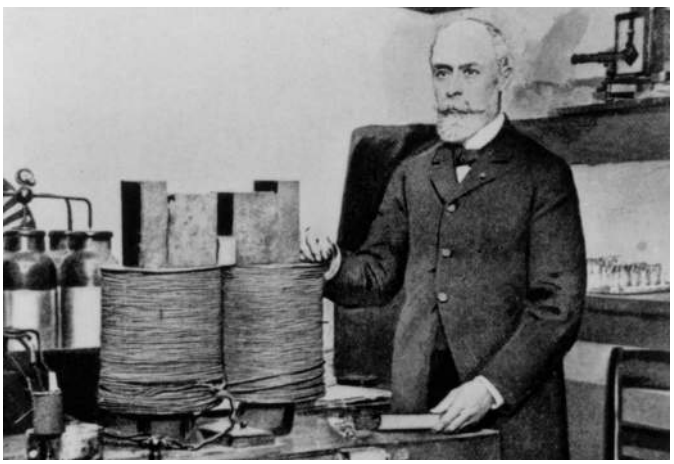


Figure 16.12 Henri Becquerel, the discoverer of radioactivity, in his laboratory. His father and grandfather had been professors of physics in Paris before him.

Radiation from radioactive substances

There are three types of radiation which are emitted by radioactive substances: alpha (α), beta (β) and gamma (γ) radiations come from the unstable nuclei of atoms. Nuclei consist of protons and neutrons, and if the balance between these two types of particles is too far to one side, the nucleus may emit α- or β-radiation as a way of achieving greater stability. Gamma-radiation is usually emitted after α or β decay, to release excess energy from the nuclei.

In fact, there are two types of β-radiation. The more familiar is beta-minus (β^-) radiation, which is simply an electron, with negative charge of $-e$. However, there are also many unstable nuclei that emit beta-plus (β^+) radiation. This radiation is in the form of **positrons**, similar to electrons in terms of mass but with positive charge of $+e$. Positrons are a form of **antimatter**. When a positron collides with an electron, they annihilate each other. Their mass is converted into electromagnetic energy in the form of two gamma photons (Figure 16.13).

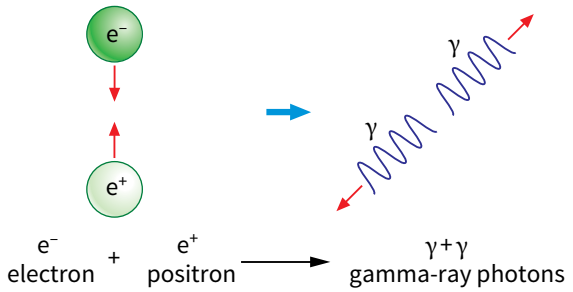


Figure 16.13 Energy is released in the annihilation of matter and antimatter.

Table 16.4 shows the basic characteristics of the different types of radiation. The masses are given relative to the mass of a proton; charge is measured in units of e, the elementary charge.

Radiation	Symbol	Mass (relative to proton)	Charge	Typical speed
α-particle	α, $^4_2$He	4	$+2e$	'slow' ($10^6\,m\,s^{-1}$)
β^--particle	β, β^-, e, $^0_{-1}$e	$\frac{1}{1840}$	$-e$	'fast' ($10^8\,m\,s^{-1}$)
β^+-particle	β, β^+, e+, $^0_{+1}$e	$\frac{1}{1840}$	$+e$	'fast' ($10^8\,m\,s^{-1}$)
γ-ray	γ	0	0	speed of light ($3 \times 10^8\,m\,s^{-1}$)

Table 16.4 The basic characteristics of ionising radiations.

Note the following points:

- α- and β-radiation are particles of matter. A γ-ray is a photon of electromagnetic radiation, similar to an X-ray. (X-rays are produced when electrons are decelerated; γ-rays are produced in nuclear reactions.)
- An α-particle consists of two protons and two neutrons; it is a nucleus of helium-4. A β^--particle is simply an electron and a β^+-particle is a positron.
- The mass of an α-particle is nearly 10 000 times that of an electron and it travels at roughly one-hundredth of the speed of an electron.

Discovering neutrinos

There is a further type of particle which we need to consider. These are the **neutrinos**. When β decay was first studied, it was realised that β-particles were electrons coming from the **nucleus** of an atom. There are no electrons in the nucleus (they 'orbit' outside the nucleus), so the process was pictured as the decay of a neutron to give a proton and an electron.

It was noticed that β-particles were emitted with a **range** of speeds – some travelled more slowly than others. It was deduced that some other particle must be carrying off some of the energy and momentum released in the decay. This particle is now known as the **antineutrino** (or, more correctly, the electron antineutrino), with symbol $\bar{\nu}$. The decay equation for β^- decay is written as:

beta-minus (β^-) decay: $^{1}_{0}\text{n} \rightarrow {}^{1}_{1}\text{p} + {}^{0}_{-1}\text{e} + \bar{\nu}$

Neutrinos are bizarre particles. They have very little mass (much less than an electron) and no electric charge, which makes them very difficult to detect. The Austrian physicist Wolfgang Pauli predicted their existence in 1930, long before they were first detected in 1956.

In β^+ decay, a proton decays to become a neutron and an **electron neutrino** (symbol ν) is released:

beta-plus (β^+) decay: $^{1}_{1}\text{p} \rightarrow {}^{1}_{0}\text{n} + {}^{0}_{+1}\text{e} + \nu$

The two equations highlighted above show two important features of radioactive decay. Firstly, nucleon number A is conserved; that is, there are as many nucleons after the decay as there were before. In β^- decay, a neutron has become a proton so that the total number of nucleons is unchanged. In β^+ decay, a proton becomes a neutron, so again A is conserved.

Secondly, proton number Z is also conserved. In β^- decay, we start with a neutron ($Z = 0$). After the decay, we have a proton ($Z = +1$) and a β^- particle ($Z = -1$). Together these have $Z = 1 - 1 = 0$. Since Z tells us about the charge of each particle, we would be surprised if we had a different amount of charge after the decay than before the decay. A similar analysis shows that Z is conserved in β^+ decay.

Do these conservation laws apply to α decay? Here is an equation that represents a typical α decay:

$$^{222}_{86}\text{Rn} \rightarrow {}^{218}_{84}\text{Po} + {}^{4}_{2}\text{He}$$

In α decay, an alpha particle (two protons and two neutrons) is emitted by a nucleus. Although these nucleons are now outside the nucleus, the equation shows that there is the same number of nucleons after the decay (218 + 4) as before the decay (222). So nucleon number A is conserved. Similarly, proton number Z is conserved (84 + 2 = 86).

The conservation of nucleon number and proton number are important laws in nuclear physics. They apply to all nuclear changes, not just to α and β decay.

There is a third quantity that is conserved. You might expect mass to be conserved, but this is not so. For example, in the α decay equation given above, the combined mass of the polonium nucleus and the alpha particle is slightly less than that of the original radon nucleus. The 'lost' mass has become energy – this is where the fast-moving alpha particle gets its kinetic energy. The relationship between mass m and energy E is given by Einstein's equation $E = mc^2$, where c is the speed of light in free space. So, instead of saying that mass is conserved in nuclear processes, we have to say that mass–energy is conserved. There is much more about this in Chapter 31.

Fundamental families

Electrons and neutrinos both belong to the family of fundamental particles called **leptons**. These are particles that do not feel the strong nuclear force. Recall from page 230 that particles that experience the strong force are hadrons, and that these are made up of fundamental particles called quarks.

So we have two families of fundamental particles, quarks and leptons. How can we understand β decay in terms of these particles?

Consider first β^- decay, in which a neutron decays. A neutron consists of three quarks (up, down, down or u d d). It decays to become a proton (u u d). Comparing these shows that one of the down quarks has become an up quark. In the process, it emits a β-particle and an antineutrino:

$$\text{d} \rightarrow \text{u} + {}^{0}_{-1}\text{e} + \bar{\nu}$$

In β^+ decay, a proton decays to become a neutron. In this case, an up quark becomes a down quark:

$$\text{u} \rightarrow \text{d} + {}^{0}_{+1}\text{e} + \nu$$

Fundamental forces

The nucleus is held together by the strong nuclear force, acting against the repulsive electrostatic or Coulomb force between protons. This force explains α decay, when a positively charged α-particle flies out of the nucleus, leaving it with less positive charge.

However, the strong force cannot explain β decay. Instead, we have to take account of a further force within the nucleus, the **weak interaction**, also known as the weak nuclear force. This is a force that acts on both quarks and leptons. The weak interaction is responsible for β decay.

QUESTIONS

10 The equation ${}^{1}_{1}\text{p} \rightarrow {}^{1}_{0}\text{n} + {}^{0}_{+1}\text{e} + \nu$ represents β^+ decay. Use the equation to explain why the neutrino ν can have no charge and very little mass.

11 What are the differences between a proton, a positron and a photon? You can describe how their masses differ, how their charges differ, or whether they are particles or antiparticles.

12 State the names of:

a all the hadrons that are mentioned in this chapter

b all the leptons that are mentioned in this chapter.

13 State **two** differences between hadrons and leptons.

Properties of ionising radiation

Radiation affects the matter it passes through by causing ionisation. Both α- and β-particles are fast-moving charged particles, and if they collide with or pass close to atoms, they may knock or drag electrons away from the atoms (Figure 16.14). The resulting atoms are said to be **ionised**, and the process is called ionisation. In the process, the radiation loses some of its kinetic energy. After many ionisations, the radiation loses all of its energy and no longer has any ionising effect.

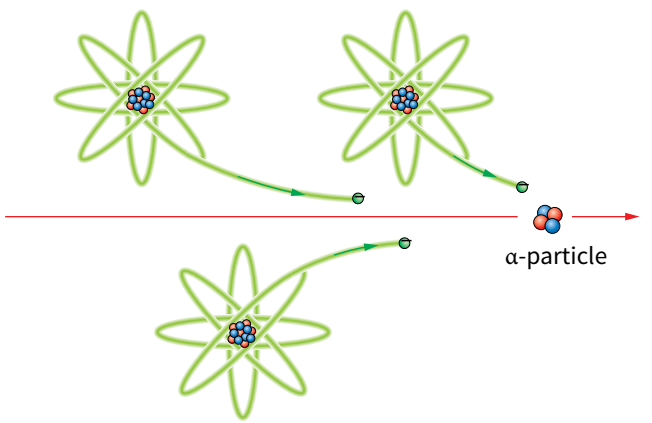


Figure 16.14 As an α-particle passes through a material, it causes ionisation of atoms.

Alpha-radiation is the most strongly ionising, because the mass and charge of an α-particle are greater than those of a β-particle, and it usually travels more slowly. This means that an α-particle interacts more strongly with any atom that it passes, and so it is more likely to cause ionisation. Beta-particles are much lighter and faster, and so their effect is smaller. Gamma-radiation also causes ionisation, but not as strongly as α- and β-particles, as γ-rays are not charged.

QUESTION

14 **a** Explain why you would expect β^--particles to travel further through air than α-particles.

b Explain why you would expect β^--particles to travel further through air than through metal.

Behaviour of radiations in electric and magnetic fields

Because α-, β^-- and γ-radiations have different charges, or no charge, they behave differently in electric and magnetic fields. This can be used to distinguish one kind of radiation from another.

Figure 16.15 shows the effect of an electric field. A mixture of α-, β^-- and γ-radiations is passing through the gap between two parallel plates; the electric field in this space is uniform (Chapter 8). Since α- and β^--particles are charged, they are attracted to the plate that has the opposite charge to their own. β^--particles are deflected more than α-particles, since their mass is so much less. Gamma-rays are undeflected since they are uncharged.

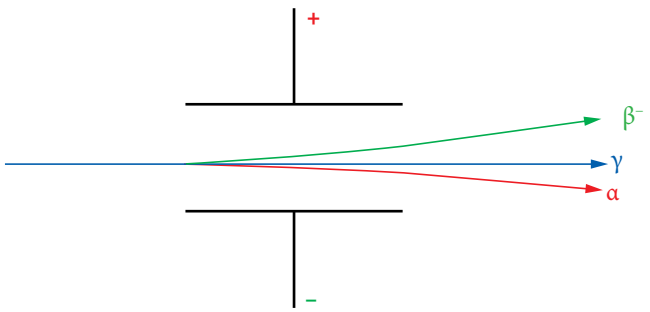


Figure 16.15 An electric field can be used to separate α-, β^-- and γ- radiations. (The deflection of the α-radiation has been greatly exaggerated here.)

Figure 16.16 shows the effect of a magnetic field. In this case, the deflecting force on the particles is at right angles to their motion.

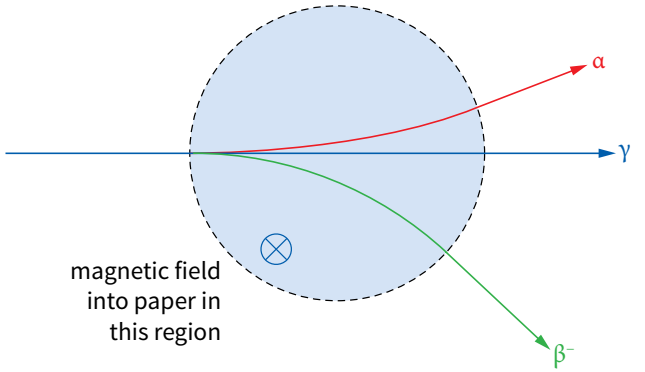


Figure 16.16 A magnetic field may also be used to separate α-, β^-- and γ- radiations. The deflection of the α-radiation has been greatly exaggerated here.

QUESTION

15 **a** Some radioactive substances emit α-particles having two different speeds. Draw a diagram similar to Figure 16.16 to show how these particles would move in a uniform electric field. Label your diagram to show the tracks of the faster and slower α-particles.

b A β^--emitting radioactive substance emits β^--particles with a range of speeds. Add to the diagram you drew in **a** to show how these particles would behave in a uniform electric field.

Radiation penetration

Safety note

When working with radioactive sources, it is essential to follow the relevant safety regulations, which your teacher will explain to you.

Alpha-radiation

Because α-radiation is highly ionising, it cannot penetrate very far into matter. A cloud chamber can be used to show the tracks of α-particles in air (Figure 16.17). The tracks are very dense, because of the dense concentration of ions produced, and they extend for only a few centimetres into the air. By the time the α-particles have travelled this far, they have lost virtually all of their kinetic energy. The α-particle, which is a nucleus of helium-4, grabs two drifting electrons in the air and becomes a neutral atom of helium gas.

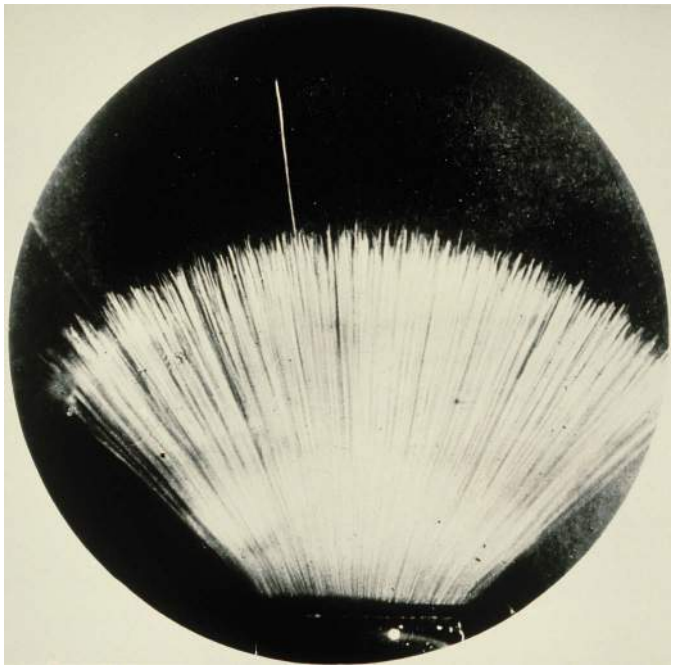


Figure 16.17 Alpha-particle tracks show up in this photograph of a cloud chamber. Notice that all the particles travel roughly the same distance through the air, indicating that they all have roughly the same initial kinetic energy.

Alpha-particles can also be detected by a solid-state detector, or by a Geiger–Müller (GM) tube with a thin end-window (Figure 16.18) connected to an electronic counter. By moving the source back and forth in front of the detector, it is simple to show that the particles only penetrate 5 or 6 cm of air. Similarly, with the source close to the detector, it can be shown that a single sheet of paper is adequate to absorb all of the α-radiation.

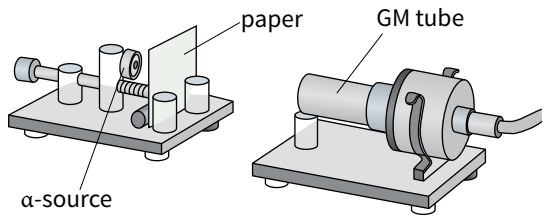


Figure 16.18 Alpha-radiation can be absorbed by a single sheet of paper.

Beta-radiation

A Geiger–Müller tube can detect β^--radiation. The source is placed close to the tube, and different materials are positioned between source and tube. Paper has little effect; a denser material such as aluminium or lead is a more effective absorber. A few millimetres of aluminium will almost completely absorb β^--radiation.

Gamma-radiation

Since γ-radiation is the least strongly ionising, it is the most penetrating. Lead can be used to absorb γ-rays. The intensity of the radiation decreases gradually as it passes through the lead. In principle, an infinite thickness of lead would be needed to absorb the radiation completely; in practice, a couple of centimetres of lead will reduce the intensity by half and 10 cm will reduce the intensity to a safe level in most situations.

The different penetrating properties of α-, β^-- and γ-radiations can be summarised as follows:

- α-radiation is absorbed by a thin sheet of paper or a few centimetres of air.
- β^--radiation is absorbed by a few millimetres of metal.
- γ-radiation is never completely absorbed but a few centimetres of lead, or several metres of concrete, greatly reduces the intensity.

This is illustrated in Figure 16.19.

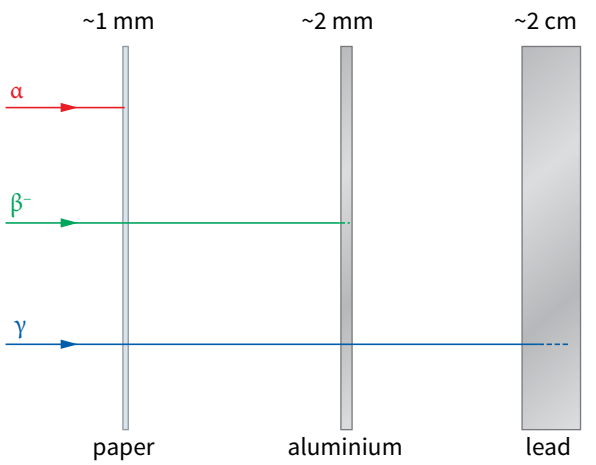


Figure 16.19 A summary of the penetrating powers of α-, β^-- and γ-radiations. The approximate thickness of the absorbing material is also shown.

The electronvolt (eV)

Alpha and beta particles move quickly; gamma photons travel at the speed of light. These types of radiation all have energy, but the energy of a single particle or photon is very small and far less than a joule. So we use another, much smaller unit of energy, the electronvolt, when considering the energy of individual particles or photons.

When an electron (with a charge of magnitude 1.60×10^{-19} C) travels through a potential difference, energy is transferred. The energy change W is given by:

$$W = QV = 1.60 \times 10^{-19} \times 1 = 1.60 \times 10^{-19}\,\text{J}$$

So we define the electronvolt as follows:

> One electronvolt (1 eV) is the energy transferred when an electron travels through a potential difference of one volt.

Therefore:

$$1\ \text{eV} = 1.60 \times 10^{-19}\,\text{J}$$

There is more about the electronvolt and its use in energy calculations in Chapter 30.

QUESTIONS

16 Explain why the most strongly ionising radiation (α-particles) is the least penetrating, while the least ionising (γ-rays) is the most penetrating.

17 A smoke detector (Figure 16.20) uses a source of α-radiation to detect the presence of smoke in the air. Find out how the smoke detector works and suggest why an α source is more suitable for this than a β^- or γ source.

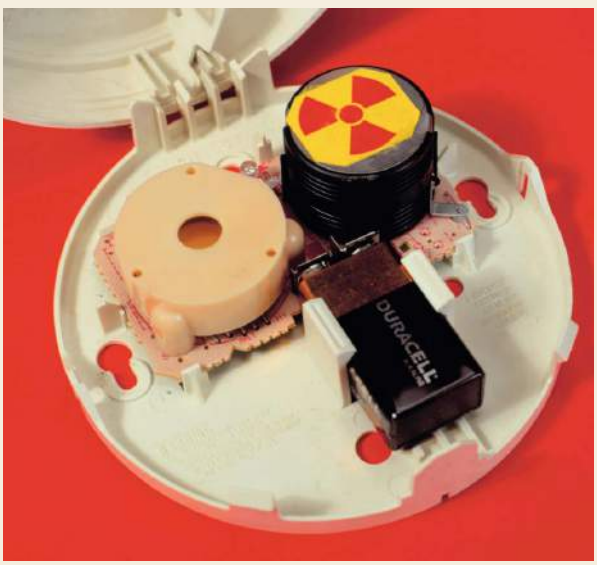


Figure 16.20 A smoke detector that uses the absorption of α-radiation as the principle of its operation.

Summary

- The α-particle scattering experiment provides evidence for the existence of a small, massive and positively charged nucleus at the centre of the atom.
- Most of the mass of an atom is concentrated in its nucleus.
- The nucleus consists of protons and neutrons, and is surrounded by a cloud of electrons.
- The number of protons and neutrons in the nucleus of an atom is called its nucleon number *A*.
- The number of protons in the nucleus of an atom is called its proton number (or atomic number) *Z*.
- Hadrons (e.g. the neutron) are particles that consist of quarks and hence are affected by the strong nuclear force. Leptons (e.g. the electron) are particles that are unaffected by the strong nuclear force.
- The weak interaction acts between quarks and is responsible for β-decay.
- Isotopes are nuclei of the same element with a different number of neutrons but the same number of protons.
- Different isotopes (or nuclides, if referring to the nucleus only) can be represented by the notation ${}^{A}_{Z}\mathrm{X}$, where X is the chemical symbol for the element.
- There are three types of ionising radiation produced by radioactive substances: α-particles, β-particles and γ-rays.
- In radioactive decay, the following quantities are conserved: proton number, nucleon number and mass–energy.
- The most strongly ionising, and hence the least penetrating, is α-radiation. The least strongly ionising is γ-radiation.
- Because of their different charges, masses and speeds, the different types of radiation can be identified by the effect of an electric or magnetic field.

End-of-chapter questions

1 Before Rutherford's model, scientists believed that the atom was made up of negatively charged electrons embedded in a 'plum pudding' of positive charge that was spread throughout the atom. Explain how the α-particle scattering experiment proved that this old model of the atom was incorrect. [3]

2 A nucleus of strontium has a nucleon number of 90 and a proton number of 38. Describe the structure of this strontium nucleus. [1]

3 State the changes that take place in a nucleus when it emits an α-particle and then two β^--particles. [5]

4 The nuclide of iodine with a nucleon number of 131 and a proton number 53 emits a β^--particle. Write a nuclear equation for this decay. [3]

5 An isotope of carbon ${}^{14}_{6}\mathrm{C}$ emits a β^--particle and changes into an isotope of nitrogen (N).

a What are β^--particles? [1]

b Write a nuclear decay equation for the decay. [2]

c Draw a graph with the *y*-axis representing nucleon numbers between 10 and 16 and the *x*-axis representing proton numbers between 4 and 10. On your graph mark:

i the isotope ${}^{14}_{6}\mathrm{C}$ [2]

ii the daughter nucleus produced in the decay. [1]

6 The uranium isotopes U-236 and U-237 both emit radioactive particles. A nucleus of uranium-237 may be written as $^{237}_{92}U$ and emits a β^--particle. A nucleus of uranium-236 emits an α-particle. The number of protons in a nucleus of uranium is 92.

a Describe the differences between an α-particle and a β^--particle. [4]

b Explain how uranium can exist in a number of different isotopes. [2]

c Write down the nuclear equation for the decay of U-236. [2]

7 Approximate values for the radius of a gold atom and the radius of a gold nucleus are 10^{-10} m and 10^{-15} m, respectively.

a Estimate the ratio of the volume of a gold atom to the volume of a gold nucleus. [2]

b The density of gold is 19 000 kg m^{-3}. Estimate the density of a gold nucleus, stating any assumptions that you make in your answer. [3]

8 The nuclide of lead $^{210}_{82}Pb$ decays in three separate stages by α- and β^- emission to another lead nuclide, $^{206}_{82}Pb$.

a Describe the structure of a nucleus of $^{210}_{82}Pb$. [2]

b α- and β^--particles are known as **ionising radiations**. State and explain why such radiations can be described as **ionising**. [2]

c The two lead nuclides are shown in the graph in Figure 16.22 of nucleon number *A* against proton number *Z*.

Copy the graph and on your copy draw **three** arrows to represent one possible route for the three decays between the two isotopes of lead. Label each arrow to show whether an α-particle or a β^--particle is emitted. [3]

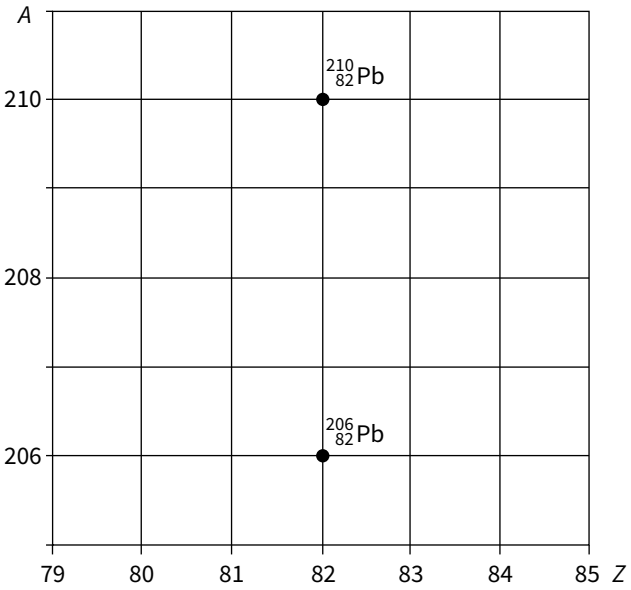


Figure 16.22 For End-of-chapter Question 8.

9 Geiger and Marsden carried out an experiment to investigate the structure of the atom. In this experiment α-particles were scattered by a thin film of gold.

a When Rutherford analysed their results, what conclusions did he draw about the distribution of mass and charge in the atom? [2]

b Describe and explain the experimental observations that led to these conclusions. [3]

10 Beta decay occurs as either β^+-decay or β^--decay. An isotope of calcium $^{46}_{20}Ca$ decays by β-emission into the isotope $^{46}_{21}Sc$, and an isotope of magnesium $^{23}_{12}Mg$ decays by β-emission into the isotope $^{23}_{11}Na$.

a Copy and complete the following decay equations for the calcium and magnesium isotopes.

i decay of calcium: $^{46}_{20}Ca \rightarrow {}^{\cdots}_{\cdots}e + {}^{\cdots}_{\cdots}Sc^+ + \ldots$

ii decay of magnesium: $^{23}_{12}Mg \rightarrow {}^{\cdots}_{\cdots}e + {}^{\cdots}_{\cdots}Na^+ + \ldots$ [3]

b State what happens in each type of β-decay in terms of the quark model of nucleons.

i β^--decay [1]

ii β^+-decay. [1]

c Name the force responsible for β-decay. [1]

11 **a** A quark is a fundamental particle but a neutron is not. Explain what this statement means.

b A proton and a neutron each contain three quarks, either up or down quarks.

i Copy and complete the table to show the charge on a proton and a neutron and the quarks that they contain.

	Charge	**Quarks**
proton		
neutron		

[2]

ii Using information from your table, suggest why some quarks must have a positive charge and some quarks a negative charge. [2]

c State what interaction is responsible for holding the nucleus together. [1]

d When a neutron decays it produces an electron and two other particles. Copy and complete the decay equation for a neutron.

$^1_0n \rightarrow$ ..... . [2]

e The electron and the neutron belong to different groups of particles.

Copy and complete the table to show the group of particles to which the electron and neutron belong and state the name of another member of each group.

	Group to which it belongs	**Another particle in the same group**
electron		
neutron		

[2]

P1: Practical skills at AS level

Learning outcomes

You should be able to:

- understand random, systematic and zero errors
- calculate uncertainties in measurements made with a range of instruments
- produce best fit straight-line graphs and obtain the intercept and gradient
- use readings to test a relationship
- suggest problems and improvements in a range of experiments

Practical work in physics

Throughout your A level physics course you will develop your skills in practical work, and they will be assessed at both AS and A level. This chapter outlines the skills you will develop in the first year of the course; it includes some activities to test your understanding as you go along.

The sciences differ from most other subjects in that they involve not only theory but also practical work. The very essence of science is that theory can be tested by practical experiment. So the ability to carry out practical exercises in a logical and scientific manner is essential.

Using apparatus and following instructions

You need to familiarise yourself with the use of simple measuring instruments such as metre rules, balances, protractors, stopwatches, ammeters and voltmeters, and even more complicated ones such as a micrometer screw gauge and calipers.

When using measuring instruments like these you need to ensure that you are fully aware of what each division on a scale represents. If you look at Figure P1.1 you will see that on the first ruler each division is 1 mm, and on the second each division is 2 mm.

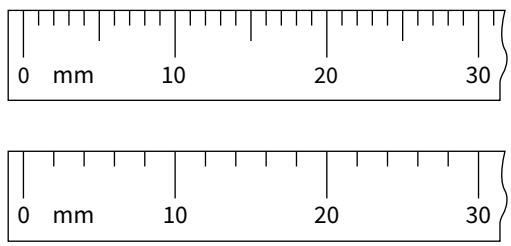


Figure P1.1 When reading from a scale, make sure that you know what each division on the scale represents.

If you use instruments incorrectly, you may introduce errors into your readings. For example, when taking a reading your line of sight should always be perpendicular to the scale that you are using. Otherwise you will introduce a parallax error; this is shown in Figure P1.2. Looking from point A the length of the rod appears to be 21 mm, from point C it appears to be 25 mm and from point B, the correct position, the length is 23 mm.

A rule, for example a metre rule, or a ruler, for example an ordinary school ruler of length 30 cm, are simple measuring instruments with a smallest division of 1 mm. Other instruments have a greater precision because their smallest scale division is less than 1 mm. Here we will look at two of these.

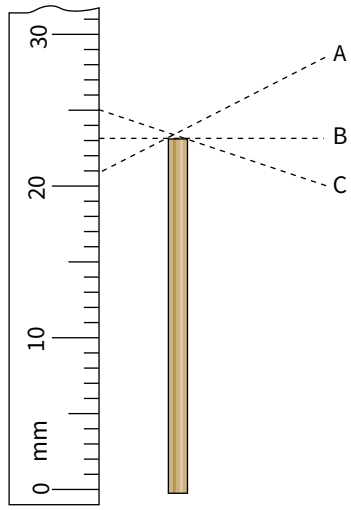


Figure P1.2 Parallax error.

Calipers

Calipers are designed to grip an object with two jaws and, in the example shown in Figure P1.3, to measure the diameter of the object. They can also be used to measure the internal diameter of a tube, for example, if the two prongs are placed inside the tube and the moving part of the calipers is adjusted until the prongs just grip the inside of the tube.

The calipers shown in Figure P1.3 are dial calipers, although other versions such as vernier calipers are still sometimes used. As the sliding scale moves along, one rotation of the dial moves the jaws 1 mm further apart. Since the dial shown has 100 divisions, each of these divisions is $\frac{1}{100} = 0.01$ mm. The object shown has a diameter of 12 mm on the fixed scale and 25 divisions or 0.25 mm on the dial, so the diameter of the object is 12.25 mm.

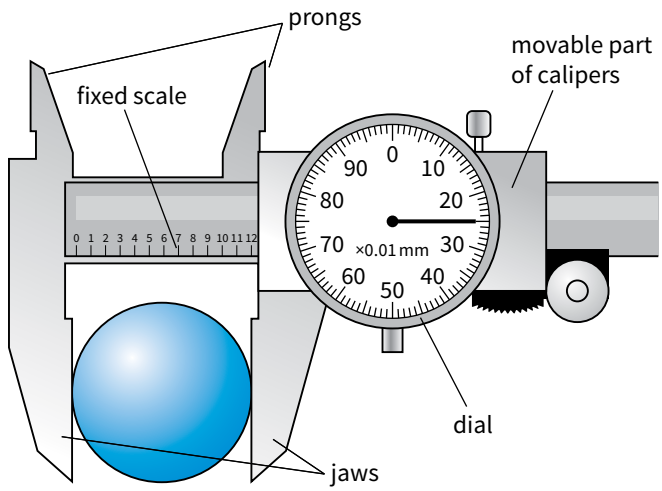


Figure P1.3 Using dial calipers.

Micrometer screw gauge

A micrometer screw gauge, or more simply a micrometer, is shown in Figure P1.4. This has also has two scales. The main scale is on the shaft and the fractional scale is on the rotating barrel. One rotation of the barrel moves the end of the barrel 0.50 mm along the shaft. The barrel has 50 divisions so each division represents $\frac{0.50}{50} = 0.01$ mm.

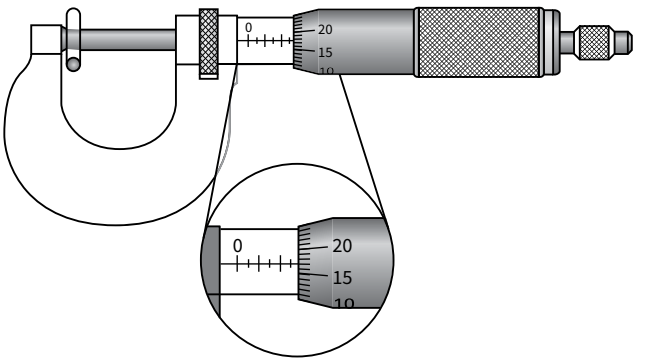


Figure P1.4 Using a micrometer screw gauge.

To use the micrometer, turn the barrel until the jaws just tighten on the object. Some micrometers have a ratchet or slip mechanism to prevent the user from tightening too hard and damaging the micrometer or object. Read the main scale to the nearest 0.5 mm, then read the number of divisions on the sleeve, which will be in 0.01 mm, and finally add the two readings. You should realise that the smallest division on the micrometer is 0.01 mm.

Before you start to use a micrometer or dial calipers it is usual to check if there is a zero error. This is done by bringing the jaws together without any object between them. Obviously the reading should be zero, but if the instrument is worn or has been used badly the reading may not be zero. When you have taken this zero error reading it should be added or subtracted to every other reading that you take with the instrument. If the jaws do not quite close to the zero mark there is a positive zero error, and this zero error reading should be subtracted. The zero error is an example of a systematic error, which is dealt with later in this chapter.

It is also important that you become familiar with setting up apparatus. When instructions are given, the only way to become confident is through practice. You may face a variety of tasks, from setting up a pendulum system to measuring the angle at which a tilted bottle falls.

You should also learn to set up simple circuits from circuit diagrams. The most common error in building circuits comes where components need to be connected in parallel. A good piece of advice here is to build the main circuit first, and then add the components which need to be connected in parallel.

Gathering evidence

When gathering evidence you should take into account the range of results that you are going to obtain. If you are investigating the extension of a spring with load, for loads of between 0 N and 20 N, you should take a fair spread of readings throughout that range. For instance, six readings between 12 N and 20 N would not be sensible because you are not investigating what happens with smaller loads. Equally, taking three readings below 5 N and three more between 15 N and 20 N does not test what happens with intermediate loads.

A sensible set of readings might be at 0 N, 4 N, 8 N, 12 N, 16 N and 20 N. This covers the whole range in equal steps.

QUESTION

1 You are investigating how the current through a resistor depends on its resistance when connected in a circuit. You are given resistors of the following values:

50 Ω, 100 Ω, 150 Ω, 200 Ω, 250 Ω, 300 Ω, 350 Ω, 400 Ω, 450 Ω, 500 Ω

You are asked to take measurements with just six of these resistors. Which six resistors would you choose? Explain your choice.

Precision, accuracy, errors and uncertainties

Whenever you make a measurement, you are trying to find the true value of a quantity. This is the value you would find if your measurement was perfect. However, no measurement can ever be perfect; there will always be some uncertainty. Your equipment may be imperfect or your technique may be capable of improvement. So, whenever you carry out practical work, you should think about two things:

- how the equipment or your technique could be improved to give better results, with less uncertainty
- how to present the uncertainty in your findings.

As you will see later in this chapter, both of these need to be reflected in the way you present your findings.

We will first consider the **precision** of a measurement. The level of precision is high if you make several measurements of a quantity and they are all very similar. A precise measurement, when repeated, will be the same, or nearly so. However, if your measurements are spread widely around the average, they are less precise. This can arise because of practical difficulties in making the measurements.

Precision is reflected in how the results are recorded. If a distance is quoted as '15 m' then it implies that it was only measured to the nearest metre, whereas if it is quoted as '15.0 m' then it suggests that it was measured to the nearest 0.1 m.

Take care not to confuse precision with **accuracy**. A measurement is described as 'accurate' if the value obtained is close to the true value. Even if a measurement is precise, and always produces the same result, it may not be accurate because every reading may have the same error. For example, you can make very precise measurements of the diameter of a wire using a micrometer screw gauge to the nearest 0.01 mm, but every reading may be inaccurate if the gauge has a zero error.

Figure P1.5 shows two attempts at making holes in the centre of a target. Imagine that the positions of the holes represent readings, with the true value at the centre. On the left, the readings are close together so we can say that they are precise. However, they are not accurate as the average is far from the centre. In the second, the measurement can be said to be accurate as the average position of the holes is close to the centre, but the readings are not precise as the holes are spread out.

Whenever you make a measurement you should be aware of the uncertainty in the measurement. It will often, but not always, be determined by the smallest division on the measuring instrument. On a metre ruler which is graduated in millimetres we should be able to read to the nearest half millimetre, but beware! If we are measuring the length of a rod there are two readings to be taken, one at each end of the rod. Each of these readings has an uncertainty of 0.5 mm, giving a total uncertainty of 1 mm.

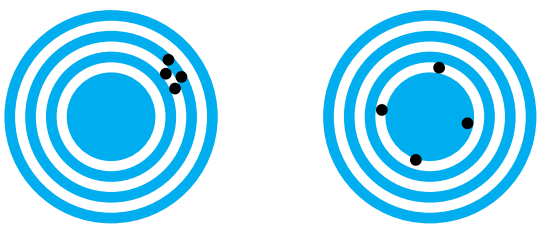


Figure P1.5 The left-hand diagram represents readings which are precise but not accurate; the right-hand diagram represents readings which are accurate but without precision.

The uncertainty will depend not only on the precision of the calibrations on the instrument you are using, but also on your ability to observe and on **errors** introduced by less than perfect equipment or poor technique in taking the observations. Here are some examples of where uncertainties might arise:

- **Systematic error** – A spring on a force meter might, over time, become weaker so that the force meter reads consistently high. Similarly the magnet in an ammeter might, over the years, become weaker and the needle may not move quite as far round the scale as might be expected. Parallax errors, described earlier, may be another example of a systematic error if one always looks from the same angle, and not directly from above, when taking a measurement. In principle, systematic errors can be corrected for by recalibrating the instrument or by correcting the technique being used.
- **Zero error** – The zero on a ruler might not be at the very beginning of the ruler. This will introduce a fixed error into any reading unless it is allowed for. This is a type of systematic error.
- **Random errors** – When a judgement has to be made by the observer, a measurement will sometimes be above and sometimes below the true value. Random errors can be reduced by making multiple measurements and averaging the results.

Good equipment and good technique will reduce the uncertainties introduced, but difficulties and judgements in making observations will limit the precision of your measurements. Here are two examples of how difficulties in observation will determine the uncertainty in your measurement.

Example 1: Using a stopwatch

Tambo has a digital stopwatch which measures to the nearest one-hundredth of a second. He is timing Nana in a 100 metre race (Figure P1.6). He shows her the stopwatch, which reads 11.87 s. She records in her notebook the time 11.9 s. She explains to Tambo that he cannot possibly



Figure P1.6 Uncertainty in timing using a stopwatch.

measure to the nearest one-hundredth of a second as he has to judge both when the starting pistol was fired and the exact moment at which she crossed the finishing line. To do this to any closer than the nearest one-tenth of a second is impossible. In addition, sometimes he will press the button too early and sometimes too late.

Example 2: Measuring displacement of a pendulum

Fatima is asked to measure the maximum displacement of a pendulum bob as it oscillates, as shown in Figure P1.7. She uses a ruler calibrated in millimetres. She argues that she can measure the displacement to the nearest millimetre. Joanne, however, correctly argues that she can only measure it to the nearest two millimetres, as not only is there the uncertainty at either end (0.5 mm) but she also has to judge precisely the point at which the bob is at its greatest displacement, which adds an extra millimetre to the uncertainty.

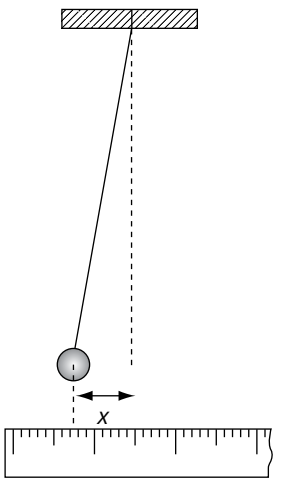


Figure P1.7 Displacement of a pendulum bob.

QUESTIONS

2 Look at Figure P1.5 on page 242. Draw similar diagrams to represent:

- **a** a target where the holes are both precise and accurate
- **b** a target where the holes are neither precise nor accurate.

3 The position of the holes in Figure P1.5 on page 242 represents attempts at measuring the position of the centre of the circle. Which one shows more random error and which shows more systematic error?

Finding the value of an uncertainty

We have used the terms uncertainty and error; they are not quite the same thing. In general an 'error' is just a problem which causes the reading to be different from the true value (although a zero error can have an actual value). The uncertainty, however, is an actual range of values around a measurement, within which you expect the true value to lie. The uncertainty is an actual number with a unit.

For example, if you happen to know that the true value of a length is 21.0 cm and an 'error' or problem causes the actual reading to be 21.5 cm, then, since the true value is 0.5 cm away from the measurement, the uncertainty is ±0.5 cm.

But how do you estimate the uncertainty in your reading without knowing the true value? Obviously if a reading is 21.5 cm and you know the true value is 21.0 cm, then the uncertainty in the reading is 0.5 cm. However, you may still have to estimate the uncertainty in your reading without knowing the true value. So how is this done?

First, it should be understood that the uncertainty is only an estimate of the difference between the actual reading and the true value. We should not feel too worried if the difference between a single measurement and the true value is as much as twice the uncertainty. Because it is an estimate, the uncertainty is likely to be given to only one significant figure. For example, we write the uncertainty as 0.5 cm and not 0.50 cm.

The uncertainty can be estimated in two ways.

Using the division on the scale – Look at the smallest division on the scale used for the reading. You then have to decide whether you can read the scale to better than this smallest division. For example, what is the uncertainty in the level of point B in Figure P1.2? The smallest division on the scale is 1 mm but is it possible to measure to better than 1 mm? This will depend on the instrument being used and whether the scale itself is accurate. In Figure P1.2, the width of the line itself is quite small but there may be some parallax error which would lead you to think that 0.5 mm or 1 mm is a reasonable uncertainty. In general, the position of a mark on a ruler can generally be measured to an uncertainty of ±0.5 mm. In Figure P1.8, the smallest division on the scale is 20 g. Can you read more accurately than this? In this case it is doubtful that every marking on the scale is accurate and so 20 g would be reasonable as the uncertainty.

You need to think carefully about the smallest division you can read on any scale. As another example, look at a protractor. The smallest division is probably 1° but it is unlikely you can use a protractor to measure an angle to better than ±0.5° with your eye.

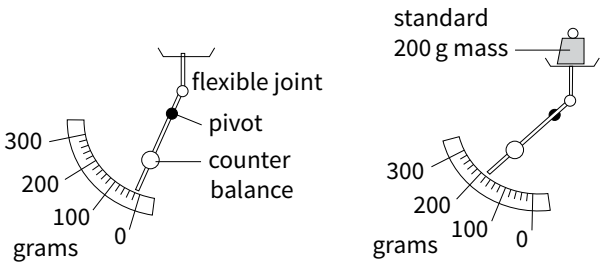


Figure P1.8 The scales on a lever-arm balance.

Repeating the readings – Repeat the reading several times. The uncertainty can then be taken as half of the range of the values obtained, in other words the smallest reading is subtracted from the largest and the result is halved. This method deals with random errors made in the readings but does not account for systematic errors. This method should always be tried, wherever possible, because it may reveal random errors and gives an easy way to estimate the uncertainty. However, if the repeated readings are all the same, do not think that the uncertainty is zero. The uncertainty can never be less than the value you obtained by looking at the smallest scale division.

Which method should you actually use to estimate the uncertainty? If possible, readings should be repeated and the second method used. But if all the readings are the same, you have to try both methods!

The uncertainty in using a stopwatch is something of a special case as you may not be able to repeat the measurement. Usually the smallest division on a stopwatch is 0.01 s, so can you measure a time interval with this uncertainty? You may know that your own reaction time is larger than this and is likely to be at least 0.1 s. The stopwatch is recording the time when you press the switch but this is not pressed at exactly the correct moment. If you do not repeat the reading then the uncertainty is likely to be at least 0.1 s, as shown in Figure P1.7. If several people take the reading at the same time you are likely to see that 0.01 s is far too small to be the uncertainty.

Even using a digital meter is not without difficulties. For example if a digital ammeter reads 0.35 A then, without any more information, the uncertainty is ± 0.01 A, the smallest digit on the meter. But if you look at the handbook for the ammeter you may well find that the uncertainty is ± 0.02 or 0.03 A (although you cannot be expected to know this).

QUESTIONS

4 Figure P1.8 shows a lever-arm balance, initially with no mass in the pan and then with a standard 200 g mass in the pan.

Explain what types of errors might arise in using this equipment.

5 Estimate the uncertainty when a student measures the length of a room using a steel tape measure calibrated in millimetres.

6 Estimate the uncertainty when a girl measures the temperature of a bath of water using the thermometer in Figure P1.9.

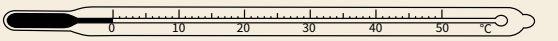


Figure P1.9 For Question 6.

7 A student is asked to measure the wavelength of waves on a ripple tank using a metre rule which is graduated in millimetres. Estimate the uncertainty in his measurement.

8 Estimate the uncertainty when a student attempts to measure the time for a single swing of a pendulum.

9 What is the average value and uncertainty in the following sets of readings? All are quoted to be consistent with the smallest scale division used.

a 20.6, 20.8

b 20, 30, 36

c 0.6, 1.0, 0.8, 1.2

d 20.5, 20.5

WORKED EXAMPLE

1 A length is measured five times with a ruler whose smallest division is 0.1 cm and the readings obtained, in cm, are 22.9, 22.7, 22.9, 23.0, 23.1. What is the reading obtained and the uncertainty?

Step 1 Find the average by adding the values and dividing by the number of values:

$$\frac{22.9 + 22.7 + 22.9 + 23.0 + 23.1}{5} = 22.92\,\text{cm}$$

This is written to 4 significant figures. At this stage you are not sure how many figures to write in the answer.

Step 2 The maximum value is 23.1 and the minimum value is 22.7. Use these values to find half the range.

$$\text{half the range} = \frac{23.1 - 22.7}{2} = 0.2\,\text{cm}$$

Step 3 Check that the uncertainty calculated in Step 2 is larger than the smallest division you can read on the scale.

Step 4 Write down the average value, the uncertainty to a reasonable number of significant figures and the unit. Obviously the last digit in 22.92 is meaningless as it is much smaller than the uncertainty; it should not be written down.

the final value is 22.9 ± 0.2 cm

You do not usually write down the final value of the answer to a greater number of decimal places than the uncertainty. Uncertainties are usually quoted to 1 or perhaps 2 significant figures.

Percentage uncertainty

The uncertainties we have found so far are sometimes called absolute uncertainties, but percentage uncertainties are also very useful.

The percentage uncertainty expresses the absolute uncertainty as a fraction of the measured value and is found by dividing the uncertainty by the measured value and multiplying by 100%.

$$\text{percentage uncertainty} = \frac{\text{uncertainty}}{\text{measured value}} \times 100\%$$

For example, suppose a student times a single swing of a pendulum. The measured time is 1.4 s and the estimated uncertainty is 0.2 s. Then we have:

$$\text{percentage uncertainty} = \frac{\text{uncertainty}}{\text{measured value}} \times 100\%$$

$$= \frac{0.2}{1.4} \times 100\% = 14\%$$

This gives a percentage uncertainty of 14%. We can show our measurement in two ways:

- with absolute uncertainty:
  time for a single swing = 1.4 s ± 0.2 s
- with percentage uncertainty:
  time for a single swing = 1.4 s ± 14%

(Note that the absolute uncertainty has a unit whereas the percentage uncertainty is a fraction, shown with a % sign.)

A percentage uncertainty of 14% is very high. This could be reduced by measuring the time for 20 swings. In doing so the absolute uncertainty remains 0.2 s (it is the uncertainty in starting and stopping the stopwatch which is the important thing here, not the accuracy of the stopwatch itself), but the total time recorded might now be 28.4 s.

$$\text{percentage uncertainty} = \frac{0.2}{28.4} \times 100\% = 0.7\%$$

So measuring 20 oscillations rather than just one reduces the percentage uncertainty to less than 1%. The time for one swing is now calculated by dividing the total time by 20, giving 1.42 s. Note that, with a smaller uncertainty, we can give the result to 2 decimal places. The percentage uncertainty remains at 0.7%:

time for a single swing = 1.42 s ± 0.7%

QUESTIONS

10 The depth of water in a bottle is measured as 24.3 cm, with an uncertainty of 0.2 cm. (This could be written as 24.3 ± 0.2 cm.) Calculate the percentage uncertainty in this measurement.

11 The angular amplitude of a pendulum is measured as 35 ± 2°.

a Calculate the percentage uncertainty in the measurement of this angle.

b The protractor used in this measurement was calibrated in degrees. Suggest why the user only feels confident to give the reading to within 2°.

12 A student measures the potential difference across a battery as 12.4 V and states that his measurement has a percentage uncertainty of 2%. Calculate the absolute uncertainty in his measurement.

Recording results

It is important that you develop the skill of recording results in a clear and concise manner.

Generally, numerical results will be recorded in a table. The table should be neatly drawn using a ruler and each heading in the table should include both the quantity being measured and the unit it is measured in.

Table P1.1 shows how a table may be laid out. The measured quantities are the length of the wire and the current though it; both have their units included. Similarly the calculated quantity, $\frac{1}{\text{current}}$, is included and this too has a unit, A^{-1}.

When recording your results, you need to think once more about the precision to which the quantities are measured. In the example in Table P1.1, the length of the wire might be measured to the nearest millimetre and the current might be measured to the nearest milliampere.

Note how '.0' is included in the second result for the length of the wire, to show that the measurement is to the nearest millimetre, not the nearest centimetre. Similarly the zero after the 0.35 shows that it is measured to the nearest milliampere or $\frac{1}{1000}$ of an ampere.

The third column is calculated and should show the same number of significant figures, or one more than the quantity (or quantities) it is calculated from. In this example the current is measured to three significant figures so the inverse of the current is calculated to three significant figures.

Length of wire / cm	Current / A	$\frac{1}{\text{current}}$ / A^{-1}
10.3	0.682	1.47
19.0	0.350	2.86

Table P1.1 A typical results table.

Analysing results

When you have obtained your results the next thing to do is to analyse them. Very often this will be done by plotting a graph.

You may be asked to plot a graph in a particular way however, the general rule is that the variable you control (the **independent variable**) is plotted on the *x*-axis and the variable which that changes as a result (the **dependent variable**) is plotted on the *y*-axis.

In the example in Table P1.1, the length of the wire would be plotted on the *x*-axis and the current (or $\frac{1}{\text{current}}$) would be plotted on the *y*-axis.

You should label your axes with both the quantities you are using and their units. You should then choose your scales to use as much of the graph paper as possible. However, you also need to keep the scales simple. Never choose scales which are multiples of 3, 7, 11 or 13. Try and stick to scales which are simple multiples of 1, 2 or 5.

Plot your points carefully using small crosses; dots tend to disappear into the page and larger dots become blobs, the centre of which is difficult to ascertain.

QUESTION

13 A ball is allowed to roll down a ramp from different starting points. Figure P1.10 shows the apparatus used. The ramp is placed at a fixed height above the floor. You are asked to measure the vertical height *h* of the starting point above the bottom of the ramp and the horizontal distance *d* the ball travels after it leaves the ramp.

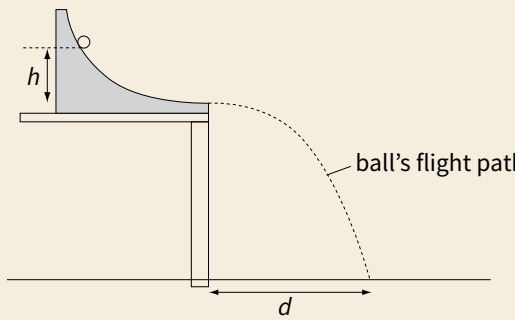


Figure P1.10 For Question 13.

You are also asked to find the square of the horizontal distance the ball travels after it leaves the ramp.

Table P1.2 shows the raw results for the experiment. Copy and complete the table.

h / cm	*d* / cm	d^2 /
1.0	18.0	
2.5	28.4	
4.0	35.8	
5.5	41.6	
7.0	47.3	
9.0	53.6	

Table P1.2 For Question 13.

Many, but not all, graphs you meet will be straight lines. The points may not all lie exactly on the straight line and it is your job to choose the **best fit line**. Choosing this line is a skill which you will develop through the experience of doing practical work.

Generally, there should be equal points either side of the line (but not three on one side at one end and three on the other at the other end). Sometimes all the points, bar one, lie on the line. The point not on the line is often referred to as an anomalous point, and it should be checked if possible. If it still appears to be off the line it might be best to ignore it and use the remaining points to give the best line. It is best to mark it clearly as 'anomalous'.

In Figure P1.11, the line chosen on the first graph is too shallow. By swinging it round so that it is steeper, it goes closer to more points and they are more evenly distributed above and below the line.

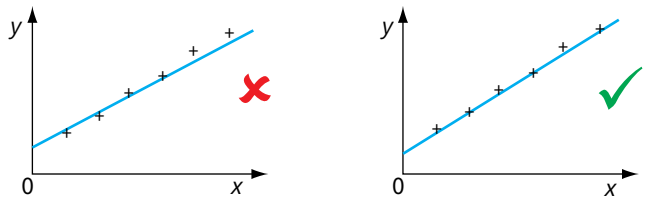


Figure P1.11

Deductions from graphs

There are two major points of information that can be obtained from straight-line graphs: the gradient and the intercept with the y-axis. When measuring the gradient a triangle should drawn, as in Figure P1.12, using at least half of the line that has been drawn.

$$\text{gradient} = \frac{\text{change in } y}{\text{change in } x} = \frac{\Delta y}{\Delta x}$$

In the mathematical equation $y = mx + c$, m is equal to the gradient of the graph and c is the intercept with the y-axis. If c is equal to zero, the graph passes through the origin, the equation becomes $y = mx$ and we can say that y is proportional to x.

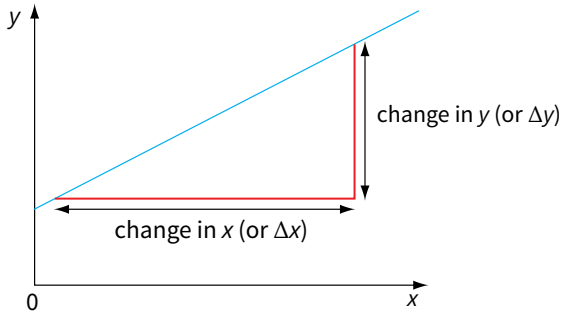


Figure P1.12

QUESTION

14 a Use your results from Question 13 to plot a graph of the height h from which the ball is released against the square of the horizontal distance d^2 it travels before it hits the ground. Draw the best fit line.

b Determine the gradient of the line on your graph and the intercept with the y-axis. Remember, both the gradient and the intercept have units; these should be included in your answer.

Curves and tangents

You also need to develop the skill of drawing smooth curves through a set of points, and drawing tangents to those points. When drawing curves you need to draw a single smooth curve, without any jerks or feathering. As with a straight line, not every point will lie precisely on the curve, and there should be a balance of points on either side.

In the first graph of Figure P1.13, the student has joined each of the points using a series of straight lines. This should never be done. The second graph is much better, although there is some feathering at the left-hand side, as two lines can be seen. The third graph shows a well-drawn curve.

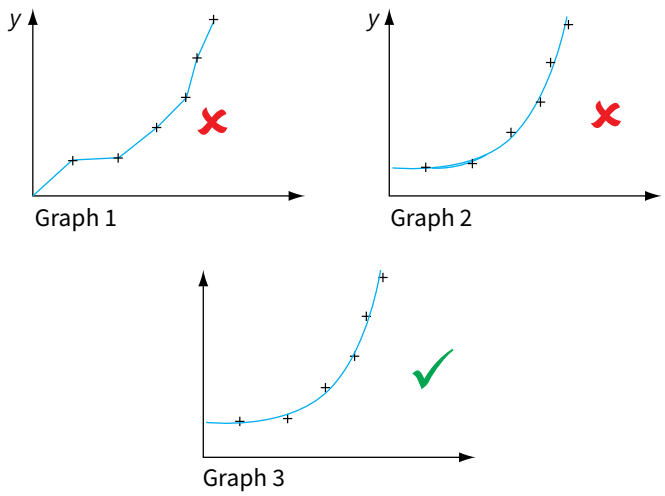


Figure P1.13

Testing a relationship

The readings from an experiment are often used to test a relationship between two quantities, typically whether two quantities are proportional or inversely proportional.

You should know that if two quantities y and x are directly proportional:

- the formula that relates them is $y = kx$, where k is a constant
- if a graph is plotted of y against x then the graph is a straight line through the origin and the gradient is the value of k.

If the two quantities are inversely proportional then $y = k/x$ and a graph of y against $1/x$ gives a straight line through the origin.

These statements can be used as a basis for a test. If a graph of y against x is a straight line through the origin, then y and x are directly proportional. If you know the values of y and x for two points, you can then calculate two values of k with the formula $k = y/x$ and see whether these two values of k are actually the same. But what if the points are not exactly on a straight line or the two values of k are not exactly the same – is the relationship actually false or is it just that errors caused large uncertainties in the readings?

In Chapter P2 we will look at how to combine the uncertainties in the values for y and x to find an uncertainty for k. However, you can use a simple check to see whether the difference in the two values of k may be due to the uncertainties in the readings. For example, if you found that the two values of k differ by 2% but the uncertainties in the readings of y and x are 5%, then you cannot say that the relationship is proved false. Indeed, you are able to say that the readings are consistent with the relationship.

You should first write down a criterion for checking whether the values of k are the same. This criterion is just a simple rule you can invent for yourself and use to compare the two values of k with the uncertainties in the readings. If the criterion is obeyed, you can then write down that the readings are consistent with the relationship.

Criterion 1

A simple approach is to assume that the percentage uncertainty in the value of k is about equal to the percentage uncertainty in either x or y; choose the larger percentage uncertainty of x or y.

You first look at the percentage uncertainty in both x and y and decide which is bigger. Let us assume that the larger percentage uncertainty is in x. Your stated criterion is then that 'if the difference in the percentage uncertainty in the two values of k is less than the percentage uncertainty in x, then the readings are consistent with the relationship'.

> If the % difference in k values is less than the % uncertainty in x (or y), the readings are consistent with the relationship.

Criterion 2

Another criterion is to state that the k values should be the same within 10% or 20%, depending on the experiment and the uncertainty that you think sensible. It is helpful if the figure of 10% or 20% is related to some uncertainty in the actual experiment.

Whatever criterion you use, it should be stated clearly and a clear conclusion given. The procedure to check whether two values of k are reasonably constant is as follows:

- Calculate two values of the constant k. The number of significant figures chosen when writing down these values should be equal to the least number of significant figures in the data used. If you are asked to justify the number of significant figures you give for your value of k, state the number of significant figures that x and y were measured to and that you will choose the smallest. Do not quote your values of k to 1 significant figure to make them look equal when x and y were measured to 2 significant figures.
- Calculate the percentage difference in the two calculated values of k. It is worthwhile using one more significant figure in each actual value of k than is completely justified in this calculation.
- Compare the percentage difference in the two values of k with your clearly stated criterion. You could compare your percentage difference in k values with the larger of the percentage differences in x and y.

WORKED EXAMPLES

2 A student investigates the depth D of a crater made when ball-bearings of different diameters d are dropped into sand. He drops two ball bearings from the same height and measures the depth of the craters using a 30 cm ruler. The results are shown in Table P1.3.

Diameter of ball bearing d / mm	Depth of the crater D / mm	D/d
5.42 ± 0.01	36 ± 2	6.64
3.39 ± 0.01	21 ± 2	6.19

Table P1.3 For Worked example 2.

It is suggested that the depth D of the crater is directly proportional to the diameter d of the ball-bearing, that is:

$D = kd$ or $\frac{D}{d} = k$

Do the readings support this hypothesis?

Step 1 Calculate the values of $k = \frac{D}{d}$. These values are shown in the third column in Table P1.3, although they should only be given to 2 significant figures as values of D are given to 2 significant figures and values of d to 3 significant figures. The more precise values for k are to be used in the next step.

Step 2 Calculate the percentage difference in the k values. The percentage difference is:

$\frac{0.45}{6.19} \times 100\% = 7.2\%$

So the k values differ by 7% of the smaller value.

Step 3 State a criterion and check it.

'My criterion is that, if the hypothesis is true, then the percentage difference in the k values will be less than the percentage uncertainty in D. I chose D as it obviously has the higher percentage uncertainty.'

The uncertainty in the smaller measurement of D can be calculated as:

uncertainty in $D = \frac{2}{21} \times 100\% = 9.5\%$

The percentage difference in the k values is less than the uncertainty in the experimental results; therefore the experiment is consistent with the hypothesis.

Of course we cannot say for sure that the hypothesis is correct. To do that we would need to greatly reduce the percentage uncertainties.

3 A student obtains data shown in Table P1.4.

x / cm	d / cm
2.0	3.0
3.5	8.0

Table P1.4 For Worked example 3.

The first reading of x was found to have an uncertainty of ± 0.1. Do the results show that d is proportional to x?

Step 1 Calculate the ratio of $\frac{d}{x}$ in both cases:

$\left(\frac{d}{x}\right)_1 = 1.50 \quad \left(\frac{d}{x}\right)_2 = 2.29$

Step 2 Calculate how close to each other the two ratios are:

$2.29 - 1.50 = 0.79$

So the two values of $\left(\frac{d}{x}\right)$ are $\frac{0.79}{1.5} = 53\%$ different.

Step 3 Compare the values and write a conclusion.

The uncertainty in the first value of x is 5% and, since the percentage difference between the ratios is 53% is much greater, the evidence does not support the suggested relationship.

QUESTIONS

15 A student obtains the following data for two variables T and m (Table P1.5).

T / s	m / kg
4.6	0.90
6.3	1.20

Table P1.5 Data for Question 15.

The first value of T has an uncertainty of ± 0.2 s. Do the results show that T is proportional to m?

16 A student obtains the following values of two variables r and t (Table P1.6).

r / cm	t / s
6.2	4.6
12.0	6.0

Table P1.6 Data for Question 16.

The first value of r has an uncertainty of ± 0.2 cm, which is much greater than the percentage uncertainty in t. Do the results show that t^2 is proportional to r?

Identifying limitations in procedures and suggesting improvements

No experiment is perfect and the ability to see weaknesses in the experimental setup and the techniques used is an important skill. You should also take the opportunity to think of ways to improve the experimental technique, thereby reducing the overall percentage uncertainty.

In this section, we will look at five experiments and discuss problems which might arise and the improvements that might be made to overcome them. It will help if you try out some of the experiments yourself so that you get a feel for the methods described. The table for each experiment is a summary of ideas that you might use in your answer.

Experiment 1: Ball bearings and craters

In Worked example 2, the student dropped a ball bearing of diameter d into sand and measured the depth D of the crater produced. He dropped two ball bearings of different diameters from the same height and measured the depth of the crater using a 30 cm ruler. Table P1.7 suggests some of the problems with the simple method used, together with some improvements.

Suggestion	Problem	Improvement
1	'Two results are not enough to draw a valid conclusion.'	'Take more results and plot a graph of D against d.'
2	'The ruler is too wide to measure the depth of the crater.'	'Use a knitting needle and mark the sand level on the needle and then measure with a ruler.'
3	'There may be a parallax error when measuring the top level of the crater.'	'Keep the eye parallel to the horizontal level of the sand, or use a stiff card.'
4	'It is difficult to release the ball bearing without giving it a sideways velocity, leading to a distorted crater.'	'Use an electromagnet to release the ball.'
5	'The crater lip is of varying height.'	'Always measure to the highest point.'

Table P1.7 Suggestions for improving Experiment 1.

It is worth making some points regarding these suggestions.

1. This is a simple idea but it is important to explain how the extra results are to be used. In this case a graph is suggested – alternatively the ratio $\frac{D}{d}$ could be calculated for each set of readings.
2. The problem is clearly explained. It is not enough to just say that the depth is difficult to measure.
3. It is not enough to just say 'parallax errors'. We need to be specific as to where they might occur. Likewise, make sure you make it clear where you look from when you suggest a cure.
4. There is no evidence that this will affect the crater depth, but it is a point worthy of consideration.
5. An interesting point: does the crater depth include the lip or is it just to the horizontal sand surface? Consistency in measurement is what is needed here.

Experiment 2: Timing with a stopwatch

Many years ago, Galileo suggested that heavy and light objects take the same time to fall to the ground from the same height, as illustrated in Figure P1.14. Imagine that you want to test this hypothesis.

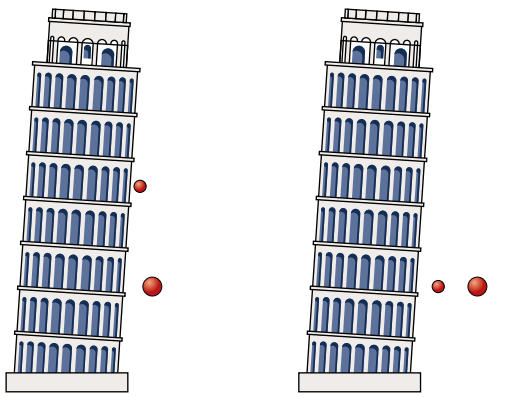


Figure P1.14 It was reputed that Galileo dropped two different masses from the top of the Leaning Tower of Pisa to prove his idea, but this is now thought unlikely to have taken place.

This is an experiment you can do yourself with two objects and a stopwatch, or even a digital wrist watch or a cell phone with a timing app. Drop two different objects, for example two stones, and measure the time they take to fall the same distance to the ground.

Of course the times you obtain are likely to be different. Does this prove Galileo wrong? You can test the relationship and establish whether your readings are consistent with his hypothesis. However, if you improve the experiment and reduce the uncertainties, the conclusion will be much more useful.

Problem	Improvement
'Taking readings for just two masses was not enough.'	'I should use a great range of different masses and plot a graph of the average time to fall to the ground against the mass of the object.'
'It was difficult to start the stopwatch at the same instant that I dropped the stone and to stop it exactly as it hit the ground. I may have been late because of my reaction time.'	'Film the fall of each stone with a video camera which has a timer in the background. When the video is played back, frame by frame, I will see the time when the ball hits the ground on the timer.' (Alternatively you can use light gates connected to a timer to measure the time electronically. You should draw a diagram, explaining that the timer starts when the first light gate is broken and stops when the second is broken.)
'My hand was not steady and so I may not have dropped the stones from exactly the same height each time.'	'Use iron objects which hang from an electromagnet. When the current in the electromagnet is switched off, the object falls.' (A diagram would help – see Figure P.15.)
'The heavier stone was larger in size and it was important that the bottom of each stone started at the same height. There may have been parallax error.'	'Clamp a metre rule vertically and start the bottom of each stone at exactly the top of the ruler each time. To avoid parallax error, I will make sure my line of sight is horizontal, at right angles to the rule.' (A diagram will show this clearly – see Figure P1.15.)
'The times that I measured were very short – not much greater than my reaction time – so reaction time had a great effect.'	'Increase the distance of fall so that the times are larger. This will make the uncertainty in each time measurement smaller in proportion to the time being measured.'

Table P1.8 Suggestions for improving Experiment 2.

When you consider improving an experiment, first consider any practical difficulties and possible sources of inaccuracy. Write them down in detail. Do not just write, for example, 'reaction time' or 'parallax error'. It is always a good idea to start with the idea that more readings need to be taken, possibly over a greater range (for example, in this case, if the masses of the stones were almost equal). Table P1.8 gives other possibilities.

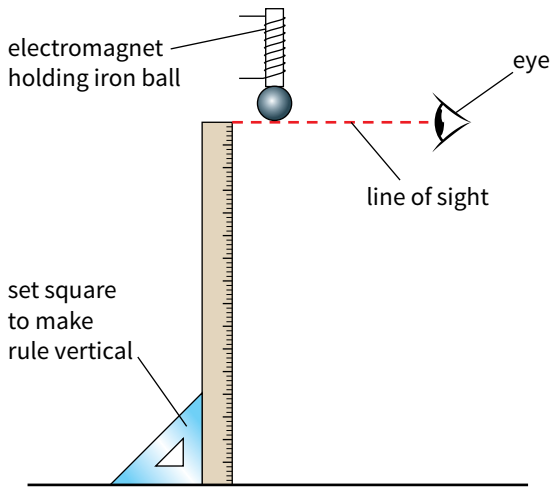


Figure P1.15 Using an electromagnet to release iron objects. The line of sight is clearly shown.

QUESTION

17 Use a stopwatch and a rule to measure the average speed as an object falls from a table to the ground. What are the difficulties and how might they be reduced? Some of the suggestions will be the same as those above but you should also consider difficulties in measuring the distance to the ground and how they can be avoided. Remember, rules have battered ends and the ends may not be at 0 and 100 cm.

Experiment 3: Timing oscillations

In physics, the study of oscillations is of great importance. Indeed, the observation of a pendulum led Galileo to study time intervals and allowed pendulum clocks to be developed.

One skill you will need to develop is finding the time for an oscillation. Figure P1.16 shows a simple pendulum and one complete oscillation. The pendulum is just a small weight, the bob, which hangs on a string.

Figure P1.16 shows that one complete oscillation can be measured in two ways. Which way is better? In fact the second way is better. This is because it is difficult to judge exactly when the pendulum bob is at the end of its swing. It is easier to start timing when the bob is moving quickly past a point; this happens in the middle of the swing. To time from the middle of the swing you should use a

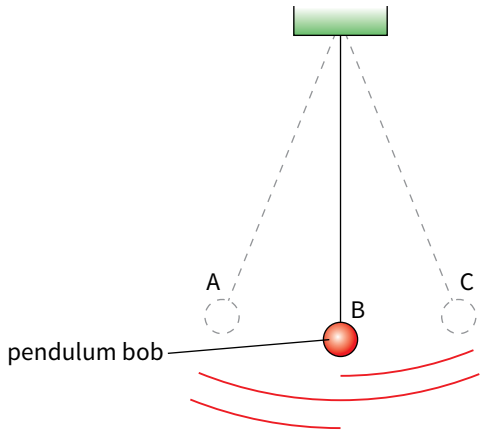


Figure P1.16 One complete oscillation is either from A to C and then back to A, or from B to C then back to B, then to A and back to B, as shown.

fiducial mark. This can be a line on the bench underneath the bob at the centre of the swing, or it can be another object in the laboratory which appears to be in line with the bob when it hangs stationary, as seen from where you are standing. As long as you do not move your position, every time the bob passes this point it passes the centre.

Another way to reduce the uncertainty in the time for one oscillation is to time more than one swing, as explained above in the section on percentage uncertainty.

A simple practical task is to test the hypothesis that the time for one oscillation T is related to the length l of a simple pendulum by the formula $T^2 = kl$, where k is a constant.

What difficulties would you face and what are possible improvements? Table P1.9 gives some possibilities.

QUESTION

18 Hang a mass from a spring or from a rubber band. Use a stopwatch to time the mass as it oscillates up and down. Measure the time for just one oscillation, the time for 10 oscillations and the time for 20 oscillations. Repeat each reading several times. Use your readings to find the time for one complete oscillation and the uncertainty in each time. Draw up a table to show the problems of such measurements and how to reduce them.

Experiment 4: Using force meters

You need to be able to read instruments, estimating the uncertainty, looking for sources of error and trying to improve their use. One such instrument is a force meter or newton-meter, shown in Figure P1.17.

In this experiment, the block is pulled using the force meter to find the force F needed to make a block just start to move. An extra mass is added on top of the block to see whether the relationship $F = km$ is obeyed, where m is the total mass of the block and k is a constant.

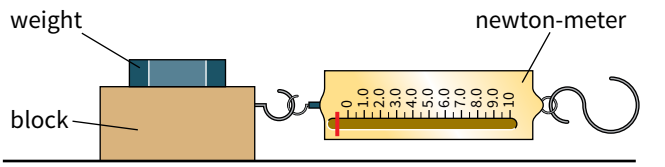


Figure P1.17 A newton-meter, just before it pulls a block along the bench.

Problem	Improvement
'Taking readings for just two lengths was not enough.'	'Use more than two lengths and plot a graph of the average time squared against the length of the string.'
'It was difficult to judge the end of the swing.'	'Use a fiducial mark at the centre of the oscillation as the position to start and stop the stopwatch.' 'Use an electronic timer placed at the centre of the oscillation to measure the time.' 'Make a video of the oscillation with a timer in the background and play it back frame by frame.'
'The oscillations died away too quickly.'	'Use a heavier mass which swings longer.'
'The times were too small to measure accurately, as my reaction time was a significant fraction of the total time.'	'Use longer strings.' 'Time 20 rather than 10 oscillations.'
'It was difficult to measure the length to the centre of gravity of the weight accurately.'	'Use a longer string so any errors are less important.' 'Measure the length to the top of the weight and use a micrometer to measure the diameter of the bob and add on half the diameter to the length of the string.'

Table P1.9 Suggestions for improving Experiment 3.

Problem	Improvement
'Taking readings for just two masses was not enough.'	'Use more than two masses and plot a graph of the force against the mass.'
'It was difficult to zero the newton-meter used horizontally.'	'Use a force sensor and computer.' 'Use a pulley and string to connect a tray to the block. Then tip sand onto a tray until the block starts to move. The weight of the sand and tray is then the force.'
'The reading of F was very low on the scale and gave a large percentage uncertainty.'	'Use heavier masses on top of the block.'
'The block starts to move suddenly and it is difficult to take the reading as this happens.'	'Video the experiment and play back frame by frame to see the largest force.' 'Use a force sensor and computer.'
'Different parts of the board are rougher than others.'	'Mark round the block with a pencil at the start and put it back in the same place each time.'

Table P1.10 Suggestions for improving Experiment 4.

Look closely at Figure P1.17. When reading the meter, the uncertainty is the smallest scale division on the meter, unless one can reasonably read between the markings. This is difficult and so an uncertainty of 0.5 N, the smallest scale division, is reasonable.

Another problem in using the meter is that it reads less than zero before it is pulled. It needs a small force to bring the meter to zero. This is a zero error and all the actual readings will be too large by the same amount. This is probably because the meter was adjusted to read zero when hanging vertically and it is now being used horizontally.

QUESTION

19 A laboratory thermometer can be used to measure an increase in temperature.

a Suggest a value for the uncertainty in such a reading. (You may need to look at some different thermometers.)

If you grip the bulb of the thermometer gently in your fingers, the reading rises to a new value. The reading will be different depending on whether you cover the bulb entirely or only partially with your fingers.

b Describe how you would test whether the temperature rise is proportional to the area of the bulb covered by your fingers You can take the surface area of the bulb to be 1 cm^2 and when you cover half of the bulb the area covered is 0.5 cm^2. The exact value of the surface area is not important; just the ratio is important.

c Suggest difficulties with this experiment, and how it might be improved. One problem with a thermometer is that it takes time for the reading to rise. What can you do about this?

Fortunately the meter can be adjusted to read zero before starting to pull.

Table P1.10 describes the problems which may be encountered with this experiment, together with suggested improvements.

Experiment 5: Electrical measurements

Electrical experiments have their own problems. Figure P1.18 shows an apparatus used to test the hypothesis that the resistance R of a wire is related to its length l by the formula $R = kl$, where k is a constant. The current is kept constant and the voltmeter reading is taken at two different values of l, for $l = 0.30$ m and 0.50 m.

What problems are likely to arise when using this apparatus? Table P1.11 identifies some possible problems with this experiment, and some suggestions for improvement.

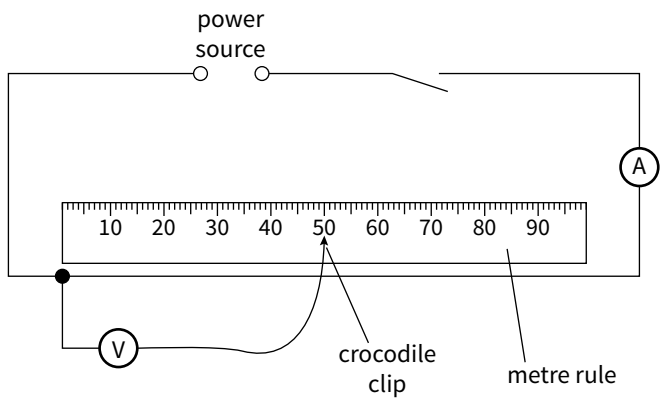


Figure P1.18 Apparatus used to check the hypothesis $R = kl$.

Problem	Improvement
'Taking readings for just two lengths was not enough.'	'Use more than two lengths and plot a graph of the voltmeter reading against the length.' 'Calculate more than just two values of k.'
'Difficult to measure the length of the wire as the clips have width and I don't know where inside they grip the wire.'	'Use narrower clips.' 'Solder the contacts onto the wire.'
'The scale is not sensitive enough and can only measure to 0.05 V.'	'Use a voltmeter that reads to 0.01 V.' 'Use a digital voltmeter.'
'The values of voltage are small, particularly at 0.30 m.'	'Use a larger current so that when l = 0.50 m the voltmeter reading is at the top of the scale.'
'The voltmeter reading fluctuates because of contact resistance.'	'Clean the wires with wire wool first.'
'Other factors may have changed the resistance; for example, the temperature may have increased because of the current.'	'Wait a long time until the wire has reached a constant temperature.' 'Use smaller currents, but with a more sensitive voltmeter.'

Table P1.11 Suggestions for improving Experiment 5.

End-of-chapter questions

1 What is the uncertainty in the following sets of readings? All of them are written down to the smallest division on the instrument used in their measurement.

a 24.6, 24.9, 30.2, 23.6 cm [1]

b 2.66, 2.73, 3.02 s [1]

c 24.0, 24.0, 24.0 g [1]

2 Electrical experiments usually involve the reading of meters such as the voltmeters shown in Figure P1.19.

a What is the reading shown by each voltmeter, and the uncertainty in each reading? [3]

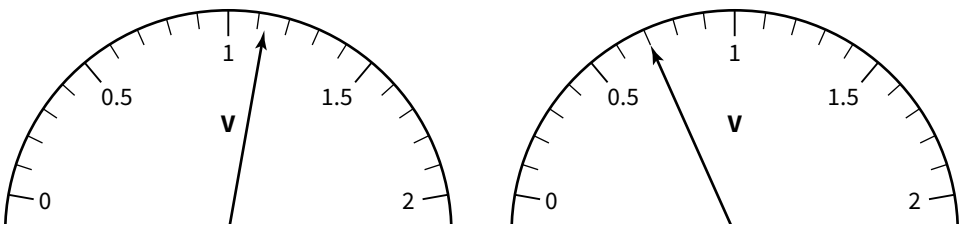


Figure P1.19 Two voltmeter scales – for End-of-chapter Question 2.

b The voltmeters show the readings obtained when they were connected across two wires which were identical apart from their different lengths. The current in each wire was 0.500 A and the length l of the wire was 30.0 cm in the right diagram and 50.0 cm in the left diagram.

Use the scale readings to test the hypothesis that the resistance R of the wire is proportional to length l. Consider the effect of the uncertainties on your conclusion. [4]

3 The apparatus shown in Figure P1.20 can be used to test the hypothesis that T, the time taken for a ball to roll down a plane from rest, is related to the distance s by the formula $T^2 = ks$, where k is a constant. The ball is timed using a stopwatch over two different values of s.

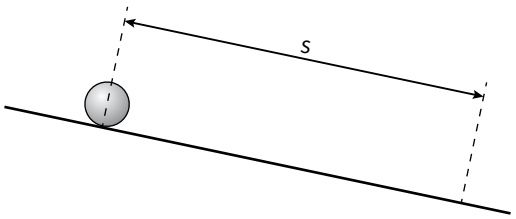


Figure P1.20 For end-of-chapter Question 3.

Suggest problems with the experiment and how they might be overcome. You should consider problems in measuring the distance as well as the time. Also note what happens to the ball; it may not roll in the way that you expect. [8]

Questions 4-6 are designed to illustrate some aspects of practical questions. They are not formal practical questions as, ideally, you should perform the experiment yourself and take some readings. This helps you to see the problems.

4 An experiment explores the relationship between the period of a vibrating spring and the mass m in a pan holder. The student is instructed to set up the apparatus as in Figure P1.21, with a mass of 200 g in the pan. He is then told to move the pan downwards by approximately 1 cm and to release it so that it vibrates in a vertical direction.

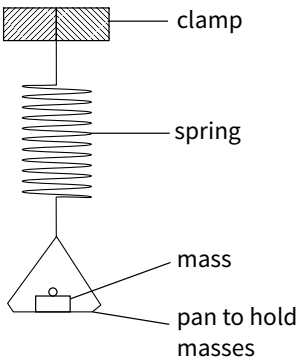


Figure P1.21 For end-of-chapter Question 4.

The student is asked to record the time taken for 20 oscillations of the spring, and then to repeat the procedure, using masses between 20 g and 200 g until he has six sets of readings. Columns are provided in the table for $\sqrt{m}$ and T, the period of the pendulum.

Table P1.12 shows the readings taken by a student with the different masses.

Mass / g	Time for 20 oscillations / s	$\sqrt{m}$	T
20	12.2		
50	15.0		
100	18.7		
150	21.8		
200	24.5		
190	24.0		

Table P1.12 For end-of-chapter Question 4.

a Copy the table and include values for $\sqrt{m}$ and T. [2]

b Plot a graph of T on the y-axis against $\sqrt{m}$ on the x-axis. Draw the straight line of best fit. [4]

c Determine the gradient and y-intercept of this line. [2]

d The quantities T and m are related by the equation

$T = C + k\sqrt{m}$

where C and k are constants.

Find the values of the two constants C and k. Give appropriate units. [2]

5 A student releases a toy car to roll down a ramp, as shown in Figure P1.22. He measures the distance l from the middle of the car as it is released to the bottom of the ramp and the distance s travelled along the straight section before the car stops. He also measures the time t taken to travel the distance s. He then repeats the experiment using a different value of l.

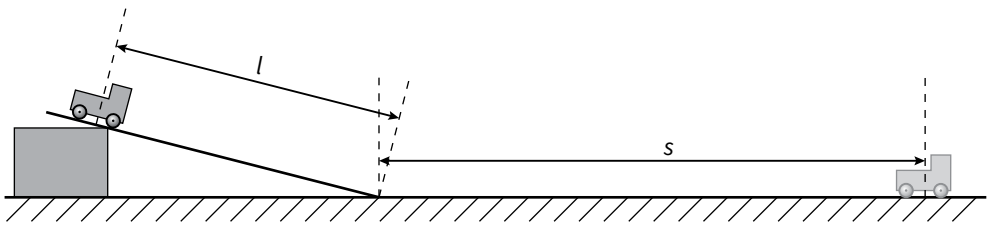


Figure P1.22 For End-of-chapter Question 5.

The student obtained readings with $l = 40$ and 60 cm, taking each reading for s and t twice. The readings were:

$l = 40.0$ cm: values for s were 124 and 130 cm; values for t were 4.6 and 4.8 s

$l = 60.0$ cm: values for s were 186 and 194 cm; values for t were 4.9 and 5.2 s.

a For the smaller value of l, obtain a value for:

i the average value of s [1]

ii the absolute and percentage uncertainty in the value of s [2]

iii the average value of t [1]

iv the absolute and percentage uncertainty in the value of t. [2]

b **i** For both values of l, calculate the average speed v of the car along the straight section of track using the relationship $v = s/t$. [1]

ii Justify the number of significant figures that you have given for your values of v. [1]

c **i** It is suggested that s is proportional to l. Explain whether the readings support this relationship. [2]

ii (HARDER) It is suggested that v^2 is proportional to l. Explain whether the readings support this relationship. [2]

d Describe **four** sources of uncertainty or limitations of the procedure for this experiment. [4]

e Describe **four** improvements that could be made to this experiment. You may suggest the use of other apparatus or different procedures. [4]

6 The apparatus in Figure P1.23 shows a resistor in some water.

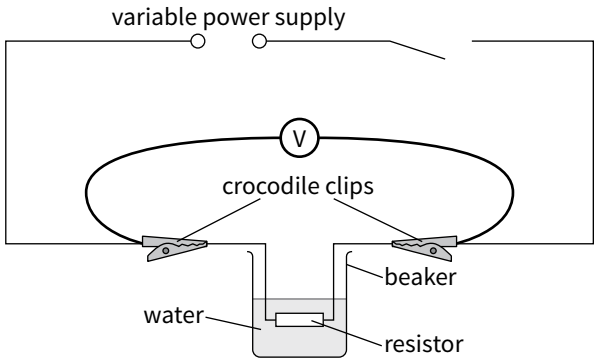


Figure P1.23 Apparatus for end-of-chapter Question 6.

A student measures the rise in temperature θ of the water in 100 s using two different values of voltage.

The student wrote:

'When the voltage was set at 6.0 V the rise in temperature of the water in 100 s was 14.5 °C. The voltmeter reading decreased by about 0.2 V during the experiment and so the final voltmeter reading was 5.8 V.

'The reading fluctuated from time to time by about 0.2 V. The smallest scale division on the thermometer was 1 °C but I could read it to 0.5 °C. I did not have time to repeat the reading.

'When the voltage was set at 12.0 V the rise in temperature in 100 s was 51.0 °C and the voltage was almost the same at the end but fluctuated by about 0.2 V.'

a Estimate the percentage uncertainty in the measurement of the first voltage. [1]

b It is suggested that θ is related to V according to the formula $\theta = kV^2$, where k is a constant.

i Calculate **two** values for k. Include the units in your answer. [2]

ii Justify the number of significant figures you have given for your value of k. [1]

iii Explain whether the results support the suggested relationship. [1]

c Describe **four** sources of uncertainty or limitations of the procedure for this experiment. [4]

d Describe **four** improvements that could be made to this experiment. You may suggest the use of other apparatus or different procedures. [4]

Chapter 17: Circular motion

Learning outcomes

You should be able to:

- express angular displacement in radians
- solve problems using the concept of angular velocity
- describe motion along a circular path as due to a perpendicular force which causes a centripetal acceleration
- recall and use equations for centripetal acceleration and for centripetal force

Moving in circles

The racing car in Figure 17.1 shows two examples of circular motion. The car's wheels spin around the axles, and the car follows a curved path as it speeds round the bend.



Figure 17.1 Circular motion: the car's wheels go round in circles as the car itself follows a curved path.

Describing circular motion

Many things move in circles. Here are some examples:

- the wheels of a car or a bicycle
- the Earth in its (approximately circular) orbit round the Sun
- the hands of a clock
- a spinning DVD in a laptop
- the drum of a washing machine.

Sometimes, things move along a path that is part of a circle. For example, the car in Figure 17.1 is travelling around a bend in the road which is an arc of a circle.

Circular motion is different from the straight-line motion that we have discussed previously in our study of kinematics and dynamics in Chapters 1–6. However, we can extend these ideas of dynamics to build up a picture of circular motion.

Around the clock

The second hand of a clock moves steadily round the clock face. It takes one minute for it to travel all the way round the circle. There are 360° in a complete circle and 60 seconds in a minute. So the hand moves 6° every second. If we know the angle θ through which the hand has moved from the vertical (12 o'clock) position, we can predict the position of the hand.

In the same way, we can describe the position of any object as it moves around a circle simply by stating the angle θ of the arc through which it has moved from its starting position. This is shown in Figure 17.2.

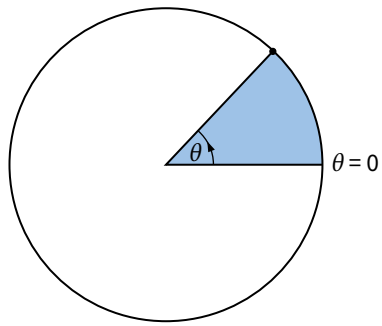


Figure 17.2 To know how far an object has moved round the circle, we need to know the angle θ.

The angle θ through which the object has moved is known as its **angular displacement**. For an object moving in a straight line, its position was defined by its displacement s, the **distance** it has travelled from its starting position. The corresponding quantity for circular motion is angular displacement θ, the **angle** of the arc through which the object has moved from its starting position.

QUESTION

1 **a** By how many degrees does the angular displacement of the hour hand of a clock change each hour?

b A clock is showing 3.30. Calculate the angular displacements in degrees from the 12.00 position of the clock to:

i the minute hand

ii the hour hand.

Angles in radians

When dealing with circles and circular motion, it is more convenient to measure angles and angular displacements in units called radians rather than in degrees. If an object moves a distance s around a circular path of radius r (Figure 17.3a), its angular displacement θ in **radians** is defined as follows:

$$\text{angle (in radians)} = \frac{\text{length of arc}}{\text{radius}}$$

or $$\theta = \frac{s}{r}$$

Since both s and r are distances measured in metres, it follows that the angle θ is simply a ratio. It is a dimensionless quantity. If the object moves twice as far around a circle of twice the radius (Figure 17.3b), its angular displacement θ will be the same.

$$\theta = \frac{\text{length of arc}}{\text{radius}} = \frac{2s}{2r} = \frac{s}{r}$$

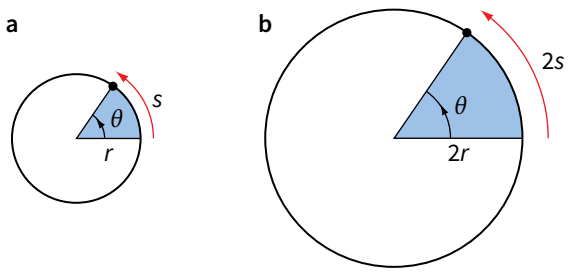


Figure 17.3 The size of an angle depends on the radius and the length of the arc. Doubling both leaves the angle unchanged.

When we define θ in this way, its units are radians rather than degrees. How are radians related to degrees? If an object moves all the way round the circumference of the circle, it moves a distance of $2\pi r$. We can calculate its angular displacement in radians:

$$\theta = \frac{\text{circumference}}{\text{radius}} = \frac{2\pi r}{r} = 2\pi$$

Hence a complete circle contains 2π radians. But we can also say that the object has moved through 360°. Hence:

$$360° = 2\pi \text{ rad}$$

Similarly, we have:

$180° = \pi$ rad $\qquad 90° = \frac{\pi}{2}$ rad

$45° = \frac{\pi}{4}$ rad $\qquad$ and so on

Defining the radian

An angle of one radian is defined as follows (see Figure 17.4):

> One **radian** is the angle subtended at the centre of a circle by an arc of length equal to the radius of the circle.

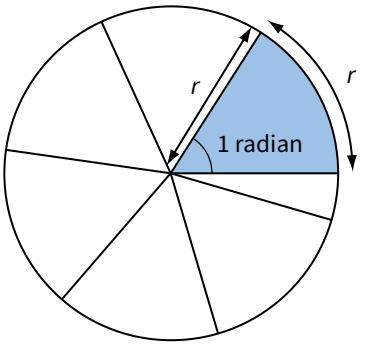


Figure 17.4 The length of the arc is equal to the radius when the angle is 1 radian.

An angle of 360° is equivalent to an angle of 2π radians. We can therefore determine what 1 radian is equivalent to in degrees.

$$1 \text{ radian} = \frac{360°}{2\pi}$$

or $$1 \text{ radian} \approx 57.3°$$

If you can remember that there are 2π rad in a full circle, you will be able to convert between radians and degrees:

- to convert from degrees to radians, multiply by $\frac{2\pi}{360°}$ or $\frac{\pi}{180°}$
- to convert from radians to degrees, multiply by $\frac{360°}{2\pi}$ or $\frac{180°}{\pi}$

Now look at Worked example 1.

WORKED EXAMPLE

1 If $\theta = 60°$, what is the value of θ in radians?

The angle θ is 60°. 360° is equivalent to 2π radians. Therefore:

$$\theta = 60 \times \frac{2\pi}{360}$$

$$= \frac{\pi}{3} = 1.05 \text{ rad}$$

(Note that it is often useful to express an angle as a multiple of π radians.)

QUESTION

2 **a** Convert the following angles from degrees into radians: 30°, 90°, 105°.

b Convert these angles from radians to degrees: 0.5 rad, 0.75 rad, π rad, $\frac{\pi}{2}$ rad.

c Express the following angles as multiples of π radians: 30°, 120°, 270°, 720°.

Steady speed, changing velocity

If we are to use Newton's laws of motion to explain circular motion, we must consider the **velocity** of an object going round in a circle, rather than its **speed**.

There is an important distinction between speed and velocity: **speed** is a scalar quantity which has magnitude only, whereas **velocity** is a vector quantity, with both magnitude and direction. We need to think about the direction of motion of an orbiting object.

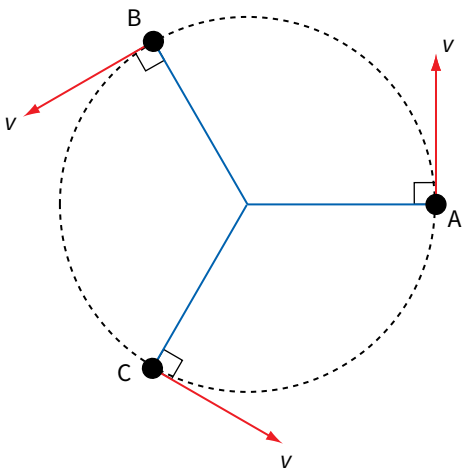


Figure 17.5 The velocity *v* of an object changes direction as it moves along a circular path.

Figure 17.5 shows how we can represent the velocity of an object at various points around its circular path. The arrows are straight and show the direction of motion at a particular instant. They are drawn as tangents to the circular path. As the object travels through points A, B, C, etc., its speed remains constant but its direction changes. Since the direction of the velocity *v* is changing, it follows that *v* itself (a vector quantity) is changing as the object moves in a circle.

Angular velocity

As the hands of a clock travel steadily around the clock face, their velocity is constantly changing. The minute hand travels round 360° or 2π radians in 3600 seconds. Although its velocity is changing, we can say that its **angular velocity** is constant, because it moves through the same angle each second:

$$\text{angular velocity} = \frac{\text{angular displacement}}{\text{time taken}}$$

$$\omega = \frac{\Delta\theta}{\Delta t}$$

We use the symbol ω (Greek letter omega) for angular velocity, measured in radians per second (rad s^{-1}). For the minute hand of a clock, we have $\omega = \frac{2\pi}{3600} \approx 0.001\,75$ rad s^{-1}.

QUESTIONS

5 Show that the angular velocity of the second hand of a clock is about 0.105 rad s^{-1}.

6 The drum of a washing machine spins at a rate of 1200 rpm (revolutions per minute).

a Determine the number of revolutions per second of the drum.

b Determine the angular velocity of the drum.

QUESTIONS

3 Explain why all the velocity arrows in Figure 17.5 are drawn the same length.

4 A toy train travels at a steady speed of 0.2 m s^{-1} around a circular track (Figure 17.6). A and B are two points diametrically opposite to one another on the track.

a Determine the change in the speed of the train as it travels from A to B.

b Determine the change in the velocity of the train as it travels from A to B.

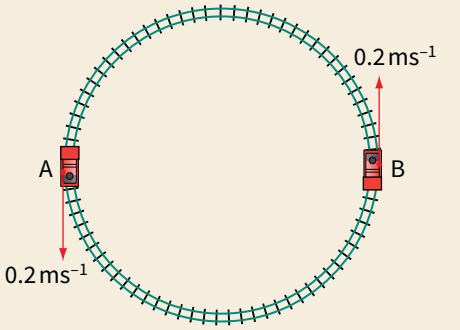


Figure 17.6 A toy train travelling around a circular track – for Question 4.

Relating velocity and angular velocity

Think again about the second hand of a clock. As it goes round, each point on the hand has the same angular velocity. However, different points on the hand have different velocities. The tip of the hand moves fastest; points closer to the centre of the clock face move more slowly.

This shows that the speed v of an object travelling around a circle depends on two quantities: its angular velocity ω and its distance from the centre of the circle r. We can write the relationship as an equation:

speed = angular velocity × radius

$$v = \omega r$$

Worked example 2 shows how to use this equation.

WORKED EXAMPLE

2 A toy train travels around a circular track of radius 2.5 m in a time of 40 s. What is its speed?

Step 1 Calculate the train's angular velocity ω. One circuit of the track is equivalent to 2π radians. The rain travels around in 10 s. Therefore:

$$\omega = \frac{2\pi}{40} = 0.157\,\text{rad}^{-1}$$

Step 2 Calculate the train's speed:

$$v = \omega r = 0.157 \times 2.5 = 0.39\,\text{m s}^{-1}$$

Hint: *You could have arrived at the same answer by calculating the distance travelled (the circumference of the circle) and dividing by the time taken.*

QUESTIONS

7 The angular velocity of the second hand of a clock is 0.105 rad s^{-1}. If the length of the hand is 1.8 cm, calculate the speed of the tip of the hand as it moves round.

8 A car travels around a 90° bend in 15 s. The radius of the bend is 50 m.

a Determine the angular velocity of the car.

b Determine the speed of the car.

9 A spacecraft orbits the Earth in a circular path of radius 7000 km at a speed of 7800 m s^{-1}. Determine its angular velocity.

Centripetal forces

When an object's velocity is changing, it has acceleration. In the case of uniform circular motion, the acceleration is rather unusual because, as we have seen, the object's speed does not change but its velocity does. How can an object accelerate and at the same time have a steady speed?

One way to understand this is to think about what Newton's laws of motion can tell us about this situation. **Newton's first law** states that an object remains at rest or in a state of uniform motion (at constant speed in a straight line) unless it is acted on by an external force. In the case of an object moving at steady speed in a circle, we have a body whose velocity is not constant; therefore, there must be a resultant (unbalanced) force acting on it.

Now we can think about different situations where objects are going round in a circle and try to find the force that is acting on them.

- Consider a rubber bung on the end of a string. Imagine whirling it in a horizontal circle above your head (Figure 17.7). To make it go round in a circle, you have to pull on the string. The pull of the string on the bung is the unbalanced force, which is constantly acting to change the bung's velocity as it orbits your head. If you let go of the string, suddenly there is no tension in the string and the bung will fly off at a tangent to the circle.
- Similarly, as the Earth orbits the Sun, it has a constantly changing velocity. Newton's first law suggests that there must be an unbalanced force acting on it. That force is the gravitational pull of the Sun. If the force disappeared, the Earth would travel off in a straight line.

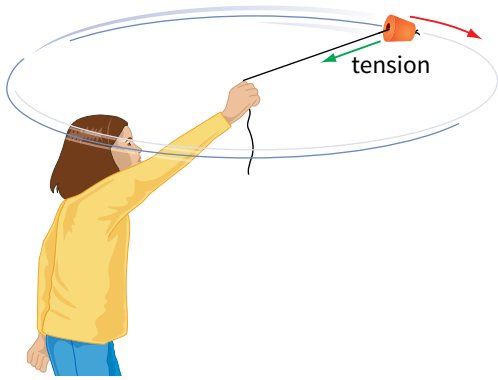


Figure 17.7 Whirling a rubber bung.

In both of these cases, you should be able to see why the direction of the force is as shown in Figure 17.8. The force on the object is directed towards the centre of the circle. We describe each of these forces as a **centripetal force** – that is, directed towards the centre.

It is important to note that the word **centripetal** is an adjective. We use it to describe a force that is making

something travel along a circular path. It does not tell us what causes this force, which might be gravitational, electrostatic, magnetic, frictional or whatever.

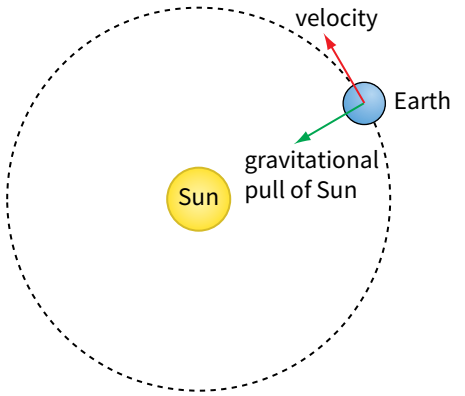


Figure 17.8 The gravitational pull of the Sun provides the centripetal force that keeps the Earth in its orbit.

QUESTIONS

10 In each of the following cases, state what provides the centripetal force:

- **a** the Moon orbiting the Earth
- **b** a car going round a bend on a flat, rough road
- **c** the weight on the end of a swinging pendulum.

11 A car is travelling along a flat road. Explain why it cannot go around a bend if the road surface is perfectly smooth. Suggest what might happen if the driver tries turning the steering wheel.

Vector diagrams

Figure 17.9a shows an object travelling along a circular path, at two positions in its orbit. It reaches position B a short time after A. How has its velocity changed between these two positions?

The change in the velocity of the object can be determined using a vector triangle. The vector triangle in Figure 17.9b shows the difference between the final velocity v_B and initial velocity v_A. The change in the velocity of the object between the points B and A is shown by the smaller arrow labelled Δv. Note that the change in the velocity of the object is (more or less):

- at right angles to the velocity at A
- directed towards the centre of the circle.

The object is accelerating because its velocity changes. Since acceleration is the rate of change of velocity:

$$a = \frac{\Delta v}{\Delta t}$$

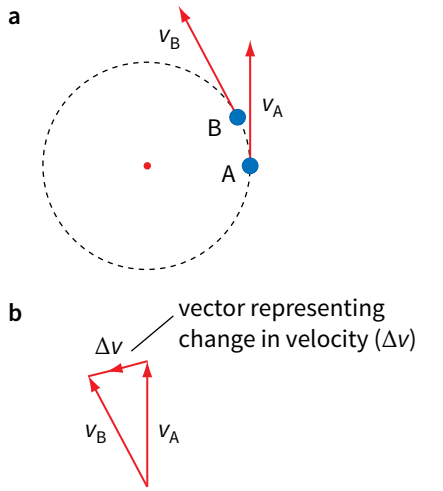


Figure 17.9 Changes in the velocity vector.

it follows that the acceleration of the object must be in the same direction as the change in the velocity – towards the centre of the circle. This is not surprising because, according to $F = ma$, the acceleration a of the object is in the same direction as the centripetal force F.

Acceleration at steady speed

Now that we know that the centripetal force F and acceleration are always at right angles to the object's velocity, we can explain why its speed remains constant. If the force is to make the object change its speed, it must have a component in the direction of the object's velocity; it must provide a push in the direction in which the object is already travelling. However, here we have a force at 90° to the velocity, so it has no component in the required direction. (Its component in the direction of the velocity is $F\cos 90° = 0$.) It acts to pull the object around the circle, without ever making it speed up or slow down.

You can also use the idea of work done to show that the speed of the object moving in a circle remains the same. The work done by a force is equal to the product of the force and the distance moved by the object in the direction of the force. The distance moved by the object in the direction of the centripetal force is zero; hence the work done is zero. If no work is done on the object, its kinetic energy must remain the same and hence its speed is unchanged.

QUESTION

12 An object follows a circular path at a steady speed. Describe how each of the following quantities changes as it follows this path: speed, velocity, kinetic energy, momentum, centripetal force, centripetal acceleration. (Refer to both magnitude and direction, as appropriate.)

Understanding circular motion

Isaac Newton devised an ingenious thought experiment that allows us to think about how an object can remain in a circular orbit around the Earth. Consider a large cannon on some high point on the Earth's surface, capable of firing objects horizontally. Figure 17.10 shows what will happen if we fire them at different speeds.

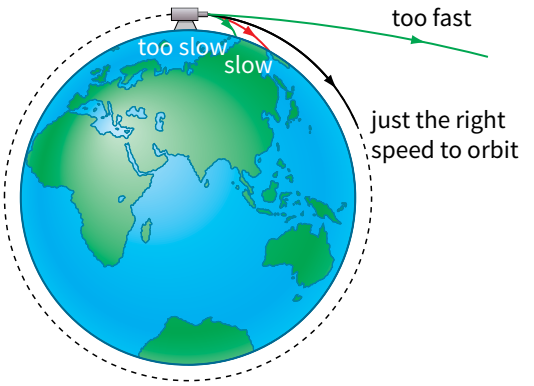


Figure 17.10 Newton's 'thought experiment'.

If the object is fired too slowly, gravity will pull it down towards the ground and it will land at some distance from the cannon. A faster initial speed results in the object landing further from the cannon.

Now, if we try a bit faster than this, the object will travel all the way round the Earth. We have to get just the right speed to do this. As the object is pulled down towards the Earth, the curved surface of the Earth falls away beneath it. The object follows a circular path, constantly falling under gravity but never getting any closer to the surface.

If the object is fired too fast, it travels off into space, and fails to get into a circular orbit. So we can see that there is just one correct speed to achieve a circular orbit under gravity. (Note that we have ignored the effects of air resistance in this discussion.)

Calculating acceleration and force

If we spin a bung around in a circle (Figure 17.7), we get a feeling for the factors which determine the centripetal force F required to keep it in its circular orbit. The greater the mass m of the bung and the greater its speed v, the greater is the force F that is required. However if the radius r of the circle is increased, F is smaller.

Now we will deduce an expression for the centripetal acceleration of an object moving around a circle with a constant speed. Figure 17.11 shows a particle moving round a circle. In time Δt it moves through an angle $\Delta\theta$ from A to B. Its speed remains constant but its velocity changes by Δv, as shown in the vector diagram. Since the narrow angle in this triangle is also $\Delta\theta$, we can say that:

$$\Delta\theta = \frac{\Delta v}{v}$$

Dividing both sides of this equation by Δt and rearranging gives:

$$\frac{\Delta v}{\Delta t} = \frac{v\Delta\theta}{\Delta t}$$

The quantity on the left is $\frac{\Delta v}{\Delta t} = a$, the particle's acceleration.

The quantity on the right is $\frac{\Delta\theta}{\Delta t} = \omega$, the angular velocity.

Substituting for these gives:

$$a = v\omega$$

Using $v = \omega r$, we can eliminate ω from this equation:

$$a = \frac{v^2}{r}$$

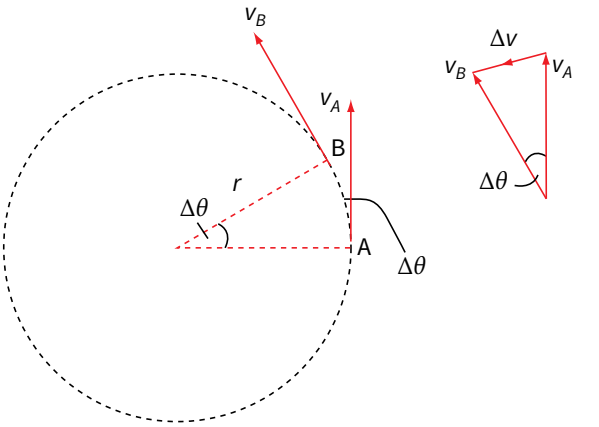


Figure 17.11 Deducing an expression for centripetal acceleration.

QUESTION

13 Show that an alternative equation for the centripetal acceleration is $a = \omega^2 r$.

Newton's second law of motion

Now that we have an equation for centripetal acceleration, we can use Newton's second law of motion to deduce an equation for centripetal force. If we write this law as $F = ma$, we find:

$$\text{centripetal force } F = \frac{mv^2}{r} = mr\omega^2$$

Remembering that an object accelerates in the direction of the resultant force on it, it follows that both F and a are in the same direction, towards the centre of the circle.

Calculating orbital speed

We can use the force equation to calculate the speed that an object must have to orbit the Earth under gravity, as in Newton's thought experiment. The necessary centripetal force $\frac{mv^2}{r}$ is provided by the Earth's gravitational pull mg. Hence:

$$mg = \frac{mv^2}{r}$$

$$g = \frac{v^2}{r}$$

where $g = 9.81\,\text{m s}^{-2}$ is the acceleration of free fall close to the Earth's surface. The radius of its orbit is equal to the Earth's radius, approximately 6400 km. Hence, we have:

$$9.81 = \sqrt{\frac{v^2}{(6.4 \times 10^6)}}$$

$$v = \sqrt{9.81 \times 6.4 \times 10^6} \approx 7.92 \times 10^3\,\text{ms}^{-1}$$

Thus if you were to throw or hit a ball horizontally at almost $8\,\text{km s}^{-1}$, it would go into orbit around the Earth.

The origins of centripetal forces

It is useful to look at one or two situations where the physical origin of the centripetal force may not be immediately obvious. In each case, you will notice that the forces acting on the moving object are not balanced – there is a resultant force. An object moving along a circular path is not in equilibrium and the resultant force acting on it is the centripetal force.

1 A car cornering on a level road (Figure 17.13). Here, the road provides two forces. The force N is the normal contact force which balances the weight mg of the car – the car has no acceleration in the vertical direction.

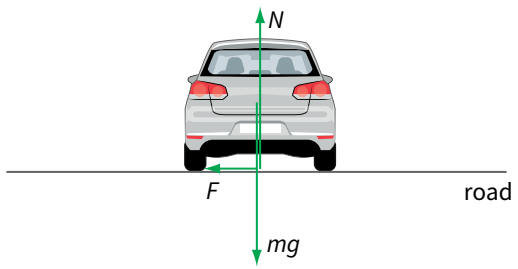


Figure 17.13 This car is moving away from us and turning to the left. Friction provides the centripetal force. N and F are the total normal contact and friction forces (respectively) provided by the contact of all four tyres with the road.

QUESTIONS

14 Calculate how long it would take a ball to orbit the Earth once, just above the surface, at a speed of $7920\,\text{m s}^{-1}$. (The radius of the Earth is 6400 km.)

15 A stone of mass 0.20 kg is whirled round on the end of a string of length 30 cm. The string will break when the tension in it exceeds 8.0 N. Calculate the maximum speed at which the stone can be whirled without the string breaking.



Figure 17.12 The view from the International Space Station, orbiting Earth over Australia.

16 The International Space Station (Figure 17.12) has a mass of 350 tonnes, and orbits the Earth at an average height of 340 km, where the gravitational acceleration is $8.8\,\text{m s}^{-2}$. The radius of the Earth is 6400 km. Calculate:

- **a** the centripetal force on the space station
- **b** the speed at which it orbits
- **c** the time taken for each orbit
- **d** the number of times it orbits the Earth each day.

17 A stone of mass 0.40 kg is whirled round on the end of a string 0.50 m long. It makes three complete revolutions each second. Calculate:

- **a** its speed
- **b** its centripetal acceleration
- **c** the tension in the string.

18 Mars orbits the Sun once every 687 days at a distance of 2.3×10^{11} m. The mass of Mars is 6.4×10^{23} kg. Calculate:

- **a** its orbital speed
- **b** its centripetal acceleration
- **c** the gravitational force exerted on Mars by the Sun.

The second force is the force of friction F between the tyres and the road surface. This is the unbalanced, centripetal force. If the road or tyres do not provide enough friction, the car will not go round the bend along the desired path. The friction between the tyres and the road provides the centripetal force necessary for the car's circular motion.

2 A car cornering on a banked road (Figure 17.14a). Here, the normal contact force N has a horizontal component which can provide the centripetal force. The vertical component of N balances the car's weight. Therefore:

vertically $N\cos\theta = mg$

horizontally $N\sin\theta = \frac{mv^2}{r}$

where r is the radius of the circular corner and v is the car's speed.

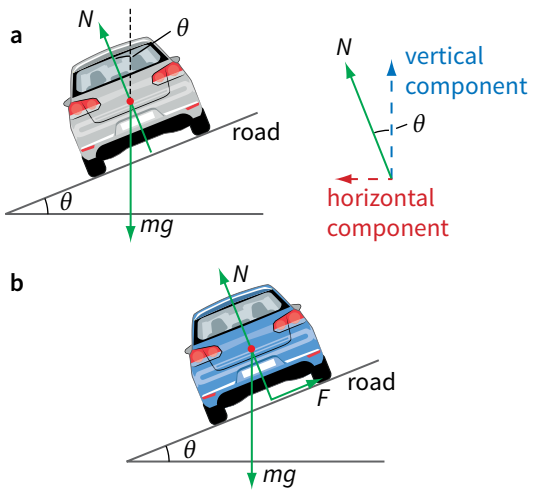


Figure 17.14 **a** On a banked road, the horizontal component of the normal contact force from the road can provide the centripetal force needed for cornering. **b** For a slow car, friction acts up the slope to stop it from sliding down.

If a car travels around the bend too slowly, it will tend to slide down the slope and friction will act up the slope to keep it on course (Figure 17.14b). If it travels too fast, it will tend to slide up the slope. If friction is insufficient, it will move up the slope and come off the road.

3 An aircraft banking (Figure 17.15a). To change direction, the pilot tips the aircraft's wings. The vertical component of the lift force L on the wings balances the weight. The horizontal component of L provides the centripetal force.

4 A stone being whirled in a horizontal circle on the end of a string – this arrangement is known as a conical pendulum (Figure 17.15b). The vertical component of the tension T is equal to the weight of the stone. The horizontal component of the tension provides the centripetal force for the circular motion.

5 At the fairground (Figure 17.15c). As the cylinder spins, the floor drops away. Friction balances your weight. The normal contact force of the wall provides the centripetal force. You feel as though you are being pushed back against the wall; what you are feeling is the push of the wall on your back.

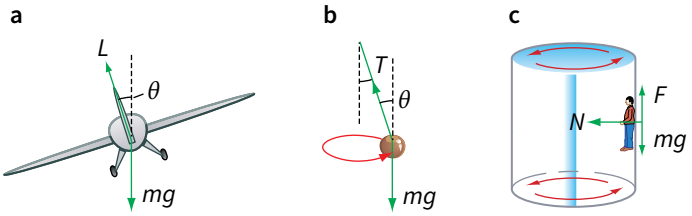


Figure 17.15 Three more ways of providing a centripetal force.

Note that the three situations shown in Figure 17.14a, Figure 17.15a and Figure 17.15b are equivalent. The moving object's weight acts downwards. The second force has a vertical component, which balances the weight, and a horizontal component, which provides the centripetal force.

QUESTIONS

19 Explain why it is impossible to whirl a bung around on the end of a string in such a way that the string remains perfectly horizontal.

20 Explain why an aircraft will tend to lose height when banking, unless the pilot increases its speed to provide more lift.

21 If you have ever been down a water-slide (a flume) (Figure 17.16) you will know that you tend to slide up the side as you go around a bend. Explain how this provides the centripetal force needed to push you around the bend. Explain why you slide higher if you are going faster.



Figure 17.16 A water-slide is a good place to experience centripetal forces.

Summary

- Angles can be measured in radians. An angle of 2π rad is equal to 360°.
- An object moving at a steady speed along a circular path has uniform circular motion.
- The angular displacement θ is a measure of the angle through which an object moves in a circle.
- The angular velocity ω is the rate at which the angular displacement changes: $\omega = \frac{\Delta\theta}{\Delta t}$
- For an object moving with uniform circular motion, speed and angular velocity are related by $v = \omega r$.
- An object moving in a circle is not in equilibrium; it has a resultant force acting on it.
- The resultant force acting on an object moving in a circle is called the centripetal force. This force is directed towards the centre of the circle and is at right angles to the velocity of the object.
- An object moving in a circle has a centripetal acceleration a given by:

$$a = \frac{v^2}{r} = r\omega^2$$

- The magnitude of the centripetal force F acting on an object of mass m moving at a speed v in a circle of radius r is given by:

$$F = \frac{mv^2}{r} = mr\omega^2$$

End-of-chapter questions

1 **a** Explain what is meant by a **radian**. [1]

b A body moves round a circle at a constant speed and completes one revolution in 15 s. Calculate the angular velocity of the body. [2]

2 Figure 17.17 shows part of the track of a roller-coaster ride in which a truck loops the loop. When the truck is at the position shown, there is no reaction force between the wheels of the truck and the track. The diameter of the loop in the track is 8.0 m.

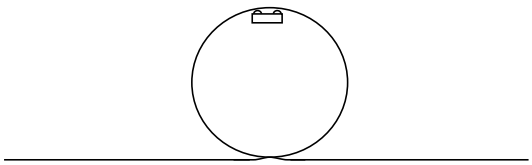


Figure 17.17 For End-of-chapter Question 2.

a Explain what provides the centripetal force to keep the truck moving in a circle. [1]

b Given that the acceleration due to gravity g is 9.8 m s^{-2}, calculate the speed of the truck. [3]

3 **a** Describe what is meant by **centripetal force**. [1]

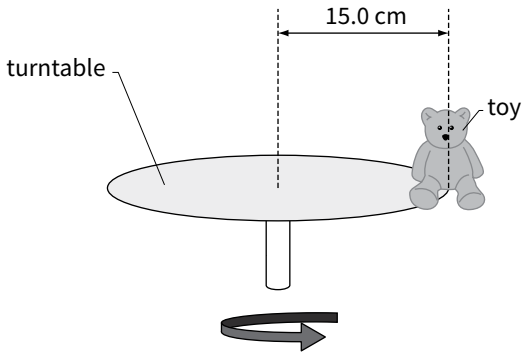


Figure 17.18 For End-of-chapter Question 3.

b Figure 17.18 shows a toy of mass 60 g placed on the edge of a rotating turntable.

i The diameter of the turntable is 15.0 cm. The turntable rotates, making 20 revolutions every minute. Calculate the centripetal force acting on the toy. [4]

ii Explain why the toy falls off when the speed of the turntable is increased. [2]

4 One end of a string is secured to the ceiling and a metal ball of mass 50 g is tied to its other end. The ball is initially at rest in the vertical position. The ball is raised through a vertical height of 70 cm (see Figure 17.19). The ball is then released. It describes a circular arc as it passes through the vertical position.

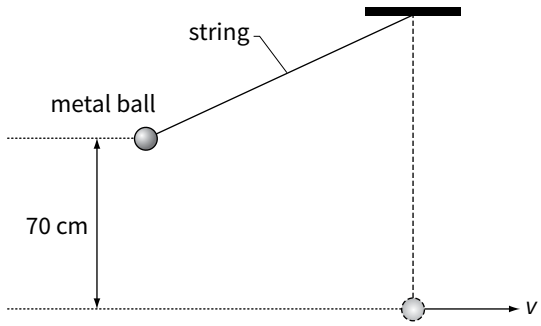


Figure 17.19 For End-of-chapter Question 4.

The length of the string is 1.50 m.

a Ignoring the effects of air resistance, determine the speed v of the ball as it passes through the vertical position. [2]

b Calculate the tension T in the string when the string is vertical. [4]

c Explain why your answer to **b** is not equal to the weight of the ball. [2]

5 A car is travelling round a bend when it hits a patch of oil. The car slides off the road onto the grass verge. Explain, using your understanding of circular motion, why the car came off the road. [2]

6 Figure 17.20 shows an aeroplane banking to make a horizontal turn. The aeroplane is travelling at a speed of 75 m s^{-1} and the radius of the turning circle is 80 m.

a Copy the diagram. On your copy, draw and label the forces acting on the aeroplane. [2]

b Calculate the angle which the aeroplane makes with the horizontal. [4]

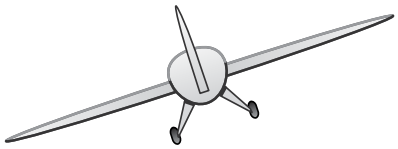


Figure 17.20 For End-of-chapter Question 6.

7 **a** Explain what is meant by the term **angular velocity**. [2]

b Figure 17.21 shows a rubber bung, of mass 200 g, on the end of a length of string being swung in a horizontal circle of radius 40 cm. The string makes an angle of 56° with the vertical.

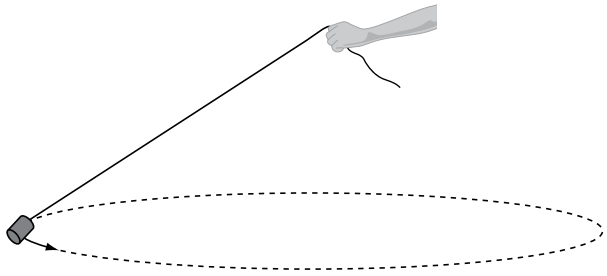


Figure 17.21 For End-of-chapter Question 7.

Calculate:

i the tension in the string [2]

ii the angular velocity of the bung [3]

iii the time it takes to make one complete revolution. [1]

8 **a** Explain what is meant by a **centripetal force**. [2]

b A teacher swings a bucket of water, of total mass 5.4 kg, round in a vertical circle of diameter 1.8 m.

i Calculate the minimum speed which the bucket must be swung at so that the water remains in the bucket at the top of the circle. [3]

ii Assuming that the speed remains constant, what will be the force on the teacher's hand when the bucket is at the bottom of the circle? [2]

9 In training, military pilots are given various tests. One test puts them in a seat on the end of a large arm which is then spun round at a high speed, as shown in Figure 17.22.

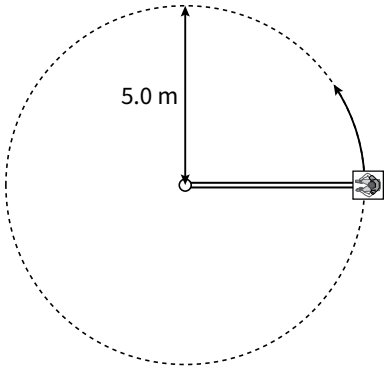


Figure 17.22 For End-of-chapter Question 9.

a Describe what the pilot will feel and relate this to the centripetal force. [3]

b At top speed the pilot will experience a centripetal force equivalent to six times his own weight (6 *mg*).

i Calculate the speed of the pilot in this test. [3]

ii Calculate the number of revolutions of the pilot per minute. [2]

c Suggest why it is necessary for pilots to be able to be able to withstand forces of this type. [2]

10 **a** Show that in one revolution there are 2π radians. [2]

b Figure 17.23 shows a centrifuge used to separate solid particles suspended in a liquid of lower density. The container is spun at a rate of 540 revolutions per minute.

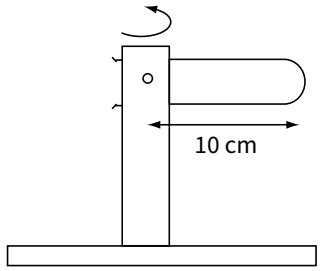


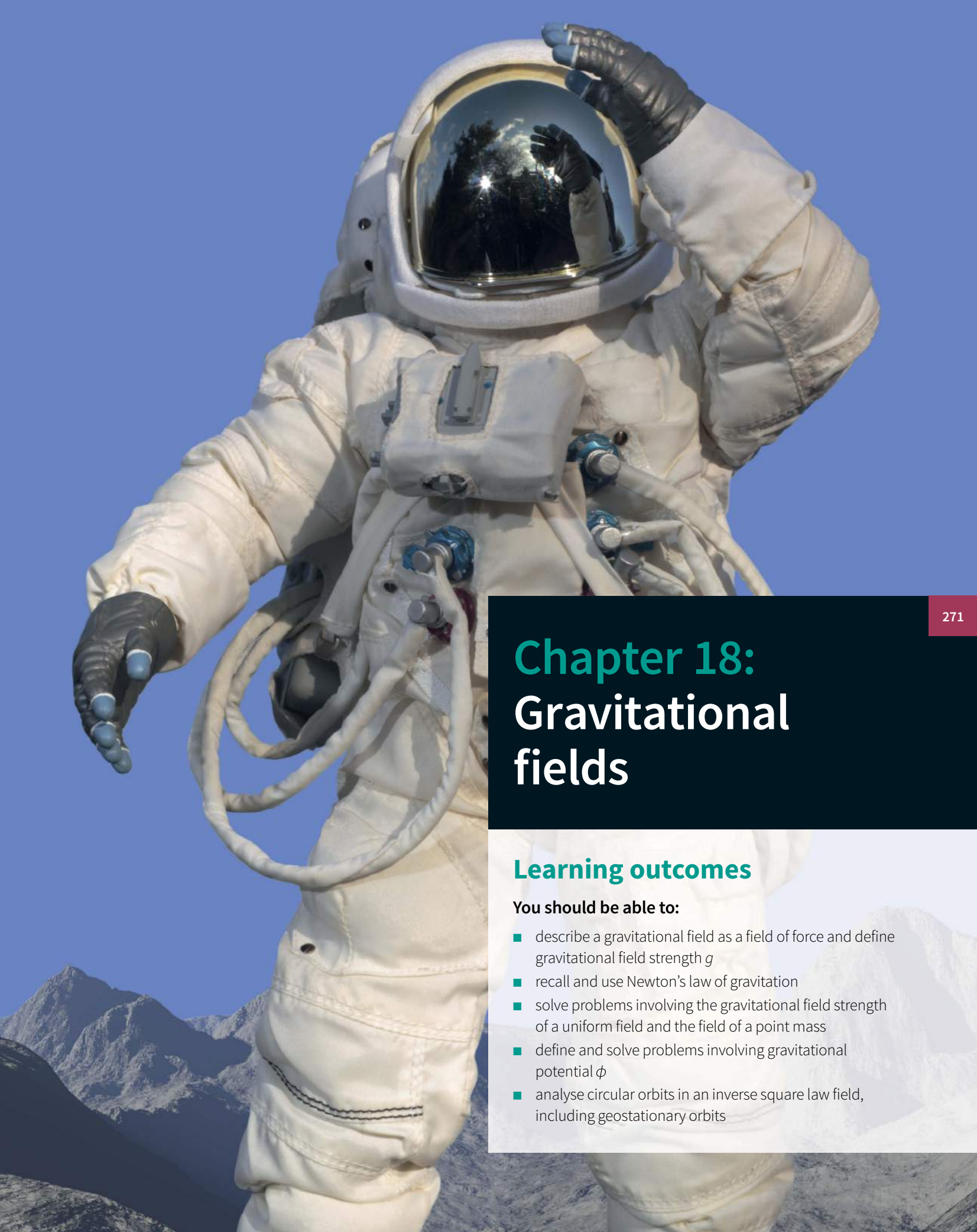
Figure 17.23 For End-of-chapter Question 10.

i Calculate the angular velocity of the container. [2]

ii Calculate the centripetal force on a particle of mass 20 mg at the end of the test tube. [2]

c An alternative method of separating the particles from the liquid is to allow them to settle to the bottom of a stationary container under gravity.

By comparing the forces involved, explain why the centrifuge is a more effective method of separating the mixture. [2]



Chapter 18: Gravitational fields

Learning outcomes

You should be able to:

- describe a gravitational field as a field of force and define gravitational field strength g
- recall and use Newton's law of gravitation
- solve problems involving the gravitational field strength of a uniform field and the field of a point mass
- define and solve problems involving gravitational potential ϕ
- analyse circular orbits in an inverse square law field, including geostationary orbits

Gravitational forces and fields



Figure 18.1 Skydivers balance the forces of gravity and air resistance.

We live our lives with the constant experience of gravity. We know that things fall when we drop them. The free-fall parachutists in Figure 18.1 are enjoying the experience of falling through the air under the influence of gravity.

The Earth's gravitational force extends well beyond its surface. The Moon stays in its orbit, at a distance of about 400 000 km away, because of the Earth's gravitational pull (Figure 18.2). The Earth orbits the Sun at a distance of 150 000 000 km because of the gravitational force between them.

According to Newton, all masses create a gravitational field in the space around them. This field gives rise to a force on any object having mass which is placed in this field. The Moon orbits the Earth because it experiences a gravitational force due to the Earth's gravitational field.

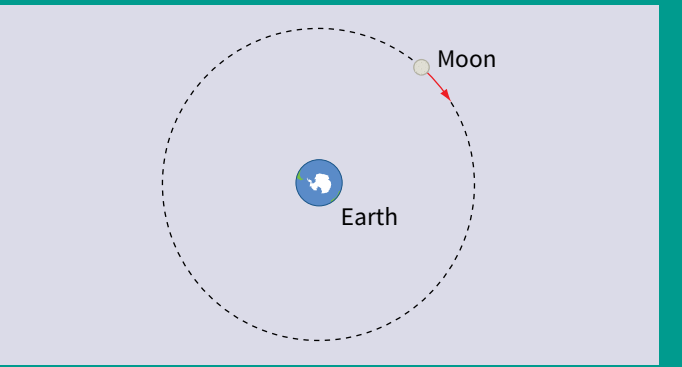


Figure 18.2 The Moon orbits the Earth. There is an attractive gravitational force acting on the Moon due to its mass and the mass of the Earth.

In physics, the idea of a field is a very general one. If an object is placed in a gravitational field, a force will act on the object because of its mass. In Chapter 8, we saw that a charged object experiences a force when it is placed in an electric field. You are probably familiar with the idea that a magnet produces a magnetic field around itself, and that this will produce a force (attractive or repulsive) on another magnet placed nearby.

Representing a gravitational field

In Chapter 8 we saw how to represent an electric field using field lines. In a similar way, we can represent the Earth's gravitational field by drawing field lines, as shown in Figure 18.3. The field lines show two things:

- The arrows on the field lines show us the direction of the gravitational force on a mass placed in the field.
- The spacing of the field lines indicates the strength of the gravitational field – the further apart they are, the weaker the field.

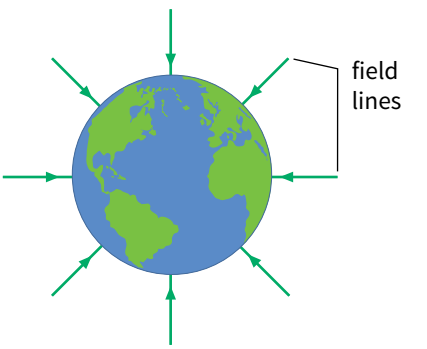


Figure 18.3 The Earth's gravitational field is represented by field lines.

The drawing of the Earth's gravitational field shows that all objects are attracted towards the centre of the Earth. This is true even if they are below the surface of the Earth. The gravitational force gets weaker as you get further away from the Earth's surface – this is shown by the greater separation between the field lines. The Earth is almost a uniform spherical mass – it bulges a bit at the equator. The gravitational field of the Earth is as if its entire mass was concentrated at its centre. As far as any object beyond the Earth's surface is concerned, the Earth behaves as a point mass.

Figure 18.4 shows the Earth's gravitational field closer to its surface. The gravitational field in and around a building on the Earth's surface shows that the gravitational force is directed downwards everywhere and (because the field lines are parallel and evenly spaced) the strength of the gravitational field is the same at all points in and around the building. This means that your weight is the same everywhere in this gravitational field. Your weight does not get significantly less when you go upstairs.

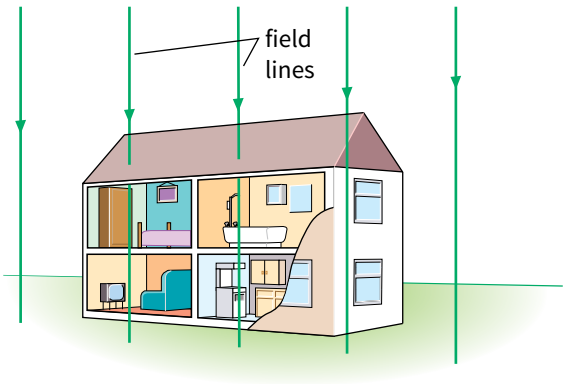


Figure 18.4 The Earth's gravitational field is uniform on the scale of a building.

We describe the Earth's gravitational field as **radial**, since the field lines diverge (spread out) radially from the centre of the Earth. However, on the scale of a building, the gravitational field is **uniform**, since the field lines are equally spaced. Jupiter is a more massive planet than the Earth and so we would represent its gravitational field by showing more closely spaced field lines.

Newton's law of gravitation

Newton used his ideas about mass and gravity to suggest a law of gravitation for two point masses (Figure 18.5). He considered two point masses M and m separated by a distance r. Each point mass attracts the other with a force F. (According to Newton's third law of motion, the point masses interact with each other and therefore exert equal but opposite forces on each other.) A statement of **Newton's law of gravitation** is shown below.

> Any two point masses attract each other with a force that is directly proportional to the product of their masses and inversely proportional to the square of their separation.

Note that the law refers to 'point masses' – you can alternatively use the term 'particles'. Things are more complicated if we think about solid bodies which occupy a volume of space. Each particle of one body attracts every particle of the other body and we would have to add all these forces together to work out the force each body has on the other. Newton was able to show that two uniform spheres attract one another with a force which is the same as if their masses were concentrated at their centres (provided their radii are much smaller than their separation).

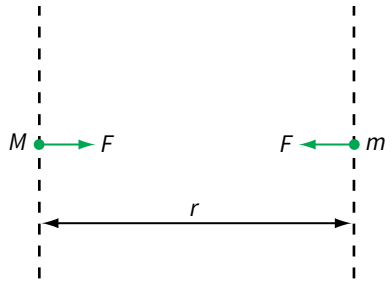


Figure 18.5 Two point masses separated by distance r.

According to Newton's law of gravitation, we have:

$$\text{force} \propto \text{product of the masses} \quad \text{or} \quad F \propto Mm$$

$$\text{force} \propto \frac{1}{\text{distance}^2} \quad \text{or} \quad F \propto \frac{1}{r^2}$$

Therefore:

$$F \propto \frac{Mm}{r^2}$$

To make this into an equation, we introduce the gravitational constant G:

$$F = \frac{GMm}{r^2}$$

(The force is attractive, so F is in the opposite direction to r.) The gravitational constant G is sometimes referred to as the universal gravitational constant because it is believed to have the same value, $6.67 \times 10^{-11}\,\text{N m}^2\text{kg}^{-2}$, throughout the Universe. This is important for our understanding of the history and likely long-term future of the Universe.

The equation above can also be applied to spherical objects (such as the Earth and the Moon) provided we remember to measure the separation r between the centres of the objects. You may also come across the equation in the form:

$$F = \frac{Gm_1m_2}{r^2}$$

where m_1 and m_2 are the masses of the two bodies.

Let us examine this equation to see why it seems reasonable. First, each of the two masses is important. Your weight (the gravitational force on you) depends on your mass and on the mass of the planet you happen to be standing on.

Second, the further away you are from the planet, the weaker its pull. Twice as far away gives one-quarter of

the force. This can be seen from the diagram of the field lines in Figure 18.6. If the distance is doubled, the lines are spread out over four times the surface area, so their concentration is reduced to one-quarter. This is called an inverse square law – you may have come across a similar law for radiation such as light or γ-rays spreading out uniformly from a point source.

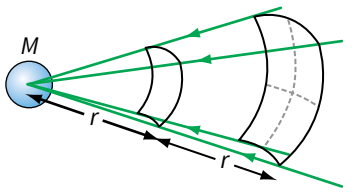


Figure 18.6 Field lines are spread out over a greater surface area at greater distances, so the strength of the field is weaker.

We measure distances from the centre of mass of one body to the centre of mass of the other (Figure 18.7). We treat each body as if its mass were concentrated at one point. The two bodies attract each other with equal and opposite forces, as required by Newton's third law of motion. The Earth pulls on you with a force (your weight) directed towards the centre of the Earth; you attract the Earth with an equal force, directed away from its centre and towards you. Your pull on an object as massive as the Earth has little effect on it. The Sun's pull on the Earth, however, has a very significant effect.

QUESTIONS

1 Calculate the gravitational force of attraction between:
   - **a** two objects separated by a distance of 1.0 cm and each having a mass of 100 g
   - **b** two asteroids separated by a distance of 4.0×10^9 m and each having a mass of 5.0×10^{10} kg
   - **c** a satellite of mass 1.4×10^4 kg orbiting the Earth at a distance of 6800 km from the Earth's centre. (The mass of the Earth is 6.0×10^{24} kg.)

2 Estimate the gravitational force of attraction between two people sitting side by side on a park bench. How does this force compare with the gravitational force exerted on each of them by the Earth, i.e. their weight?

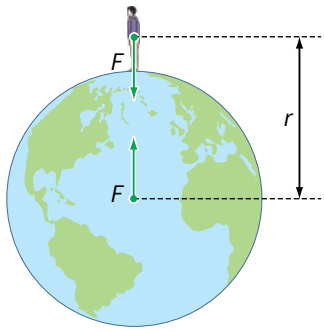


Figure 18.7 A person and the Earth exert equal and opposite attractive forces on each other.

Gravitational field strength *g*

We can describe how strong or weak a gravitational field is by stating its **gravitational field strength**. We are used to this idea for objects on or near the Earth's surface. The gravitational field strength is the familiar quantity *g*. Its value is approximately 9.81 m s^{-2}. The weight of a body of mass *m* is *mg*.

To make the meaning of *g* clearer, we should write it as 9.81 N kg^{-1}. That is, each 1 kg of mass experiences a gravitational force of 9.81 N.

The gravitational field strength *g* at any point in a gravitational field is defined as follows:

> The gravitational field strength at a point is the gravitational force exerted per unit mass on a small object placed at that point.

This can be written as an equation:

$$g = \frac{F}{m}$$

where *F* is the gravitational force on the object and *m* is the mass of the object. Gravitational field strength has units of N kg^{-1}. This is equivalent to m s^{-2}.

We can use the definition above to determine the gravitational field strength for a point (or spherical) mass. The force between two point masses is given by:

$$F = \frac{GMm}{r^2}$$

The gravitational field strength *g* due to the mass *M* at a distance of *r* from its centre is thus:

$$g = \frac{F}{m} = \frac{GM\cancel{m}}{r^2\cancel{m}}$$

or

$$g = \frac{GM}{r^2}$$

Since force is a vector quantity, it follows that gravitational field strength is also a vector. We need to give its direction as well as its magnitude in order to specify it completely. The field strength g is not a constant; it decreases as the distance r increases. The field strength obeys an inverse square law with distance. The field strength will decrease by a factor of four when the distance from the centre is doubled. Close to the Earth's surface, the magnitude of g is about $9.81\,\text{N kg}^{-1}$. Even if you climbed Mount Everest, which is 8.85 km high, the field strength will only decrease by 0.3%.

So the gravitational field strength g at a point depends on the mass M of the body causing the field, and the distance r from its centre (see Worked example 1).

Gravitational field strength g also has units m s^{-2}; it is an acceleration. Another name for g is 'acceleration of free fall'. Any object that falls freely in a gravitational field has this acceleration, approximately $9.81\,\text{m s}^{-2}$ near the Earth's surface. In Chapter 2, you learned about different ways to determine an experimental value for g, the local gravitational field strength.

WORKED EXAMPLE

1 The Earth has radius 6400 km. The gravitational field strength on the Earth's surface is $9.81\,\text{N kg}^{-1}$. Use this information to determine the mass of the Earth and its mean density.

Step 1 Write down the quantities given:

$r = 6.4 \times 10^6\,\text{m} \quad g = 9.81\,\text{N kg}^{-1}$

Step 2 Use the equation $g = \frac{GM}{r^2}$ to determine the mass of the Earth.

$$g = \frac{GM}{r^2}$$

$$9.8 = \frac{6.67 \times 10^{-11}\,M}{(6.4 \times 10^6)^2}$$

$$\text{mass of Earth} = M = 9.71 \times \frac{(6.4 \times 10^6)^2}{6.67 \times 10^{-11}}$$

$$= 6.0 \times 10^{24}\,\text{kg}$$

Step 3 Use the equation $\text{density} = \frac{\text{mass}}{\text{volume}}$ to determine the density of the Earth.

The Earth is a spherical mass. Its volume can be calculated using $\frac{4}{3}\pi r^3$:

$$\text{density} = \rho = \frac{M}{V} = \frac{6.0 \times 10^{24}}{\frac{4}{3} \times \pi \times (6.4 \times 10^6)^3}$$

$$= 5500\,\text{kg m}^3$$

QUESTIONS

You will need the data in Table 18.1 to answer these questions.

Body	Mass / kg	Radius / km	Distance from Earth / km
Earth	6.0×10^{24}	6400	–
Moon	7.4×10^{22}	1740	3.8×10^5
Sun	2.0×10^{30}	700 000	1.5×10^8

Table 18.1 Data for Questions 3–9.

3 Mount Everest is approximately 9.0 km high. Calculate how much less a mountaineer of mass 100 kg (including backpack) would weigh at its summit, compared to her weight at sea level. Would this difference be measurable with bathroom scales?

4 **a** Calculate the gravitational field strength:
- **i** close to the surface of the Moon
- **ii** close to the surface of the Sun.

b Suggest how your answers above help to explain why the Moon has only a thin atmosphere, while the Sun has a dense atmosphere.

5 **a** Calculate the Earth's gravitational field strength at the position of the Moon.

b Calculate the force the Earth exerts on the Moon. Hence determine the Moon's acceleration towards the Earth.

6 Jupiter's mass is 320 times that of the Earth and its radius is 11.2 times the Earth's. The Earth's surface gravitational field strength is $9.81\,\text{N kg}^{-1}$. Calculate the gravitational field strength close to the surface of Jupiter.

7 The Moon and the Sun both contribute to the tides on the Earth's oceans. Which has a bigger pull on each kilogram of seawater, the Sun or the Moon?

8 Astrologers believe that the planets exert an influence on us, particularly at the moment of birth. (They don't necessarily believe that this is an effect of gravity!)

a Calculate the gravitational force on a 4.0 kg baby caused by Mars when the planet is at its closest to the Earth at a distance of 100 000 000 km. Mars has mass 6.4×10^{23} kg.

b Calculate the gravitational force on the same baby due to its 50 kg mother at a distance of 0.40 m.

QUESTION

9 There is a point on the line joining the centres of the Earth and the Moon where their combined gravitational field strength is zero. Is this point closer to the Earth or to the Moon? Calculate how far it is from the centre of the Earth.

Energy in a gravitational field

As well as the force on a mass in a gravitational field, we can think about its energy. If you lift an object from the ground, you increase its gravitational potential energy (g.p.e.). The higher you lift it, the more work you do on it and so the greater its g.p.e. The object's change in g.p.e. can be calculated as $mg\Delta h$, where Δh is the change in its height (as we saw in Chapter 5).

This approach is satisfactory when we are considering objects close to the Earth's surface. However, we need a more general approach to calculating gravitational energy, for two reasons:

- If we use g.p.e. = $mg\Delta h$, we are assuming that an object's g.p.e. is zero on the Earth's surface. This is fine for many practical purposes but not, for example, if we are considering objects moving through space, far from Earth. For these, there is nothing special about the Earth's surface.
- If we lift an object to a great height, g decreases and we would need to take this into account when calculating g.p.e.

For these reasons, we need to set up a different way of thinking about gravitational potential energy. We start by picturing a mass at infinity, that is, at an infinite distance from all other masses. We say that here the mass has zero potential energy. This is a more convenient way of defining the zero of g.p.e. than using the surface of the Earth.

Now we picture moving the mass to the point where we want to know its g.p.e. As with lifting an object from the ground, we determine the work done to move the mass to the point. The work done on it is equal to the energy transferred to it, i.e. its g.p.e., and that is how we can determine the g.p.e. of a particular mass.

Gravitational potential

In practice, it is more useful to talk about the gravitational potential at a point. This tells us the g.p.e. per unit mass at the point (just as field strength g tells us the force per unit mass at a point in a field). The symbol used for potential is ϕ (Greek letter phi), and unit mass means one kilogram. **Gravitational potential** is defined as follows:

> The gravitational potential at a point is the work done per unit mass in bringing a mass from infinity to the point.

For a point mass M, we can write an equation for ϕ at a distance r from M:

$$\phi = -\frac{GM}{r}$$

where G is the gravitational constant as before. Notice the minus sign; gravitational potential is always negative. This is because, as a mass is brought towards another mass, its g.p.e. decreases. Since g.p.e. is zero at infinity, it follows that, anywhere else, g.p.e. and potential are less than zero, i.e. they are negative.

Picture a spacecraft coming from a distant star to visit the solar system. The variation of the gravitational potential along its path is shown in Figure 18.8. We will concentrate on three parts of its journey:

1. As the craft approaches the Earth, it is attracted towards it. The closer it gets to Earth, the lower its g.p.e. becomes and so the lower its potential.
2. As it moves away from the Earth, it has to work against the pull of the Earth's gravity. Its g.p.e. increases and so we can say that the potential increases. The Earth's gravitational field creates a giant 'potential well' in space. We live at the bottom of that well.
3. As it approaches the Sun, it is attracted into a much deeper well. The Sun's mass is much greater than the Earth's and so its pull is much stronger and the potential at its surface is more negative than on the Earth's surface.

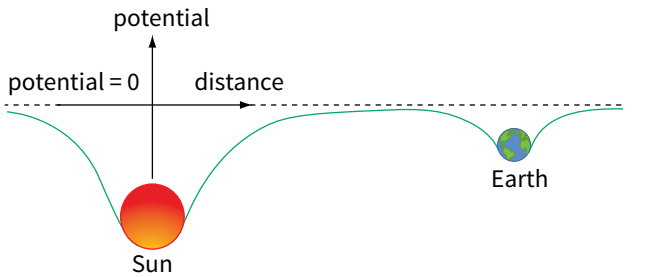


Figure 18.8 The gravitational potential is zero at infinity (far from any mass), and decreases as a mass is approached.

WORKED EXAMPLE

2 A planet has a diameter of 6800 km and a mass of 4.9×10^{23} kg. A rock of mass 200 kg, initially at rest and a long distance from the planet, accelerates towards the planet and hits the surface of the planet.

Calculate the change in potential energy of the rock and its speed when it hits the surface.

Step 1 Write down the quantities given.

$r = 3.4 \times 10^6$ m $\quad M = 4.9 \times 10^{23}$ kg

Step 2 The equation $\varphi = -\frac{GM}{r}$ gives the potential at the surface of the planet, that is, the gravitational potential energy per unit mass at that point. So the gravitational potential energy of the rock of mass m at that point is given by:

$$\text{g.p.e.} = -\frac{GMm}{r}$$

The g.p.e. of the rock when it is far away is zero, so the value we calculate using this equation gives the decrease in the rock's g.p.e. during its fall to hit the planet.

$$\text{change in g.p.e.} = \frac{6.67 \times 10^{-11} \times 4.9 \times 10^{23} \times 200}{3.4 \times 10^6}$$

$$= 1.92 \times 10^9 \text{ J} \approx 1.9 \times 10^9 \text{ J}$$

Step 3 In the absence of an atmosphere, all of the g.p.e. becomes kinetic energy of the rock, and so:

$$\tfrac{1}{2}mv^2 = 1.92 \times 10^9 \text{ J}$$

$$v = \sqrt{\frac{1.92 \times 10^9 \times 2}{200}} = 4400 \text{ m s}^{-1}$$

Note that the rock's final speed when it hits the planet does not depend on the mass of the rock. This is because, if you equate the two equations for k.e. and the change in g.p.e., the mass m of the rock cancels.

QUESTION

You will need the data for the mass and radius of the Earth and the Moon from Table 18.1 to answer this question.

Gravitational constant $G = 6.67 \times 10^{-11}$ N m^2 kg^{-2}.

10 **a** Determine the gravitational potential at the surface of the Earth.

b Determine the gravitational potential at the surface of the Moon.

c Which has the shallower 'potential well', the Earth or the Moon? Draw a diagram similar to Figure 18.8 to compare the 'potential wells' of the Earth and the Moon.

d Use your diagram to explain why a large rocket is needed to lift a spacecraft from the surface of the Earth but a much smaller rocket can be used to launch from the Moon's surface.

Fields – terminology

The words used to describe gravitational (and other) fields can be confusing. Remember:

- **Field strength** tells us about the **force** on unit mass at a point;
- **Potential** tells us about **potential energy** of unit mass at a point.

You have already learned about field strength in connection with electric fields, where it is the force on unit charge. Similarly, when we talk about the potential difference between two points in electricity, we are talking about the difference in electrical potential energy per unit charge. You will learn more about this in Chapter 24.

Orbiting under gravity

For an object orbiting a planet, such as an artificial satellite orbiting the Earth, gravity provides the centripetal force which keeps it in orbit (Figure 18.9). This is a simple situation as there is only one force acting on the satellite – the gravitational attraction of the Earth. The satellite follows a circular path because the gravitational force is at right angles to its velocity.

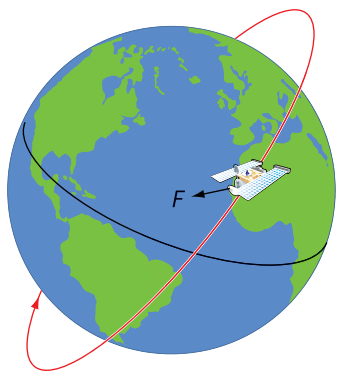


Figure 18.9 The gravitational attraction of the Earth provides the centripetal force on an orbiting satellite.

From Chapter 17, you know that the centripetal force F on a body is given by:

$$F = \frac{mv^2}{r}$$

Consider a satellite of mass m orbiting the Earth at a distance r from the Earth's centre at a constant speed v. Since it is the gravitational force between the Earth and the satellite which provides this centripetal force, we can write:

$$\frac{GMm}{r^2} = \frac{mv^2}{r}$$

where M is the mass of the Earth. (There is no need for a minus sign here as the gravitational force and the centripetal force are both directed towards the centre of the circle.)

Rearranging gives:

$$v^2 = \frac{GM}{r}$$

This equation allows us to calculate, for example, the speed at which a satellite must travel to stay in a circular orbit. Notice that the mass of the satellite m has cancelled out. The implication of this is that all satellites, whatever their masses, will travel at the same speed in a particular orbit. You would find this very reassuring if you were an astronaut on a space walk outside your spacecraft (Figure 18.10). You would travel at the same speed as your craft, despite the fact that your mass is a lot less than its mass. The equation above can be applied to the planets of our solar system – M becomes the mass of the Sun.



Figure 18.10 During this space walk, both the astronaut and the spacecraft travel through space at over 8 km s^{-1}.

Now look at Worked example 2.

WORKED EXAMPLE

3 The Moon orbits the Earth at an average distance of 384 000 km from the centre of the Earth. Calculate its orbital speed. (The mass of the Earth is 6.0×10^{24} kg.)

Step 1 Write down the known quantities.

$r = 3.84 \times 10^8$ m $\quad M = 6.0 \times 10^{24}$ kg $\quad v = ?$

Step 2 Use the equation $v^2 = \frac{GM}{r}$ to determine the orbital speed v.

$$v^2 = \frac{GM}{r}$$

$$v^2 = \frac{6.67 \times 10^{-11} \times 6.0 \times 10^{24}}{3.84 \times 10^8}$$

$$v^2 = 1.04 \times 10^6$$

Hint: *Don't forget to take the square root of v^2 to get v.*

$v = 1020\ \text{m s}^{-1} \approx 1.0 \times 10^3\ \text{m s}^{-1}$

So the Moon travels around its orbit at a speed of roughly 1 km s^{-1}.

QUESTION

11 Calculate the orbital speed of an artificial satellite travelling 200 km above the Earth's surface. (The radius of Earth is 6.4×10^6 m and its mass is 6.0×10^{24} kg.)

The orbital period

It is often more useful to consider the time taken for a complete orbit, the orbital **period** T. Since the distance around an orbit is equal to the circumference $2\pi r$, it follows that:

$$v = \frac{2\pi r}{T}$$

We can substitute this in the equation for v^2.
This gives:

$$\frac{4\pi^2 r^2}{T^2} = \frac{GM}{r}$$

and rearranging this equation gives:

$$T^2 = \left(\frac{4\pi^2}{GM}\right) r^3$$

This equation shows that the orbital period T is related to the radius r of the orbit. The square of the period is directly proportional to the cube of the radius ($T^2 \propto r^3$). This is

an important result. It was first discovered by Johannes Kepler, who analysed the available data for the planets of the solar system. It was an empirical law (one based solely on experiment) since he had no theory to explain why there should be this relationship between T and r. It was not until Isaac Newton formulated his law of gravitation that it was possible to explain this fact.

Orbiting the Earth

The Earth has one natural satellite – the Moon – and many thousands of artificial satellites – some spacecraft and a lot of debris. Each of these satellites uses the Earth's gravitational field to provide the centripetal force that keeps it in orbit. In order for a satellite to maintain a particular orbit, it must travel at the correct speed. This is given by the equation on page 278:

$$v^2 = \frac{GM}{r}$$

It follows from this equation that, the closer the satellite is to the Earth, the faster it must travel. If it travels too slowly, it will fall down towards the Earth's surface. If it travels too quickly, it will move out into a higher orbit.

QUESTION

12 A satellite orbiting a few hundred kilometres above the Earth's surface will experience a slight frictional drag from the Earth's (very thin) atmosphere. Draw a diagram to show how you would expect the satellite's orbit to change as a result. How can this problem be overcome if it is desired to keep a satellite at a particular height above the Earth?

Observing the Earth

Artificial satellites have a variety of uses. Many are used for making observations of the Earth's surface for commercial, environmental, meteorological or military purposes. Others are used for astronomical observations, benefiting greatly from being above the Earth's atmosphere. Still others are used for navigation, telecommunications and broadcasting.

Figure 18.11 shows two typical orbits. A satellite in a circular orbit close to the Earth's surface, and passing over the poles, completes about 16 orbits in 24 hours. As the Earth turns below it, the satellite 'sees' a different strip of the Earth's surface during each orbit. A satellite in an elliptical orbit has a more distant view of the Earth.

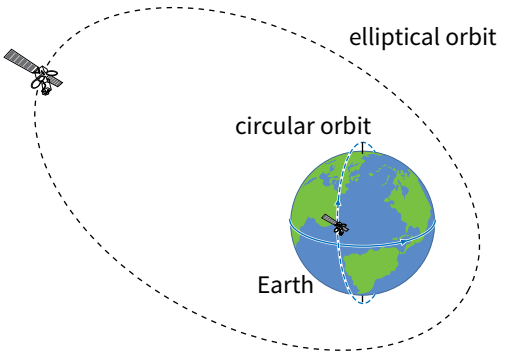


Figure 18.11 Satellites orbiting the Earth.

Geostationary orbits

A special type of orbit is one in which a satellite is positioned so that, as it orbits, the Earth rotates below it at the same rate. The satellite remains above a fixed point on the Earth's equator. This kind of orbit is called a **geostationary orbit**. There are over 300 satellites in such orbits. They are used for telecommunications (transmitting telephone messages around the world) and for satellite television transmission. A base station on Earth sends the TV signal up to the satellite, where it is amplified and broadcast back to the ground. Satellite receiver dishes are a familiar sight; you will have observed how, in a neighbourhood, they all point towards the same point in the sky. Because the satellite is in a geostationary orbit, the dish can be fixed. Satellites in any other orbits move across the sky so that a tracking system is necessary to communicate with them. Such a system is complex and expensive, and too demanding for the domestic market.

Geostationary satellites have a lifetime of perhaps ten years. They gradually drift out of the correct orbit, so they need a fuel supply for the rocket motors which return them to their geostationary position, and which keep them pointing correctly towards the Earth. Eventually they run out of fuel and need to be replaced.

We can determine the distance of a satellite in a geostationary orbit using the equation:

$$T^2 = \left(\frac{4\pi^2}{GM}\right) r^3$$

For a satellite to stay above a fixed point on the equator, it must take exactly 24 hours to complete one orbit (Figure 18.12). We know:

$G = 6.67 \times 10^{-11}\,\text{N m}^2\,\text{kg}^{-2}$

$T = 24\text{ hours} = 86\,400\,\text{s}$

$M = 6.0 \times 10^{24}\,\text{kg}$

$$r^3 = \frac{GMT^2}{4\pi^2} = \frac{6.67 \times 10^{-11} \times 6.0 \times 10^{24} \times (86\,400)^2}{4\pi^2}$$

$$= 7.66 \times 10^{22}\,\text{m}^3$$

$$r = \sqrt[3]{7.66 \times 10^{22}} \approx 4.23 \times 10^7\,\text{m}$$

So, for a satellite to occupy a geostationary orbit, it must be at a distance of 42 300 km from the centre of the Earth and at a point directly above the equator. Note that the radius of the Earth is 6400 km, so the orbital radius is 6.6 Earth radii from the centre of the Earth (or 5.6 Earth radii from its surface). Figure 18.12 has been drawn to give an impression of the size of the orbit.

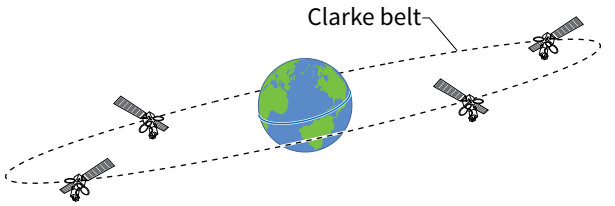


Figure 18.12 Geostationary satellites are parked in the 'Clarke belt', high above the equator. This is a perspective view; the Clarke belt is circular.

QUESTIONS

13 For any future mission to Mars, it would be desirable to set up a system of three or four geostationary (or 'martostationary') satellites around Mars to allow communication between the planet and Earth. Calculate the radius of a suitable orbit around Mars.

Mars has mass 6.4×10^{23} kg and a rotational period of 24.6 hours.

14 Although some international telephone signals are sent via satellites in geostationary orbits, most are sent along cables on the Earth's surface. This reduces the time delay between sending and receiving the signal. Estimate this time delay for communication via a satellite, and explain why it is less significant when cables are used.

You will need the following:

- radius of geostationary orbit = 42 300 km
- radius of Earth = 6400 km
- speed of electromagnetic waves in free space $c = 3.0 \times 10^8\,\text{m s}^{-1}$

Summary

- The force of gravity is an attractive force between any two objects due to their masses.
- The gravitational field strength g at a point is the gravitational force exerted per unit mass on a small object placed at that point – that is:

  $g = \frac{F}{m}$

- The external field of a uniform spherical mass is the same as that of an equal point mass at the centre of the sphere.
- Newton's law of gravitation states that:

  Any two point masses attract each other with a force that is directly proportional to the product of their masses and inversely proportional to the square of their separation.
- The equation for Newton's law of gravitation is:

  $F = \frac{GMm}{r^2}$

- The gravitational field strength at a point is the gravitational force exerted per unit mass on a small object placed at that point:
  $g = \frac{GM}{r^2}$
- On or near the surface of the Earth, the gravitational field is uniform, so the value of g is approximately constant. Its value is equal to the acceleration of free fall.
- The gravitational potential at a point is the work done in bringing unit mass from infinity to that point.
- The gravitational potential of a point mass is given by:
  $\phi = -\frac{GM}{r}$
- The orbital period of a satellite is the time taken for one orbit.
- The orbital period can be found by equating the gravitational force $\frac{GMm}{r^2}$ to the centripetal force $\frac{mv^2}{r}$.
- The orbital speed of a planet or satellite can be determined using the equation:
  $v^2 = \frac{GM}{r}$
- Geostationary satellites have an orbital period of 24 hours and are used for telecommunications transmissions and for television broadcasting.

End-of-chapter questions

1 Two small spheres each of mass 20 g hang side by side with their centres 5.00 mm apart. Calculate the gravitational attraction between the two spheres. [3]

2 It is suggested that the mass of a mountain could be measured by the deflection from the vertical of a suspended mass. Figure 18.13 shows the principle.

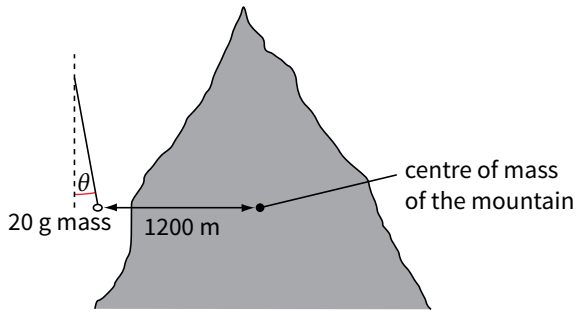


Figure 18.13 For End-of-chapter Question 2.

a Copy Figure 18.13 and draw arrows to represent the forces acting on the mass. Label the arrows. [2]

b The whole mass of the mountain, 3.8×10^{12} kg, may be considered to act at its centre of mass. Calculate the horizontal force on the mass due to the mountain. [2]

c Compare the force calculated in **b** with the Earth's gravitational force on the mass. [2]

3 Figure 18.14 shows the Earth's gravitational field.

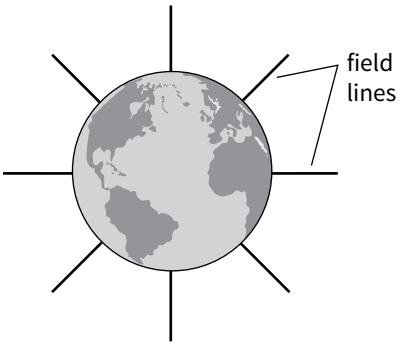


Figure 18.14 For End-of-chapter Question 3.

a Copy the diagram and add arrows to show the direction of the field. [1]

b Explain why the formula for potential energy gained ($mg\Delta h$) can be used to find the increase in potential energy when an aircraft climbs to a height of 10 000 m, but cannot be used to calculate the increase in potential energy when a spacecraft travels from the Earth's surface to a height of 10 000 km. [2]

4 Mercury, the smallest of the eight recognised planets, has a diameter of 4.88×10^6 m and a mean density of 5.4×10^3 kg m^{-3}.

a Calculate the gravitational field at its surface. [4]

b A man has a weight of 900 N on the Earth's surface. What would his weight be on the surface of Mercury? [2]

5 Calculate the potential energy of a spacecraft of mass 250 kg when it is 20 000 km from the planet Mars. (Mass of Mars = 6.4×10^{23} kg, radius of Mars = 3.4×10^6 m.) [2]

6 Ganymede is the largest of Jupiter's moons, with a mass of 1.48×10^{23} kg. It orbits Jupiter with an orbital radius of 1.07×10^6 km and it rotates on its own axis with a period of 7.15 days. It has been suggested that to monitor an unmanned landing craft on the surface of Ganymede a geostationary satellite should be placed in orbit around Ganymede.

a Calculate the orbital radius of the proposed geostationary satellite. [2]

b Suggest a difficulty that might be encountered in achieving a geostationary orbit for this moon. [1]

7 The Earth orbits the Sun with a period of 1 year at an orbital radius of 1.50×10^{11} m. Calculate:

a the orbital speed of the Earth [3]

b the centripetal acceleration of the Earth [2]

c the Sun's gravitational field strength at the Earth. [1]

8 The planet Mars has a mass of 6.4×10^{23} kg and a diameter of 6790 km.

a **i** Calculate the acceleration due to gravity at the planet's surface. [2]

ii Calculate the gravitational potential at the surface of the planet. [2]

b A rocket is to return some samples of Martian material to Earth. Write down how much energy each kilogram of matter must be given to escape completely from Mars's gravitational field. [1]

c Use you answer to **b** to show that the minimum speed that the rocket must reach to escape from the gravitational field is 5000 m s^{-1}. [2]

d Suggest why it has been proposed that, for a successful mission to Mars, the craft that takes the astronauts to Mars will be assembled at a space station in Earth orbit and launched from there, rather than from the Earth's surface. [2]

9 **a** Explain what is meant by the **gravitational potential at a point**. [2]

b Figure 18.15 shows the gravitational potential near a planet of mass M and radius R.

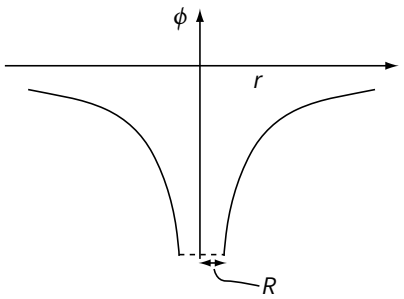


Figure 18.15 For End-of-chapter Question 9.

On a copy of the diagram, draw similar curves:

i for a planet of the same radius but of mass $2M$ – label this **i**. [2]

ii for a planet of the same mass but of radius $2R$ – label this **ii**. [2]

c Use the graphs to explain from which of these three planets it would require the least energy to escape. [2]

d Venus has a diameter of 12 100 km and a mass of 4.87×10^{24} kg.

Calculate the energy needed to lift one kilogram from the surface of Venus to a space station in orbit 900 km from the surface. [4]

10 **a** Explain what is meant by the **gravitational field strength at a point**. [2]

Figure 18.16 shows the dwarf planet, Pluto, and its moon, Charon. These can be considered to be a double planetary system orbiting each other about their joint centre of mass.

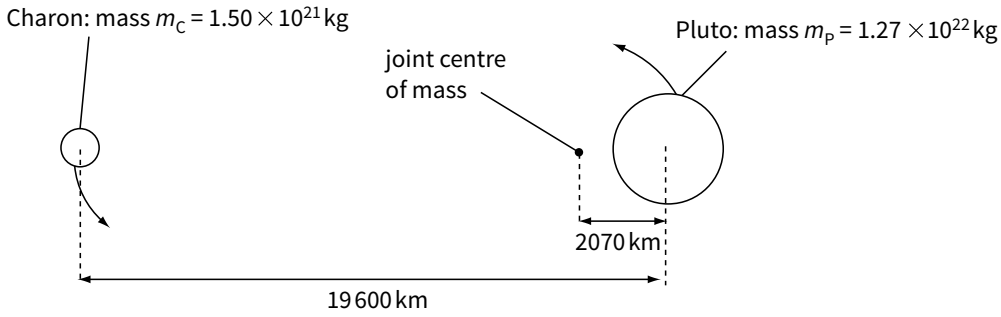


Figure 18.16 For End-of-chapter Question 10.

b Calculate the gravitational pull on Charon due to Pluto. [3]

c Use your result to **b** to calculate Charon's orbital period. [3]

d Explain why Pluto's orbital period must be the same as Charon's. [1]

11 Figure 18.17 shows the variation of the Earth's gravitational field strength with distance from its centre.

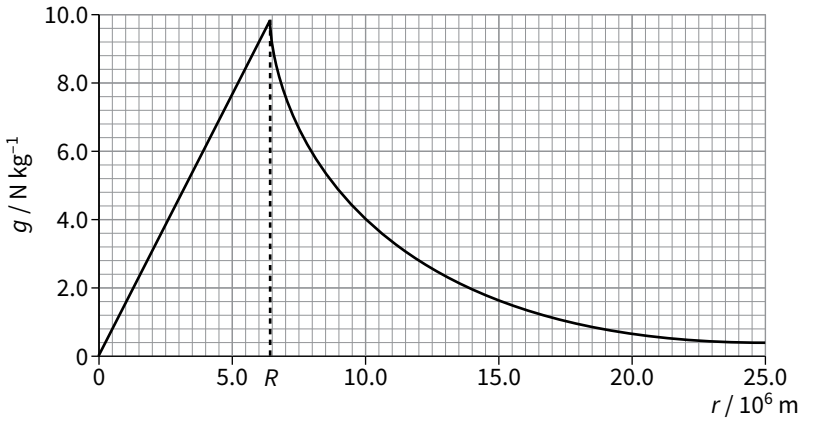


Figure 18.17 For End-of-chapter Question 11.

a Determine the gravitational field strength at a height equal to $2R$ above the Earth's surface, where R is the radius of the Earth. [1]

b A satellite is put into an orbit at this height. State the centripetal acceleration of the satellite. [1]

c Calculate the speed at which the satellite must travel to remain in this orbit. [2]

d **i** Frictional forces mean that the satellite gradually slows down after it has achieved a circular orbit. Draw a diagram of the initial circular orbital path of the satellite, and show the resulting orbit as frictional forces slow the satellite down. [1]

ii Suggest and explain why there is not a continuous bombardment of old satellites colliding with the Earth. [2]



Chapter 19: Oscillations

Learning outcomes

You should be able to:

- give examples of free and forced oscillations
- use appropriate terminology to describe oscillations
- use the equation $a = -\omega^2 x$ to define simple harmonic motion (s.h.m.)
- recall and use equations for displacement and velocity in s.h.m.
- draw and use graphical representations of s.h.m.
- describe energy changes during s.h.m.
- describe the effects of damping on oscillations, with practical examples
- give examples of forced oscillations and resonance and describe the effects of damping on resonance

Oscillations in an engine

Figure 19.1 shows a cut-away view of a modern car engine; there are four pistons which oscillate up to 5000 times per minute when the engine is operating at full power. Engineers need to understand the physics of oscillations to be able to calculate the stresses produced on the pistons when the engine is operating.



Figure 19.1 The pistons inside this car engine oscillate up and down as the engine powers the car.

Free and forced oscillations

Oscillations and vibrations are everywhere. A bird in flight flaps its wings up and down. An aircraft's wings also vibrate up and down, but this is not how it flies. The wings are long and thin, and they vibrate slightly because they are not perfectly rigid. Many other structures vibrate – bridges when traffic flows across, buildings in high winds.

A more specific term than vibration is **oscillation**. An object **oscillates** when it moves back and forth repeatedly, on either side of some equilibrium position. If we stop the object from oscillating, it returns to the equilibrium position.

We make use of oscillations in many different ways – for pleasure (a child on a swing), for music (the vibrations of a guitar string), for timing (the movement of a pendulum or the vibrations of a quartz crystal). Whenever we make a sound, the molecules of the air oscillate, passing the sound energy along. The atoms of a solid vibrate more and more as the temperature rises.

These examples of oscillations and vibrations may seem very different from one another. In this chapter, we will look at the characteristics that are shared by all oscillations.

Free or forced?

The easiest oscillations to understand are free oscillations. If you pluck a guitar string, it continues to vibrate for some time after you have released it. It vibrates at a particular frequency (the number of vibrations per unit time). This is called its **natural frequency** of vibration, and it gives rise to the particular note that you hear. Change the length of the string, and you change the natural frequency. In a similar way, the prongs of a tuning fork have a natural frequency of vibration, which you can observe when you strike it on a cork. Every oscillator has a natural frequency of vibration, the frequency with which it vibrates freely after an initial disturbance.

On the other hand, many objects can be forced to vibrate. If you sit on a bus, you may notice that the vibrations from the engine are transmitted to your body, causing you to vibrate with the same frequency. These are not free vibrations of your body; they are forced vibrations. Their frequency is not the natural frequency of vibration of your body, but the forcing frequency of the bus.

In the same way, you can force a metre ruler to oscillate by waving it up and down; however, its natural frequency of vibration will be much greater than this, as you will discover if you hold one end down on the bench and twang the other end (Figure 19.2).



Figure 19.2 A ruler vibrating freely at its natural frequency.

QUESTION

1 State which of the following are free oscillations, and which are forced:
   - **a** the wing beat of a mosquito
   - **b** the movement of the pendulum in a grandfather clock
   - **c** the vibrations of a cymbal after it has been struck
   - **d** the shaking of a building during an earthquake.

Observing oscillations

Many oscillations are too rapid or too small for us to observe. Our eyes cannot respond rapidly enough if the frequency of oscillation is more than about 5 Hz (five oscillations per second); anything faster than this appears as a blur. In order to see the general characteristics of oscillating systems, we need to find suitable systems that oscillate slowly. Box 19.1 describes three suitable situations to investigate.

BOX 19.1: Observing slow oscillations

A mass–spring system

A trolley, loaded with extra masses, is tethered by identical springs in between two clamps (Figure 19.3). Displace the trolley to one side and it will oscillate back and forth along the bench. Listen to the sound of the trolley moving. Where is it moving fastest? What happens to its speed as it reaches the ends of its oscillation? What is happening to the springs as the trolley oscillates?

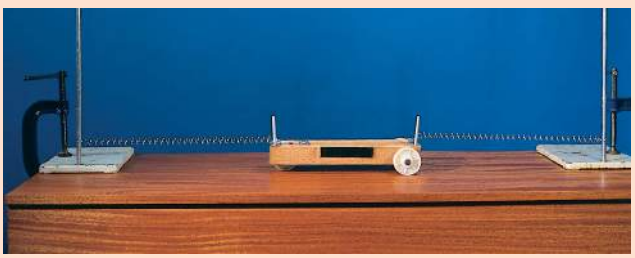


Figure 19.3 A trolley tethered between springs will oscillate freely from side to side.

A long pendulum

A string, at least 2 m long, hangs from the ceiling with a large mass fixed at the end (Figure 19.4). Pull the mass some distance to one side, and let go. The pendulum will swing back and forth at its natural frequency of oscillation. Try to note the characteristics of its motion. In what ways is it similar to the motion of the oscillating trolley? In what ways is it different?

A loudspeaker cone

A signal generator, set to a low frequency (say, 1 Hz), drives a loudspeaker so that it vibrates (Figure 19.5). You need to be able to see the cone of the loudspeaker.

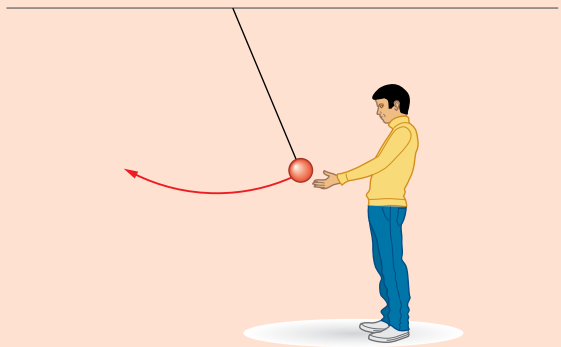


Figure 19.4 A long pendulum oscillates back and forth.

How does this motion compare with that of the pendulum and the mass–spring system? Try using a higher frequency (say, 100 Hz). Use an electronic stroboscope flashing at a similar frequency to show up the movement of the cone. (It may help to paint a white spot on the centre of the cone.) Do you observe the same pattern of movement?

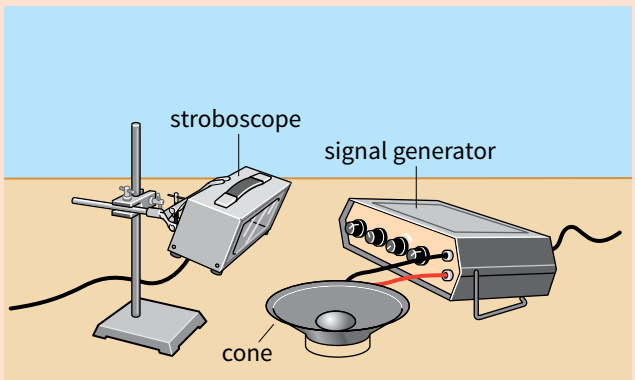


Figure 19.5 A loudspeaker cone forced to vibrate up and down.

QUESTION

2 If you could draw a velocity–time graph for any of the oscillators described in Box 19.1, what would it look like? Would it be a curve like the one shown in Figure 19.6a, or triangular (saw-toothed) like the one shown in Figure 19.6b?

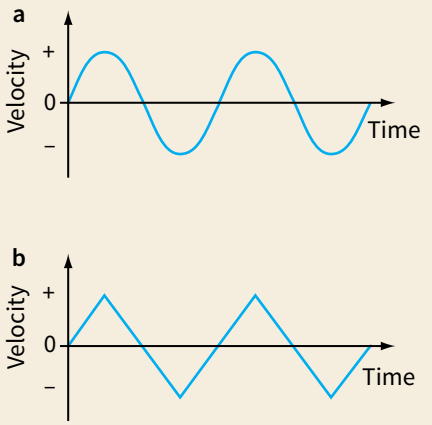


Figure 19.6 Two possible velocity–time graphs for vibrating objects.

Describing oscillations

All of these examples show the same pattern of movement. The trolley accelerates as it moves towards the centre of the oscillation. It is moving fastest at the centre. It decelerates as it moves towards the end of the oscillation. At the extreme position, it stops momentarily, reverses its direction and accelerates back towards the centre again.

Amplitude, period and frequency

Many oscillating systems can be represented by a displacement–time graph like that shown in Figure 19.7. The displacement x varies in a smooth way on either side of the midpoint. The shape of this graph is a sine curve, and the motion is described as **sinusoidal**.

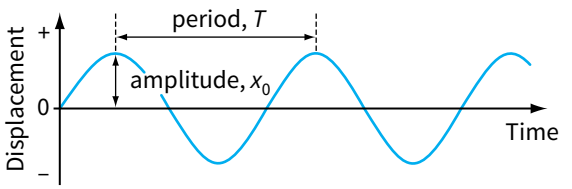


Figure 19.7 A displacement–time graph to show the meaning of amplitude and period.

Notice that the displacement changes between positive and negative values, as the object moves through the equilibrium position. The maximum displacement from the equilibrium position is called the **amplitude** x_0 of the oscillation.

The displacement–time graph can also be used to show the period and frequency of the oscillation. The period T is the time for one complete oscillation. Note that the oscillating object must go from one side to the other and back again (or the equivalent). The frequency f is the number of oscillations per unit time, and so f is the reciprocal of T:

$$\text{frequency} = \frac{1}{\text{period}} \quad \text{or} \quad f = \frac{1}{T}$$

The equation above can also be written as:

$$\text{period} = \frac{1}{\text{frequency}} \quad \text{or} \quad T = \frac{1}{f}$$

QUESTION

3 From the displacement–time graph shown in Figure 19.8, determine the amplitude, period and frequency of the oscillations represented.

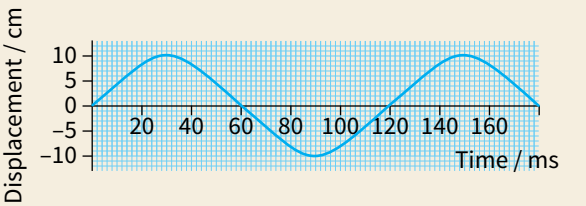


Figure 19.8 A displacement–time graph for an oscillator.

Phase

The term **phase** describes the point that an oscillating mass has reached within the complete cycle of an oscillation. It is often important to describe the **phase difference** between two oscillations. The graph of Figure 19.9a shows two oscillations which are identical except for their phase difference. They are out of step with one another. In this example, they have a phase difference of one-quarter of an oscillation. Phase difference can be measured as a fraction of an oscillation, in degrees or in radians (see Worked example 1).

QUESTION

4 **a** Figure 19.9b shows two oscillations which are out of phase. By what fraction of an oscillation are they out of phase?

b Why would it not make sense to ask the same question about Figure 19.9c?

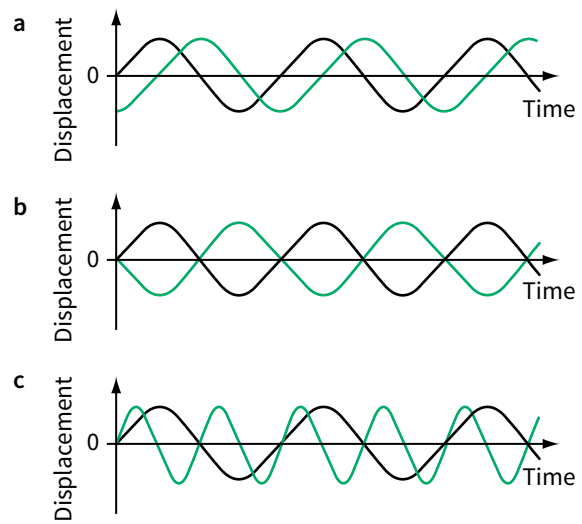


Figure 19.9 Illustrating the idea of phase difference.

Simple harmonic motion

There are many situations where we can observe the special kind of oscillations called **simple harmonic motion** (s.h.m.). Some are more obvious than others. For example, the vibrating strings of a musical instrument show s.h.m. When plucked or bowed, the strings move back and forth about the equilibrium position of their oscillation. The motion of the tethered trolley in Figure 19.3 and that of the pendulum in Figure 19.4 are also s.h.m. (Simple harmonic motion is defined in terms of the acceleration and displacement of an oscillator – see pages 294–5.)

Here are some other, less obvious, situations where simple harmonic motion can be found:

- When a pure (single tone) sound wave travels through air, the molecules of the air vibrate with s.h.m.
- When an alternating current flows in a wire, the electrons in the wire vibrate with s.h.m.
- There is a small alternating electric current in a radio or television aerial when it is tuned to a signal, in the form of electrons moving with s.h.m.
- The atoms that make up a molecule vibrate with s.h.m. (see, for example, the hydrogen molecule in Figure 19.11a).

WORKED EXAMPLE

1 Figure 19.10 shows displacement–time graphs for two identical oscillators. Calculate the phase difference between the two oscillations. Give your answer in degrees and in radians.

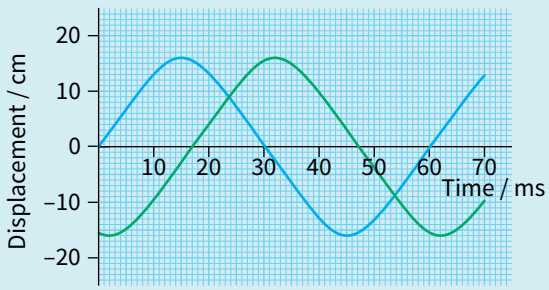


Figure 19.10 The displacement–time graphs of two oscillators with the same period.

Step 1 Measure the time interval t between two corresponding points on the graphs.

$t = 17$ ms

Step 2 Determine the period T for one complete oscillation.

$T = 60$ ms

Hint: *Remember that a complete oscillation is when the object goes from one side to the other and back again.*

Step 3 Now you can calculate the phase difference as a fraction of an oscillation.

phase difference = fraction of an oscillation

Therefore:

$$\text{phase difference} = \frac{t}{T} = \frac{17}{60} = 0.283 \text{ oscillation}$$

Step 4 Convert to degrees and radians. There are 360° and 2π rad in one oscillation.

phase difference = $0.283 \times 360° = 102° \approx 100°$

phase difference = $0.283 \times 2\pi$ rad

$= 1.78 \text{ rad} \approx 1.8 \text{ rad}$

Oscillations can be very complex, with many different frequencies of oscillation occurring at the same time. Examples include the vibrations of machinery, the motion of waves on the sea and the vibration of a solid crystal formed when atoms, ions or molecules bond together (Figure 19.11b). It is possible to break down a complex oscillation into a sum of simple oscillations, and so we will focus our attention in this chapter on s.h.m. with only one frequency. We will also concentrate on large-scale mechanical oscillations, but you should bear in mind that this analysis can be extended to the situations mentioned above, and many more besides.

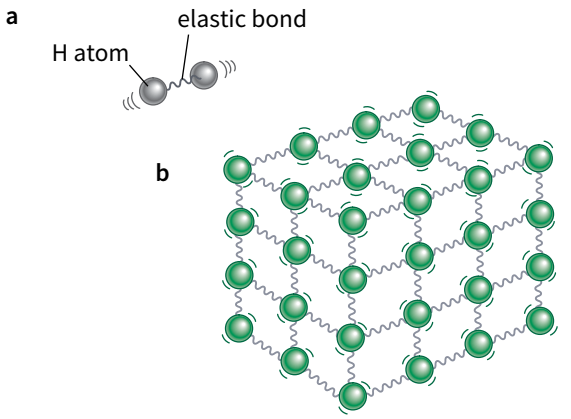


Figure 19.11 We can think of the bonds between atoms as being springy; this leads to vibrations, **a** in a molecule of hydrogen and **b** in a solid crystal.

The requirements for s.h.m.

If a simple pendulum is undisturbed, it is in equilibrium. The string and the mass will hang vertically. To start it swinging (Figure 19.12), it must be pulled to one side of its equilibrium position. The forces on the mass are unbalanced and so it moves back towards its equilibrium position. The mass swings past this point and continues until it comes to rest momentarily at the other side; the process is then repeated in the opposite direction. Note that a complete oscillation in Figure 19.12 is from right to left and back again.

The three requirements for s.h.m. of a mechanical system are:

- a mass that oscillates
- a position where the mass is in equilibrium (conventionally, displacement x to the right of this position is taken as positive; to the left it is negative)
- a restoring force that acts to return the mass to its equilibrium position; the restoring force F is directly proportional to the displacement x of the mass from its equilibrium position and is directed towards that point.

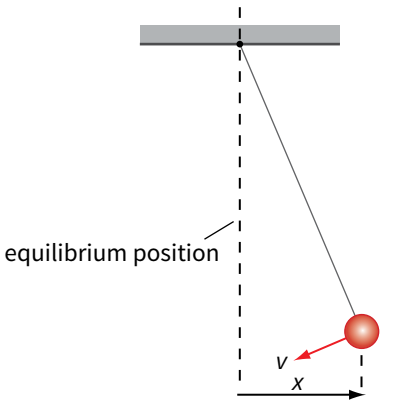


Figure 19.12 This swinging pendulum has positive displacement x and negative velocity v.

The changes of velocity in s.h.m.

As the pendulum swings back and forth, its velocity is constantly changing. As it swings from right to left (as shown in Figure 19.12) its velocity is negative. It accelerates towards the equilibrium position and then decelerates as it approaches the other end of the oscillation. It has positive velocity as it swings back from left to right. Again, it has maximum speed as it travels through the equilibrium position and decelerates as it swings up to its starting position.

This pattern of acceleration – deceleration – changing direction – acceleration again is characteristic of simple harmonic motion. There are no sudden changes of velocity. In the next section we will see how we can observe these changes and how we can represent them graphically.

QUESTIONS

5 Identify the features of the motion of the trolley in Figure 19.3 (on page 287) that satisfy the three requirements for s.h.m.

6 Explain why the motion of someone jumping up and down on a trampoline is not simple harmonic motion. (Their feet lose contact with the trampoline during each bounce.)

Representing s.h.m. graphically

If you set up a trolley tethered between springs (Figure 19.13) you can hear the characteristic rhythm of s.h.m. as the trolley oscillates back and forth. By adjusting the mass carried by the trolley, you can achieve oscillations with a period of about two seconds.

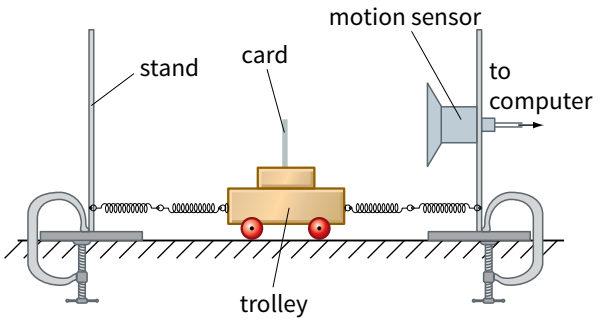


Figure 19.13 A motion sensor can be used to investigate s.h.m. of a spring–trolley system.

The motion sensor allows you to record how the displacement of the trolley varies with time. Ultrasonic pulses from the sensor are reflected by the card on the trolley and the reflected pulses are detected. This 'sonar' technique allows the sensor to determine the displacement of the trolley. A typical screen display is shown in Figure 19.14.

The computer can then determine the velocity of the trolley by calculating the rate of change of displacement. Similarly, it can calculate the rate of change of velocity to determine the acceleration.

Idealised graphs of displacement, velocity and acceleration against time are shown in Figure 19.15. We will examine these graphs in sequence to see what they tell us about s.h.m. and how the three graphs are related to one another.

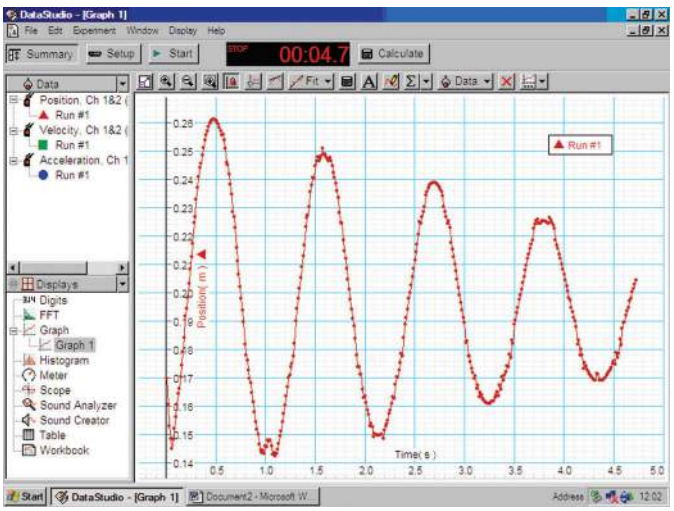


Figure 19.14 A typical displacement–time graph generated by a motion sensor.

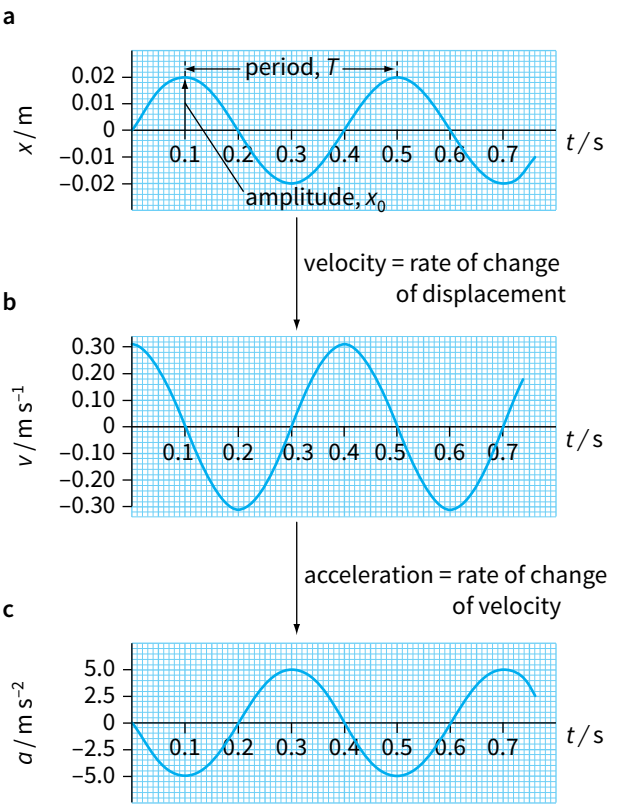


Figure 19.15 Graphs of displacement x, velocity v and acceleration a against time t for s.h.m.

Displacement–time (x–t) graph

The displacement of the oscillating mass varies according to the smooth curve shown in Figure 19.15a. Mathematically, this is a sine curve; its variation is described as **sinusoidal**. Note that this graph allows us to determine the amplitude x_0 and the period T of the oscillations. In this graph, the displacement x of the oscillation is shown as zero at the start, when t is zero. We have chosen to consider the motion to start when the mass is at the midpoint of its oscillation (equilibrium position) and is moving to the right. We could have chosen any other point in the cycle as the starting point, but it is conventional to start as shown here.

Velocity–time (v–t) graph

The velocity v of the oscillator at any time can be determined from the gradient of the displacement–time graph:

$$v = \frac{\Delta x}{\Delta t}$$

Again, we have a smooth curve (Figure 19.15b), which shows how the velocity v depends on time t. The shape of the curve is the same as for the displacement–time graph, but it starts at a different point in the cycle.

When time $t = 0$, the mass is at the equilibrium position and this is where it is moving fastest. Hence the velocity has its maximum value at this point. Its value is positive because at time $t = 0$ it is moving towards the right.

Acceleration–time (a–t) graph

Finally, the acceleration a of the oscillator at any time can be determined from the gradient of the velocity–time graph:

$$a = \frac{\Delta v}{\Delta t}$$

This gives a third curve of the same general form (Figure 19.15c), which shows how the acceleration a depends on time t. At the start of the oscillation, the mass is at its equilibrium position. There is no resultant force acting on it so its acceleration is zero. As it moves to the right, the restoring force acts towards the left, giving it a negative acceleration. The acceleration has its greatest value when the mass is displaced farthest from the equilibrium position. Notice that the acceleration graph is 'upside-down' compared with the displacement graph. This shows that:

$$\text{acceleration} \propto -\text{displacement}$$

or

$$a \propto -x$$

In other words, whenever the mass has a positive displacement (to the right), its acceleration is to the left, and vice versa.

Frequency and angular frequency

The frequency f of s.h.m. is equal to the number of oscillations per unit time. As we saw earlier, f is related to the period T by:

$$f = \frac{1}{T}$$

We can think of a complete oscillation of an oscillator or a cycle of s.h.m. as being represented by 2π radians. (This is similar to a complete cycle of circular motion, where an object moves round through 2π radians.) The phase of the oscillation changes by 2π rad during one oscillation. Hence, if there are f oscillations in unit time, there must be $2\pi f$ radians in unit time. This quantity is the **angular frequency** of the s.h.m. and it is represented by the symbol ω (Greek letter omega).

The angular frequency ω is thus related to frequency f by the following equation:

$$\omega = 2\pi f$$

Since $f = \frac{1}{T}$, the angular frequency ω is related to the period T of the oscillator by the equation:

$$\omega = \frac{2\pi}{T} \quad \text{or} \quad T = \frac{2\pi}{\omega}$$

QUESTIONS

7 Use the graphs shown in Figure 19.15 to determine the values of the following quantities:

- **a** amplitude
- **b** period
- **c** maximum velocity
- **d** maximum acceleration.

8 State at what point in an oscillation the oscillator has zero velocity but positive acceleration.

9 Look at the x–t graph of Figure 19.15a. When $t = 0.1$ s, what is the gradient of the graph? State the velocity at this instant.

10 Figure 19.16 shows the displacement–time (x–t) graph for an oscillating mass. Use the graph to determine the following quantities:

- **a** the velocity in cm s^{-1} when $t = 0$ s
- **b** the maximum velocity in cm s^{-1}
- **c** the acceleration in cm s^{-2} when $t = 1.0$ s.

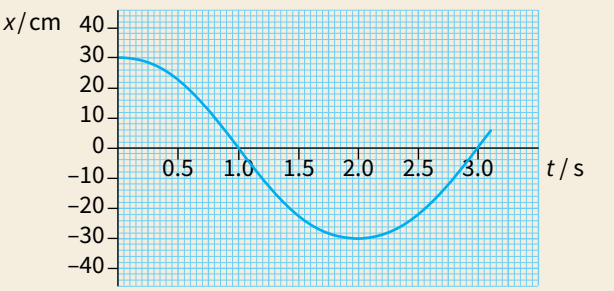


Figure 19.16 A displacement–time graph for an oscillating mass – see Question 10.

In Figure 19.17, a single cycle of s.h.m. is shown, but with the x-axis marked with the phase of the motion in radians.

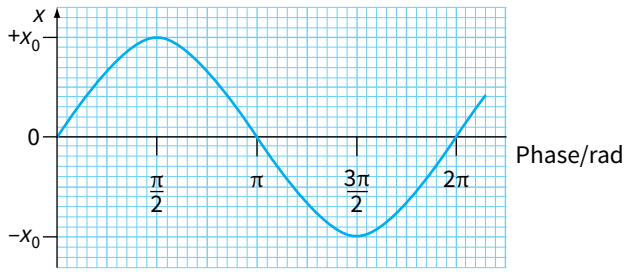


Figure 19.17 The phase of an oscillation varies from 0 to 2π during one cycle.

Equations of s.h.m.

The graph of Figure 19.15a, shown earlier, represents how the displacement of an oscillator varies during s.h.m. We have already mentioned that this is a sine curve. We can present the same information in the form of an equation. The relationship between the displacement x and the time t is as follows:

$$x = x_0 \sin \omega t$$

where x_0 is the amplitude of the motion and ω is its frequency. Sometimes the same motion is represented using a cosine function, rather than a sine function:

$$x = x_0 \cos \omega t$$

The difference between these two equations is illustrated in Figure 19.19. The sine version starts at $x = 0$, i.e. the oscillating mass is at its equilibrium position when $t = 0$.

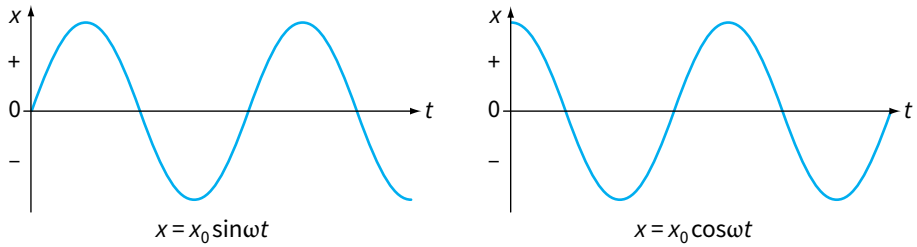


Figure 19.19 These two graphs represent the same simple harmonic motion. The difference in starting positions is related to the sine and cosine forms of the equation for x as a function of t.

QUESTIONS

11 An object moving with s.h.m. goes through two complete cycles in 1.0 s. Calculate:

- **a** the period T
- **b** the frequency f
- **c** the angular frequency ω.

12 Figure 19.18 shows the displacement–time graph for an oscillating mass. Use the graph to determine the following:

- **a** amplitude
- **b** period
- **c** frequency
- **d** angular frequency
- **e** displacement at A
- **f** velocity at B
- **g** velocity at C.

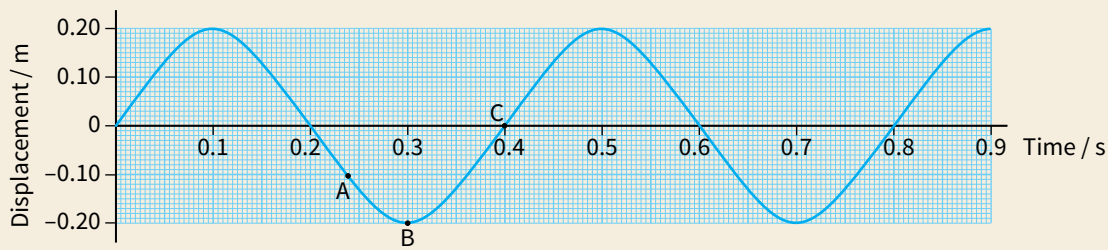


Figure 19.18 A displacement–time graph. For Question 12.

13 An atom in a crystal vibrates with s.h.m. with a frequency of 10^{14} Hz. The amplitude of its motion is 2.0×10^{-12} m.

- **a** Sketch a graph to show how the displacement of the atom varies during one cycle.
- **b** Use your graph to estimate the maximum velocity of the atom.

The cosine version starts at $x = x_0$, so that the mass is at its maximum displacement when $t = 0$.

Note that, in calculations using these equations, the quantity ωt is in radians. Make sure that your calculator is in radian mode for any calculation (see Worked example 2). The presence of the π in the equation should remind you of this.

QUESTIONS

14 The vibration of a component in a machine is represented by the equation:

$x = 3.0 \times 10^{-4} \sin(240\pi t)$

where the displacement x is in metres. Determine the **a** amplitude, **b** frequency and **c** period of the vibration.

15 A trolley is at rest, tethered between two springs. It is pulled 0.15 m to one side and, when time $t = 0$, it is released so that it oscillates back and forth with s.h.m. The period of its motion is 2.0 s.

a Write an equation for its displacement x at any time t (assume that the motion is not damped by frictional forces).

b Sketch a displacement–time graph to show two cycles of the motion, giving values where appropriate.

Acceleration and displacement

In s.h.m., an object's acceleration depends on how far it is displaced from its equilibrium position and on the magnitude of the restoring force. The greater the displacement x, the greater the acceleration a. In fact, a is proportional to x. We can write the following equation to represent this:

$$a = -\omega^2 x$$

This equation shows that a is proportional to x; the constant of proportionality is ω^2. The minus sign shows that, when the object is displaced to the **right**, the direction of its acceleration is to the **left**. The acceleration is always directed towards the equilibrium position, in the opposite direction to the displacement.

It should not be surprising that angular frequency ω appears in this equation. Imagine a mass hanging on a spring, so that it can vibrate up and down. If the spring is stiff, the force on the mass will be greater, it will be accelerated more for a given displacement and its frequency of oscillation will be higher.

The equation

$$a = -\omega^2 x$$

helps us to define simple harmonic motion. The acceleration a is directly proportional to displacement x; and the minus sign shows that it is in the opposite direction.

WORKED EXAMPLE

2 A pendulum oscillates with frequency 1.5 Hz and amplitude 0.10 m. If it is passing through its equilibrium position when $t = 0$, write an equation to represent its displacement x in terms of amplitude x_0, angular frequency ω and time t. Determine its displacement when $t = 0.50$ s.

Step 1 Select the correct equation. In this case, the displacement is zero when $t = 0$, so we use the sine form:

$x = x_0 \sin \omega t$

Step 2 From the frequency f, calculate the angular frequency ω:

$\omega = 2\pi f = 2 \times \pi \times 1.5 = 3.0\pi$

Step 3 Substitute values in the equation: $x_0 = 0.10$ m, so:

$x = 0.10 \sin(3.0\pi t)$

Hint: *Remember to put your calculator into radian mode.*

Step 4 To find x when $t = 0.50$ s, substitute for t and calculate the answer:

$x = 0.10 \sin(2\pi \times 1.5 \times 0.50)$

$= 0.10 \sin(4.713)$

$= -0.10$ m

This means that the pendulum is at the extreme end of its oscillation; the minus sign means that it is at the negative or left-hand end, assuming you have chosen to consider displacements to the right as positive.

(If your calculation went like this:

$x = 0.10 \sin(2\pi \times 1.5 \times 0.50) = 0.10 \sin(4.713)$

$= -8.2 \times 10^{-3}$ m

then your calculator was incorrectly set to work in degrees, not radians.)

Simple harmonic motion is defined as follows:

> A body executes simple harmonic motion if its acceleration is directly proportional to its displacement from its equilibrium position, and in the opposite direction to its displacement.

If a and x were in the same direction (no minus sign), the body's acceleration would increase as it moved away from the fixed point and it would move away faster and faster, never to return.

Figure 19.20 shows the acceleration–displacement (a–x) graph for an oscillator executing s.h.m. Note the following:

- The graph is a straight line through the origin ($a \propto x$).
- It has a negative slope (the minus sign in the equation $a = -\omega^2 x$). This means that the acceleration is always directed towards the equilibrium position.
- The magnitude of the gradient of the graph is ω^2.
- The gradient is independent of the amplitude of the motion. This means that the frequency f or the period T of the oscillator is independent of the amplitude and so a simple harmonic oscillator keeps steady time.

A mathematical note: we say that the equation $a = -\omega^2 x$ defines simple harmonic motion – it tells us what is required if a body is to perform s.h.m. The equation $x = x_0 \sin \omega t$ is then described as a **solution** to the equation, since it tells us how the displacement of the body varies with time. If you have studied calculus you may be able to differentiate the equation for x twice with respect to time to obtain an equation for acceleration and thereby show that the defining equation $a = -\omega^2 x$ is satisfied.

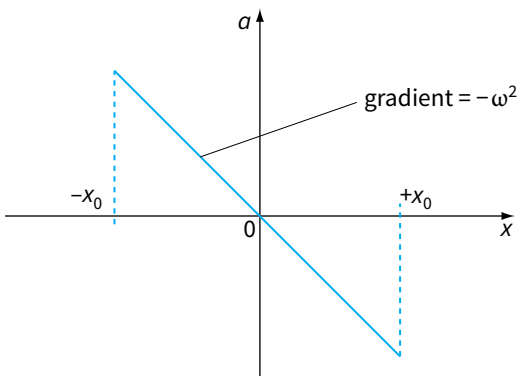


Figure 19.20 Graph of acceleration a against displacement x for an oscillator executing s.h.m.

Equations for velocity

The velocity v of an oscillator varies as it moves back and forth. It has its greatest speed when it passes through the equilibrium position in the middle of the oscillation. If we take time $t = 0$ when the oscillator passes through the middle of the oscillation with its greatest speed v_0, then we can represent the changing velocity as an equation:

$$v = v_0 \cos \omega t$$

We use the cosine function to represent the velocity since it has its maximum value when $t = 0$.

The equation $v = v_0 \cos \omega t$ tells us how v depends on t. We can write another equation to show how the velocity depends on the oscillator's displacement x:

$$v = \pm\, \omega \sqrt{x_0^2 - x^2}$$

This equation can be used to deduce the speed of an oscillator at any point in an oscillation, including its maximum speed.

Maximum speed of an oscillator

If an oscillator is executing simple harmonic motion, it has maximum speed when it passes through its equilibrium position. This is when its displacement x is zero. The maximum speed v_0 of the oscillator depends on the frequency f of the motion and on the amplitude x_0. Substituting $x = 0$ in the equation:

$$v = \pm\, \omega \sqrt{x_0^2 - x^2}$$

gives the maximum speed:

$$v_0 = \omega\, x_0$$

According to this equation, for a given oscillation:

$$v_0 \propto x_0$$

A simple harmonic oscillator has a period that is independent of the amplitude. A greater amplitude means that the oscillator has to travel a greater distance in the same time – hence it has a greater speed.

The equation also shows that:

$$v_0 \propto \omega$$

so that the maximum speed is proportional to the frequency. Increasing the frequency means a shorter period. A given distance is covered in a shorter time – hence it has a greater speed.

Have another look at Figure 19.15. The period of the motion is 0.40 s and the amplitude of the motion is 0.02 m. The frequency f can be calculated as follows:

$$f = \frac{1}{t} = \frac{1}{0.40} = 2.5\,\text{Hz}$$

We can now use the equation $v_0 = (2\pi f)x_0$ to determine the maximum speed v_0:

$$v_0 = (2\pi f)x_0 = (2\pi \times 2.5) \times 2.0 \times 10^{-2}$$

$$v_0 \approx 0.31\ \text{m s}^{-1}$$

This is how the values on Figure 19.15b were calculated.

QUESTIONS

16 A mass secured at the end of a spring moves with s.h.m. The frequency of its motion is 1.4 Hz.

a Write an equation of the form $a = -\omega^2 x$ to show how the acceleration of the mass depends on its displacement.

b Calculate the acceleration of the mass when it is displaced 0.050 m from its equilibrium position.

17 A short pendulum oscillates with s.h.m. such that its acceleration a (in m s^{-2}) is related to its displacement x (in m) by the equation $a = -300x$. Determine the frequency of the oscillations.

18 The pendulum of a grandfather clock swings from one side to the other in 1.00 s. The amplitude of the oscillation is 12 cm.

a Calculate:

i the period of its motion

ii the frequency

iii the angular frequency.

b Write an equation of the form $a = -\omega^2 x$ to show how the acceleration of the pendulum weight depends on its displacement.

c Calculate the maximum speed of the pendulum bob.

d Calculate the speed of the bob when its displacement is 6 cm.

19 A trolley of mass m is fixed to the end of a spring. The spring can be compressed and extended. The spring has a force constant k. The other end of the spring is attached to a vertical wall. The trolley lies on a smooth horizontal table. The trolley oscillates when it is displaced from its equilibrium position.

a Show that the motion of the oscillating trolley is s.h.m.

b Show that the period T of the trolley is given by the equation:

$$T = 2\pi\sqrt{\frac{m}{k}}$$

Energy changes in s.h.m.

During simple harmonic motion, there is a constant interchange of energy between two forms: potential and kinetic. We can see this by considering the mass–spring system shown in Figure 19.21.

When the mass is pulled to one side (to start the oscillations), one spring is compressed and the other is stretched. The springs store elastic potential energy. When the mass is released, it moves back towards the equilibrium position, accelerating as it goes. It has increasing kinetic energy. The potential energy stored in the springs decreases while the kinetic energy of the mass increases by the same amount (as long as there are no heat losses due to frictional forces). Once the mass has passed the equilibrium position, its kinetic energy decreases and the energy is transferred back to the springs. Provided the oscillations are undamped, the total energy in the system remains constant.

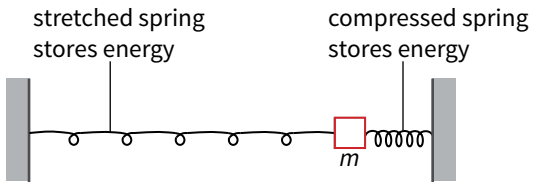


Figure 19.21 The elastic potential energy stored in the springs is converted to kinetic energy when the mass is released.

Energy graphs

We can represent these energy changes in two ways. Figure 19.22 shows how the kinetic energy and elastic potential energy change with time. Potential energy is maximum when displacement is maximum (positive or negative). Kinetic energy is maximum when displacement is zero. The total energy remains constant throughout. Note that both kinetic energy and potential energy go through two complete cycles during one period of the oscillation. This is because kinetic energy is maximum when the mass is

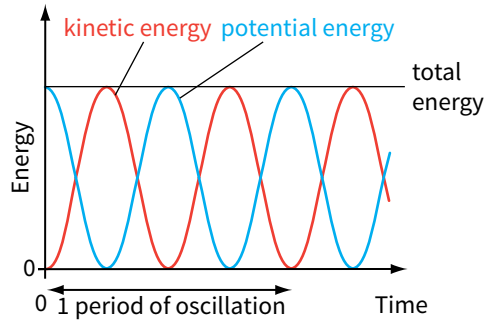


Figure 19.22 The kinetic energy and potential energy of an oscillator vary periodically, but the total energy remains constant if the system is undamped.

passing through the equilibrium position moving to the left and again moving to the right. The potential energy is maximum at both ends of the oscillation.

A second way to show this is to draw a graph of how potential energy and kinetic energy vary with displacement (Figure 19.23).
The graph shows that:

- kinetic energy is maximum when displacement $x = 0$
- potential energy is maximum when $x = \pm x_0$
- at any point on this graph, the total energy (k.e. + p.e.) has the same value.

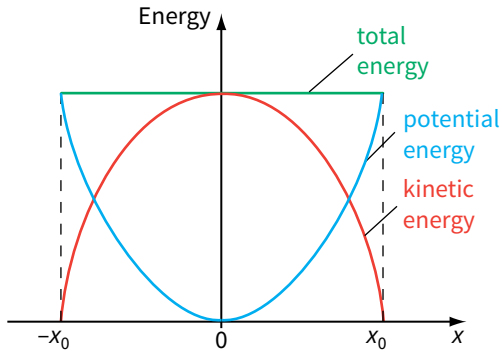


Figure 19.23 The kinetic energy is maximum at zero displacement; the potential energy is maximum at maximum displacement (x_0 and $-x_0$).

QUESTIONS

20 To start a pendulum swinging, you pull it slightly to one side.

a What kind of energy does this transfer to the mass?

b Describe the energy changes that occur when the mass is released.

21 Figure 19.23 shows how the different forms of energy change with displacement during s.h.m. Copy the graph, and show how the graph would differ if the oscillating mass were given only half the initial input of energy.

22 Figure 19.24 shows how the velocity v of a 2.0 kg mass was found to vary with time t during an investigation of the s.h.m. of a pendulum. Use the graph to estimate the following for the mass:

a its maximum velocity

b its maximum kinetic energy

c its maximum potential energy

d its maximum acceleration

e the maximum restoring force that acted on it.

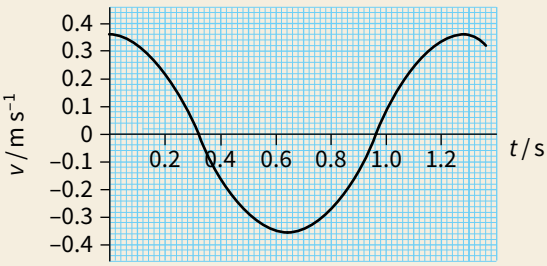


Figure 19.24 A velocity–time graph for a pendulum – see Question 22.

Damped oscillations

In principle, oscillations can go on for ever. In practice, however, the oscillations we observe around us do not. They die out, either rapidly or gradually. A child on a swing knows that the amplitude of her swinging will decline until eventually she will come to rest, unless she can put some more energy into the swinging to keep it going.

This happens because of friction. On a swing, there is friction where the swing is attached to the frame and there is friction with the air. The amplitude of the child's oscillations decreases as the friction transfers energy away from her to the surroundings.

We describe these oscillations as **damped**. Their amplitude decreases according to a particular pattern. This is shown in Figure 19.25.

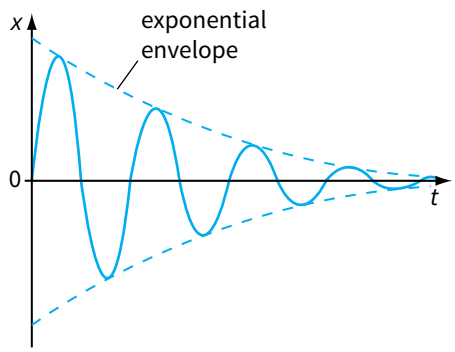


Figure 19.25 Damped oscillations.

The amplitude of damped oscillations does not decrease linearly. It decays exponentially with time. An exponential decay is a particular mathematical pattern that arises as follows. At first, the swing moves rapidly. There is a lot of air resistance to overcome, so the swing loses energy quickly and its amplitude decreases at a high rate. Later, it is moving more slowly. There is less air resistance and so energy is lost more slowly – the amplitude decreases at a lower rate. Hence we get the characteristic curved shape, which is the 'envelope' of the graph in Figure 19.25.

Notice that the frequency of the oscillations does not change as the amplitude decreases. This is a characteristic of simple harmonic motion. The child may, for example, swing back and forth once every two seconds, and this stays the same whether the amplitude is large or small.

BOX 19.2: Investigating damping

You can investigate the exponential decrease in the amplitude of oscillations using a simple laboratory arrangement (Figure 19.26). A hacksaw blade or other springy metal strip is clamped (vertically or horizontally) to the bench. A mass is attached to the free end. This will oscillate freely if you displace it to one side.

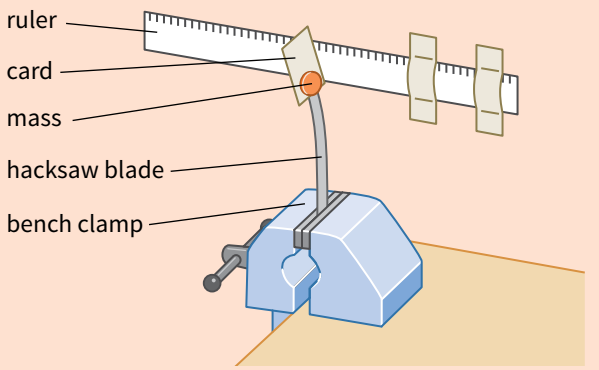


Figure 19.26 Damped oscillations with a hacksaw blade.

A card is attached to the mass so that there is significant air resistance as the mass oscillates. The amplitude of the oscillations decreases and can be measured every five oscillations by judging the position of the blade against a ruler fixed alongside.

A graph of amplitude against time will show the characteristic exponential decrease. You can find the 'half-life' of this exponential decay graph by determining the time it takes to decrease to half its initial amplitude (Figure 19.27).

By changing the size of the card, it is possible to change the degree of damping, and hence alter the half-life of the motion.

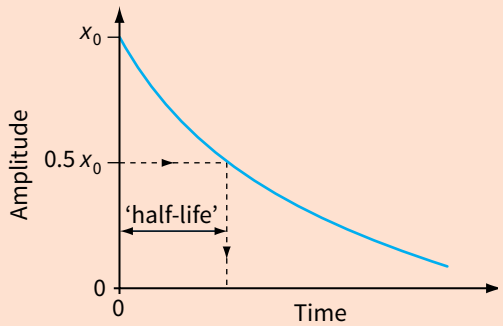


Figure 19.27 A typical graph of amplitude against time for damped oscillations.

Energy and damping

Damping can be very useful if we want to get rid of vibrations. For example, a car has springs (Figure 19.28) which make the ride much more comfortable for us when the car goes over a bump. However, we wouldn't want to spend every car journey vibrating up and down as a reminder of the last bump we went over. So the springs are damped by the shock absorbers, and we return rapidly to a smooth ride after every bump.



Figure 19.28 The springs and shock absorbers in a car suspension system form a damped system.

Damping is achieved by introducing the force of friction into a mechanical system. In an undamped oscillation, the total energy of the oscillation remains constant. There is a regular interchange between potential and kinetic energy. By introducing friction, damping has the effect of removing energy from the oscillating system, and the amplitude and maximum speed of the oscillation decrease.

QUESTION

23 **a** Sketch graphs to show how each of the following quantities changes during the course of a single complete oscillation of an undamped pendulum: kinetic energy, potential energy, total energy.

b State how your graphs would be different for a lightly damped pendulum.

Resonance

Resonance is an important physical phenomenon that can appear in a great many different situations. A dramatic example is the Millennium Footbridge in London, opened in June 2000 (Figure 19.29). With up to 2000 pedestrians walking on the bridge, it started to sway dangerously. The people also swayed in time with the bridge, and this caused the amplitude of the bridge's oscillations to increase – this is resonance. After three days the bridge was closed. It took engineers two years to analyse the problem and then add 'dampers' to the bridge to absorb the energy of its oscillations. The bridge was then reopened and there have been no problems since.

You will have observed a much more familiar example of resonance when pushing a small child on a swing. The swing plus child has a natural frequency of oscillation. A small push in each cycle results in the amplitude increasing until the child is swinging high in the air.



Figure 19.29 The 'wobbly' Millennium Footbridge in London was closed for nearly two years to correct problems caused by resonance.

BOX 19.3: Observing resonance

Resonance can be observed with almost any oscillating system. The system is forced to oscillate at a particular frequency. If the forcing frequency happens to match the natural frequency of oscillation of the system, the amplitude of the resulting oscillations can build up to become very large.

Barton's pendulums

Barton's pendulums is a demonstration of this (Figure 19.30). Several pendulums of different lengths hang from a horizontal string. Each has its own natural frequency of oscillation. The 'driver' pendulum at the end is different; it has a large mass at the end, and its length is equal to that of one of the others. When the driver is set swinging, the others gradually start to move. However, only the pendulum whose length matches that of the driver pendulum builds up a large amplitude so that it is resonating.



Figure 19.30 Barton's pendulums.

BOX 19.3: Observing resonance (continued)

What is going on here? All the pendulums are coupled together by the suspension. As the driver swings, it moves the suspension, which in turn moves the other pendulums. The frequency of the matching pendulum is the same as that of the driver, and so it gains energy and its amplitude gradually builds up. The other pendulums have different natural frequencies, so the driver has little effect.

In a similar way, if you were to push the child on the swing once every three-quarters of an oscillation, you would soon find that the swing was moving backwards as you tried to push it forwards, so that your push would slow it down.

A mass–spring system

You can observe resonance for yourself with a simple mass–spring system. You need a mass on the end of a spring (Figure 19.31), chosen so that the mass oscillates up and down with a natural frequency of about 1 Hz. Now hold the top end of the spring and move your hand up and down rapidly, with an amplitude of a centimetre or two. Very little happens. Now move your hand up and down more slowly, close to 1 Hz.

You should see the mass oscillating with gradually increasing amplitude. Adjust your movements to the exact frequency of the natural vibrations of the mass and you will see the greatest effect.



Figure 19.31 Resonance with a mass on a spring.

Defining resonance

For resonance to occur, we must have a system that is capable of oscillating freely. We must also have some way in which the system is forced to oscillate. When the forcing frequency matches the natural frequency of the system, the amplitude of the oscillations grows dramatically.

If the driving frequency does not quite match the natural frequency, the amplitude of the oscillations will increase, but not to the same extent as when resonance is achieved. Figure 19.32 shows how the amplitude of oscillations depends on the driving frequency in the region close to resonance.

In resonance, energy is transferred from the driver to the resonating system more efficiently than when resonance does not occur. For example, in the case of the Millennium Footbridge, energy was transferred from the pedestrians to the bridge, causing large-amplitude oscillations.

The following statements apply to any system in resonance:

- Its natural frequency is equal to the frequency of the driver.
- Its amplitude is maximum.
- It absorbs the greatest possible energy from the driver.

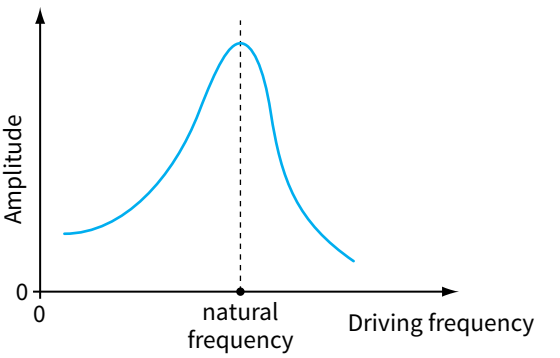


Figure 19.32 Maximum amplitude is achieved when the driving frequency matches the natural frequency of oscillation.

Resonance and damping

During earthquakes, buildings are forced to oscillate by the vibrations of the Earth. Resonance can occur, resulting in serious damage (Figure 19.33). In regions of the world where earthquakes happen regularly, buildings may be built on foundations that absorb the energy of the shock waves. In this way, the vibrations are 'damped' so that the amplitude of the oscillations cannot reach dangerous levels. This is an expensive business, and so far is restricted to the wealthier parts of the world.



Figure 19.33 Resonance during the Mexico City earthquake of 19 September 1985 caused the collapse of many buildings. The earthquake, whose epicentre was in the Pacific Ocean, measured 8.1. Many lives were lost.

Damping is thus useful if we want to reduce the damaging effects of resonance. Figure 19.34 shows how damping alters the resonance response curve of Figure 19.32. Notice that, as the degree of damping is increased, the amplitude of the resonant vibrations decreases. The resonance peak becomes broader. There is also an effect on the frequency at which resonance occurs, which becomes lower as the damping increases.

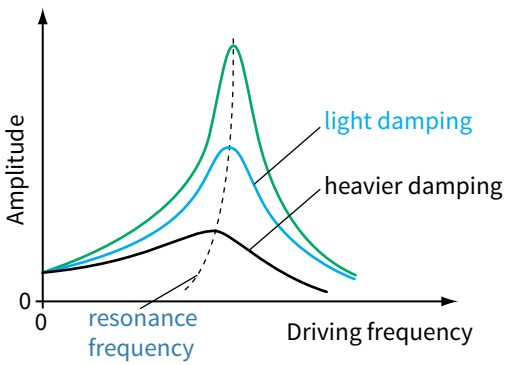


Figure 19.34 Damping reduces the amplitude of resonant vibrations.

An everyday example of damping can be seen on some doors. For example, a restaurant may have a door leading to the kitchen; this door can swing open in either direction. Such a door is designed to close by itself after someone has passed through it. Ideally, the door should swing back quickly without overshooting its closed position. To achieve this, the door hinges (or the closing mechanism) must be correctly damped. If the hinges are damped too lightly, the door will swing back and forth several times as it closes. If the damping is too heavy, it will take too long to close. With **critical damping**, the door will swing closed quickly without oscillating.

Critical damping is thus the minimum amount of damping required to return an oscillator to its equilibrium position without oscillating. Under-damping results in unwanted oscillations; over-damping results in a slower return to equilibrium (see Figure 19.35). A car's suspension system uses springs to smooth out bumps in the road. It is usually critically damped so that passengers do not experience nasty vibrations every time the car goes over a bump.

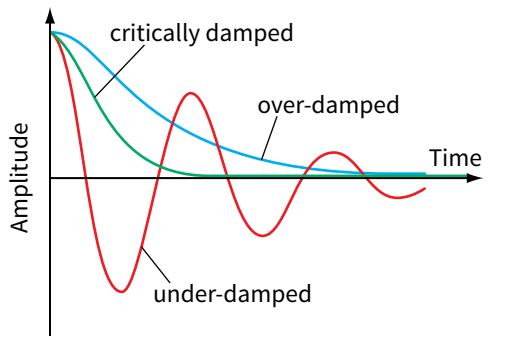


Figure 19.35 Critical damping is just enough to ensure that a damped system returns to equilibrium without oscillating.

Using resonance

As we have seen, resonance can be a problem in mechanical systems. However, it can also be useful. For example, many musical instruments rely on resonance.

Resonance is not confined to mechanical systems. It is made use of in, for example, microwave cooking. The microwaves used have a frequency that matches a natural frequency of vibration of water molecules (the microwave is the 'driver' and the molecule is the 'resonating system'). The water molecules in the food are forced to vibrate and they absorb the energy of the microwave radiation. The water gets hotter and the absorbed energy spreads through the food and cooks or heats it.

Magnetic resonance imaging (MRI) is increasingly used in medicine to produce images such as Figure 19.36, showing aspects of a patient's internal organs. Radio waves having a range of frequencies are used, and particular frequencies are absorbed by particular atomic nuclei. The frequency absorbed depends on the type of nucleus and on its surroundings. By analysing the absorption of the radio waves, a computer-generated image can be produced. (There is much more about how MRI works in Chapter 32.)

A radio or television also depends on resonance for its tuning circuitry. The aerial picks up signals of many different frequencies from many transmitters. The tuner can be adjusted to resonate at the frequency of the transmitting station you are interested in, and the circuit produces a large-amplitude signal for this frequency only.

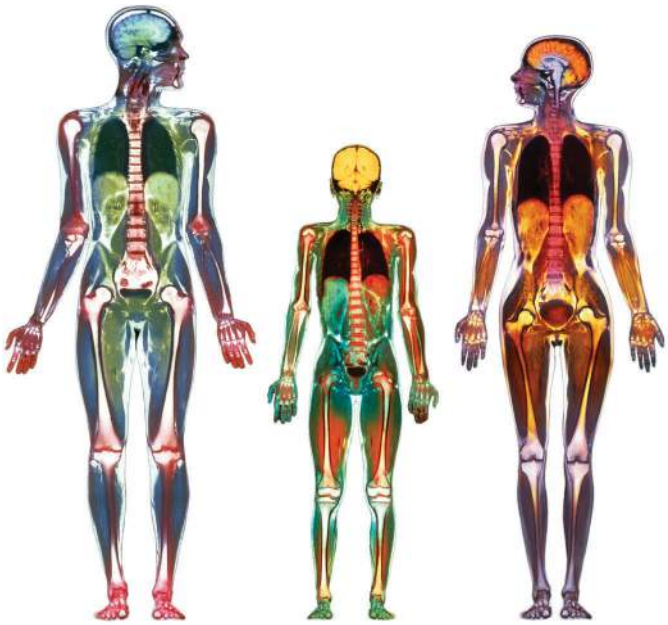


Figure 19.36 This magnetic resonance imaging (MRI) picture shows a man, a woman and a nine-year-old child. The image has been coloured to show up the bones (white), lungs (dark) and other organs.

Big ideas in physics

This study of simple harmonic motion illustrates some important aspects of physics:

- Physicists often take a complex problem (such as how the atoms in a solid vibrate) and reduce it to a simpler, more manageable problem (such as how a mass–spring system vibrates). This is simpler because we know that the spring obeys Hooke's law, so that force is proportional to displacement.
- Physicists generally feel happier if they can write mathematical equations which will give numerical answers to problems. The equation $a = -\omega^2 x$ which describes s.h.m. can be solved to give the sine and cosine equations we have considered above.
- Once physicists have solved one problem like this, they look around for other situations where they can use the same ideas all over again. So the mass–spring theory also works well for vibrating atoms and molecules, for objects bobbing up and down in water, and in many other situations.
- Physicists also seek to modify the theory to fit a greater range of situations. For example, what happens if the vibrating mass experiences a frictional force as it oscillates? (This is damping, as discussed above.) What happens if the spring doesn't obey Hooke's law? (This is a harder question to answer.)

Your A level physics course will help you to build up your appreciation of some of these big ideas – fields (magnetic, electric, gravitational), energy and so on.

QUESTION

24 List **three** examples of situations where resonance is a problem, and three others where resonance is useful. In each case, state what the oscillating system is and what forces it to resonate.

Summary

- Many systems, mechanical and otherwise, will oscillate freely when disturbed from their equilibrium position.
- Some oscillators have motion described as **simple harmonic motion** (s.h.m.). For these systems, graphs of displacement, velocity and acceleration against time are sinusoidal curves – see Figure 19.37.

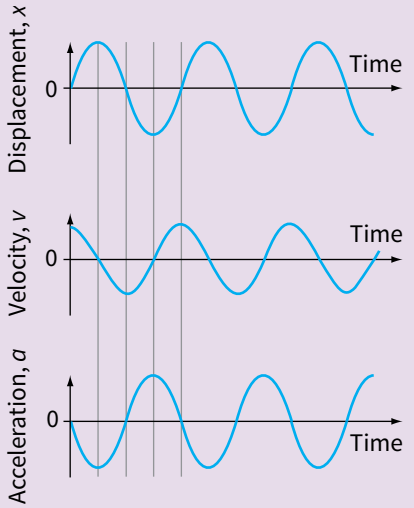


Figure 19.37 Graphs for s.h.m.

- During a single cycle of s.h.m., the phase changes by 2π radians. The angular frequency ω of the motion is related to its period T and frequency f:

  $\omega = \frac{2\pi}{T}$ and $\omega = 2\pi f$

- In s.h.m., displacement x and velocity v can be represented as functions of time t by equations of the form:

  $x = x_0 \sin \omega t$ and $v = v_0 \cos \omega t$

- A body executes simple harmonic motion if its acceleration is directly proportional to its displacement from its equilibrium position, and is always directed towards the equilibrium position.
- Acceleration a in s.h.m. is related to displacement x by the equation $a = -\omega^2 x$.
- The maximum speed v_0 in s.h.m. is given by the equation:

  $v_0 = \omega x_0$.

- The frequency or period of a simple harmonic oscillator is independent of its amplitude.
- In s.h.m., there is a regular interchange between kinetic energy and potential energy.
- Resistive forces remove energy from an oscillating system. This is known as damping. Damping causes the amplitude to decay with time.
- Critical damping is the minimum amount of damping required to return an oscillator to its equilibrium position without oscillating.
- When an oscillating system is forced to vibrate close to its natural frequency, the amplitude of vibration increases rapidly. The amplitude is maximum when the forcing frequency matches the natural frequency of the system; this is resonance.
- Resonance can be a problem, but it can also be very useful.

End-of-chapter questions

1 State and justify whether the following oscillators show simple harmonic motion:

- **a** a basketball being bounced repeatedly on the ground [2]
- **b** a guitar string vibrating [2]
- **c** a conducting sphere vibrating between two parallel, oppositely charged metal plates [1]
- **d** the pendulum of a grandfather clock. [2]

2 The pendulum of a clock is displaced by a distance of 4.0 cm and it oscillates in s.h.m. with a period of 1.0 s.

- **a** Write down an equation to describe the displacement x of the pendulum bob with time t. [2]
- **b** Calculate:
  - **i** the maximum velocity of the pendulum bob [2]
  - **ii** its velocity when its displacement is 2.0 cm. [1]

3 A 50 g mass is attached to a securely clamped spring. The mass is pulled downwards by 16 mm and released, which causes it to oscillate with s.h.m. of time period of 0.84 s.

- **a** Calculate the frequency of the oscillation. [1]
- **b** Calculate the maximum velocity of the mass. [1]
- **c** Calculate the maximum kinetic energy of the mass and state at which point in the oscillation it will have this velocity. [2]
- **d** Write down the maximum gravitational potential energy of the mass (relative to its equilibrium position). You may assume that the damping is negligible. [1]

4 In each of the examples in Figure 19.38, give the phase difference between the two curves: **i** as a fraction of an oscillation, **ii** in degrees and **iii** in radians. [9]

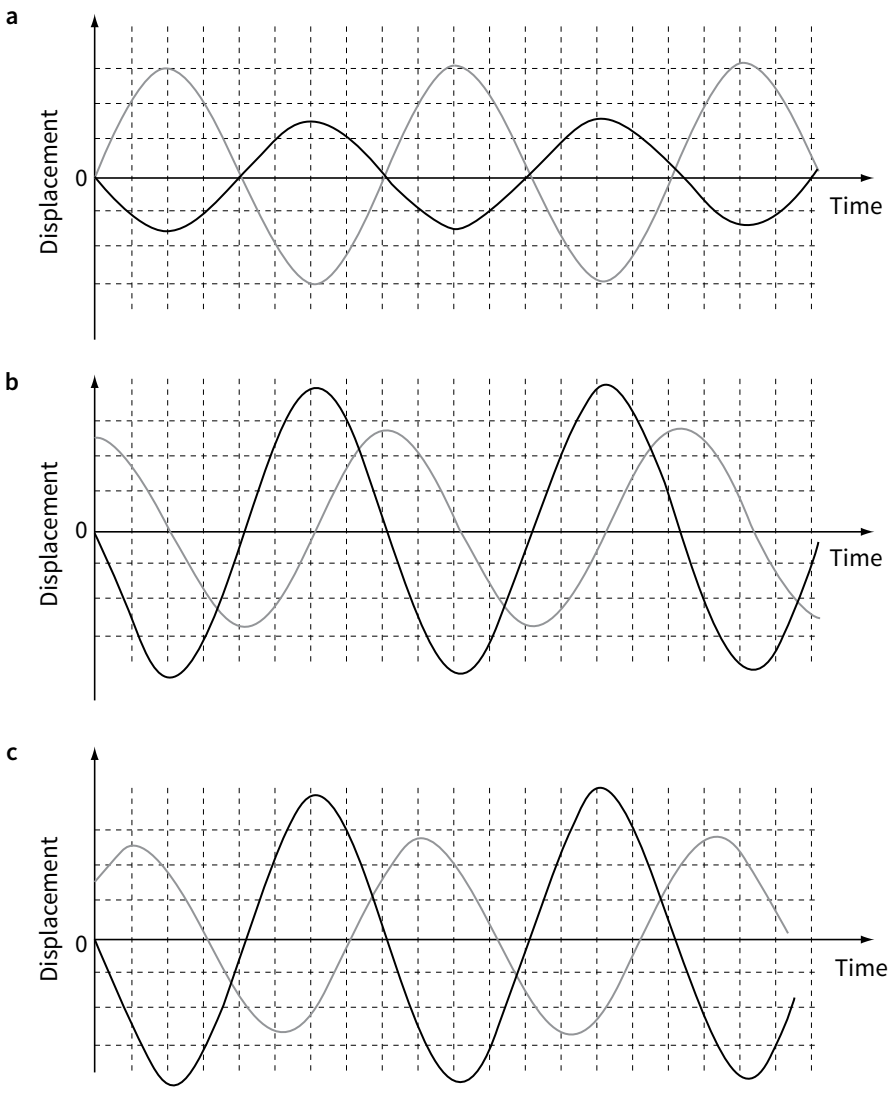


Figure 19.38 For End-of-chapter Question 4.

5 **a** Determine the frequency and the period of the oscillation described by the graph in Figure 19.39. [2]

b Use a copy of the graph and on the same axes sketch:

i the velocity of the particle [1]

ii the acceleration of the particle. [2]

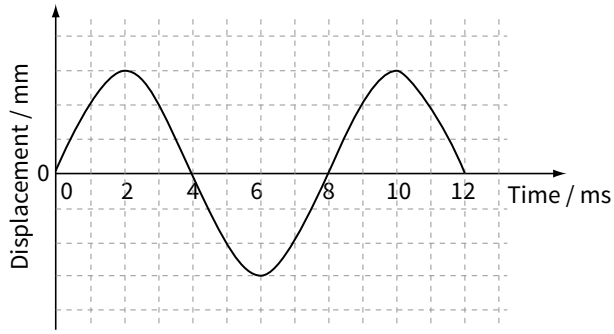


Figure 19.39 For End-of-chapter Question 5.

6 Figure 19.40 shows the displacement of a body as it vibrates between two points.

a State and explain whether the body is moving with simple harmonic motion. [1]

b Make a copy of Figure 19.40.

i On the second set of axes on your copy, show the velocity of the body as it vibrates. [1]

ii On the third set of axes on your copy, show the acceleration of the body. [2]

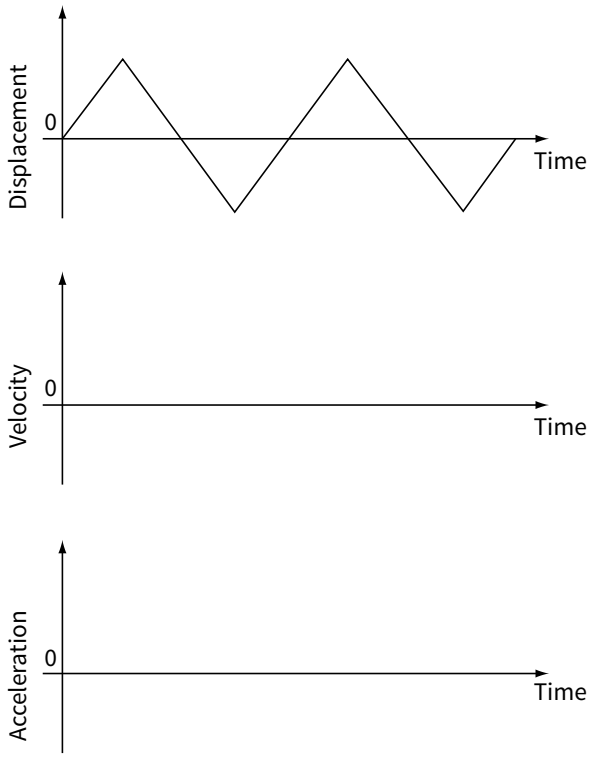


Figure 19.40 For End-of-chapter Question 6.

7 Figure 19.41 shows the piston of a small car engine which oscillates in the cylinder with a motion which approximates simple harmonic motion at 4200 revs per minute (1 rev = 1 cycle). The mass of the piston is 0.24 kg.

a Explain what is meant by **simple harmonic motion.** [2]

b Calculate the frequency of the oscillation. [1]

c The amplitude of the oscillation is 12.5 cm. Calculate:

i the maximum speed at which the piston moves [2]

ii the maximum acceleration of the piston [2]

iii the force required on the piston to produce the maximum acceleration. [1]

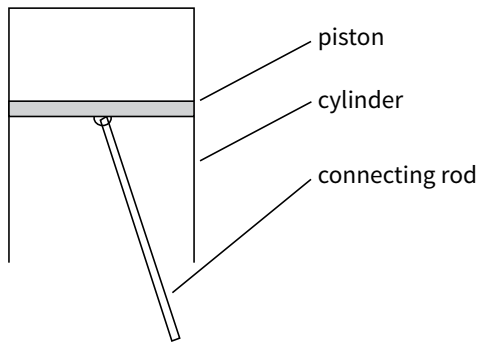


Figure 19.41 For End-of-chapter 7.

8 Figure 19.42 shows a turntable with a rod attached to it a distance 15 cm from the centre. The turntable is illuminated from the side so that a shadow is cast on a screen.

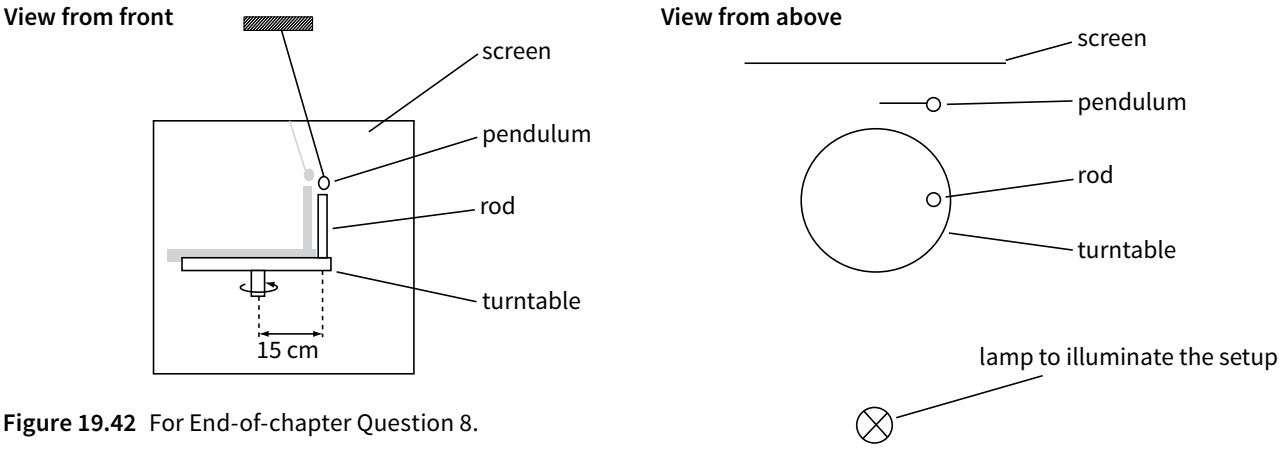


Figure 19.42 For End-of-chapter Question 8.

A simple pendulum is placed behind the turntable and is set oscillating so that it has an amplitude equal to the distance of the rod from the centre of the turntable.

The speed of rotation of the turntable is adjusted. When it is rotating at 1.5 revolutions per second the shadow of the pendulum and the rod are found to move back and forth across the screen exactly in phase.

a Explain what is meant by the term **in phase**. [1]

b Write down an equation to describe the displacement x of the pendulum from its equilibrium position and the angular frequency of the oscillation of the pendulum. [1]

c The turntable rotates through 60° from the position of maximum displacement shown in the diagram.

i Calculate the displacement (from its equilibrium position) of the pendulum at this point. [3]

ii Calculate its speed at this point. [2]

iii Through what further angle must the turntable rotate before it has this speed again? [2]

9 When a cricket ball hits a cricket bat at high speed it can cause a standing wave to form on the bat. In one such example the handle of the bat moved with a frequency of 60 Hz with an amplitude of 2.8 mm.

The vibrational movement of the bat handle can be modelled on simple harmonic motion.

a State the conditions for simple harmonic motion. [2]

b Calculate the maximum acceleration of the bat handle. [2]

c Given that the part of the bat handle held by the cricketer has a mass of 0.48 kg, calculate the maximum force produced on his hands. [1]

d The oscillations are damped and die away after about five complete cycles. Sketch a displacement–time graph to show the oscillations. [2]

10 Seismometers are used to detect and measure the shock waves which travel through the Earth due to earthquakes.

Figure 19.43 shows the structure of a simple seismometer. The shock wave will cause the mass to vibrate, causing a trace to be drawn on the paper scroll.

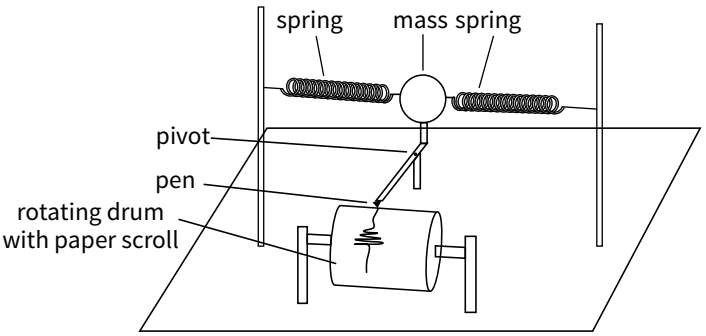


Figure 19.43 A simple seismometer.

a The frequency of a typical shock wave is between 30 and 40 Hz. Explain why the natural frequency of the spring–mass system in the seismometer should be very much less than this range of frequencies. [3]

The graph in Figure 19.44 shows the acceleration of the mass against its displacement when the seismometer is recording an earthquake.

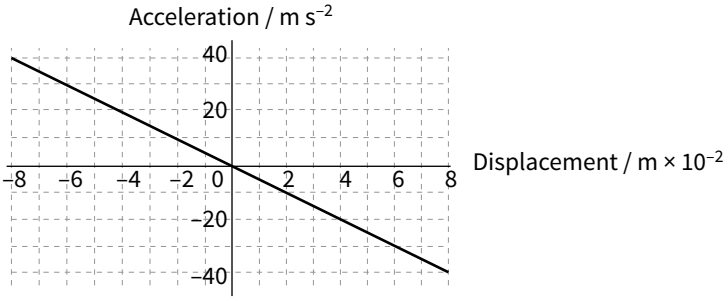
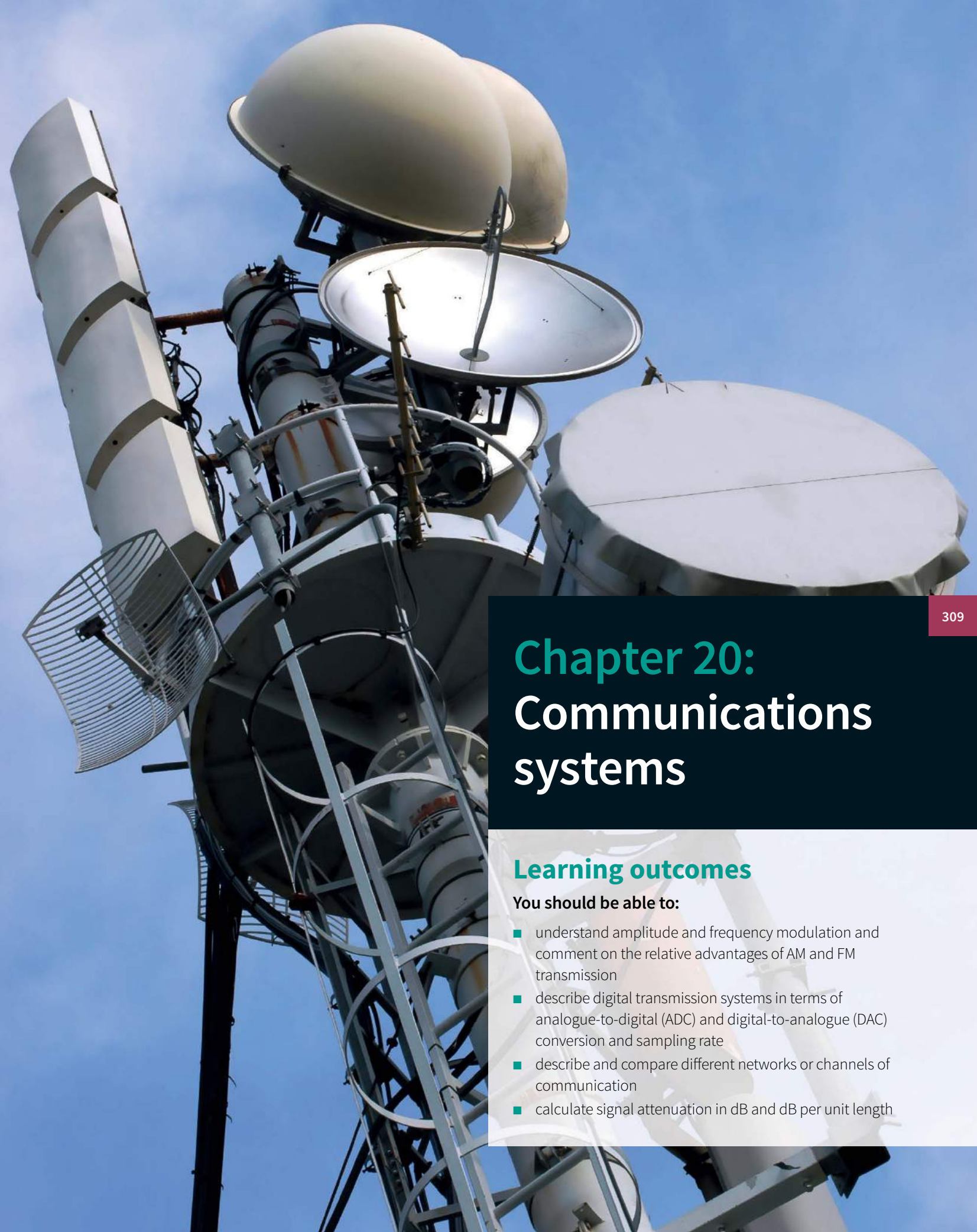


Figure 19.44 Graph of acceleration against displacement recorded during an earthquake.

b What evidence does the graph give that the motion is simple harmonic? [2]

c Use information from the graph to calculate the frequency of the oscillation. [4]



Chapter 20: Communications systems

Learning outcomes

You should be able to:

- understand amplitude and frequency modulation and comment on the relative advantages of AM and FM transmission
- describe digital transmission systems in terms of analogue-to-digital (ADC) and digital-to-analogue (DAC) conversion and sampling rate
- describe and compare different networks or channels of communication
- calculate signal attenuation in dB and dB per unit length

Communications today

The ability to communicate is an important feature of modern life. We can now speak directly to others right around the world and generate vast amounts of information every day. The invention of the mobile phone has increased our ability to talk to one another and to find out what is happening around us. This chapter looks briefly at how some communications systems work. Communications engineers have applied the principles of physics to change the present and shape the future.



Figure 20.1 A communications satellite.

Radio waves

The person listening to the radio in Figure 20.2 is at the end of a communications system. The system starts with sound or music passing into a microphone. The sound signal is converted into a radio signal and at the end of the communications system the radio signal is converted back again into a sound signal.



Figure 20.2 Part of a communications system.

Sound waves and radio waves are completely different types of wave:

- Sound waves travel as mechanical vibrations. Their speed is typically 330 m s^{-1} in air, and audible frequencies lie between 20 Hz and 20 kHz.
- Radio waves are electromagnetic waves which travel at the speed of light (3.0×10^6 m s^{-1} in a vacuum). Their frequencies are much higher than those of sound waves.

(You should recall from Chapter 13 that wave speed v, frequency f and wavelength λ are related by $v = f\lambda$.)

To allow several radio stations to broadcast, each signal has a different **carrier wave** frequency. To hear just one station at a time, you tune the radio receiver so that it separates out radio waves of a single frequency.

The information, for example a sound signal, is carried by modulating or altering the carrier wave. **Modulation** is the variation of either the **amplitude** or the **frequency** of the carrier wave. The modulated wave is the actual wave transmitted. The signal is present in either the changing amplitude or the changing frequency of the modulated wave.

Amplitude modulation

Figure 20.3 shows **amplitude modulation** (AM). The three diagrams show the carrier wave, the signal and the modulated wave. In each case, the horizontal axis represents time, shown on the axis at the bottom. The modulated wave is the carrier wave but its amplitude rises and falls to match the value of the signal at any instant. Look at the amplitude-modulated wave on its own, and notice how the amplitude variation at top and bottom has the same pattern as the signal.

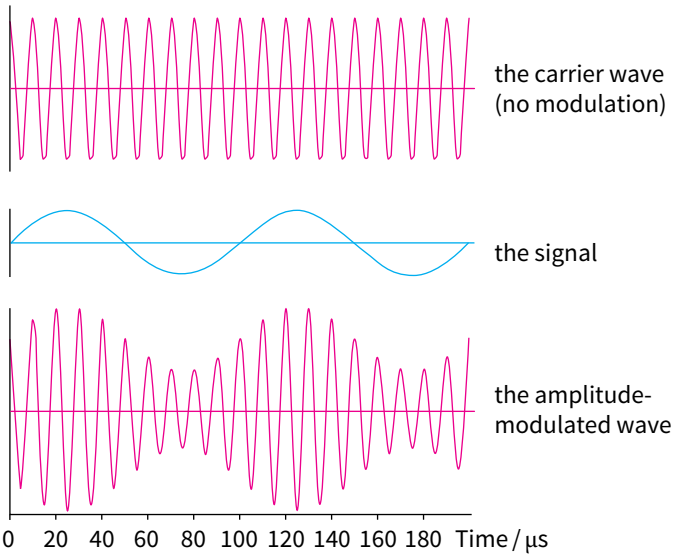


Figure 20.3 Amplitude modulation.

The amplitude of the signal must be less than half of the amplitude of the carrier wave; otherwise the variation in the amplitude at the top will be confused with the variation in amplitude at the bottom of the wave.

If a radio station carries music, the wave transmitted by the radio station will differ from the wave shown in Figure 20.3. There is only one signal frequency present in the signal in Figure 20.3. Music consists of many changing frequencies superposed so that it has a more complex wave pattern. The amplitude of the carrier wave will change as the music pattern changes. The carrier wave frequency does not change but the amplitude of the trace will change with time.

In amplitude modulation (AM), the frequency of the modulated wave is constant. The amplitude of the modulated wave is proportional to, and in phase with, the signal.

WORKED EXAMPLE

1 Calculate the frequencies of the carrier wave and signal shown in Figure 20.3.

Step 1 For the carrier wave, there are 10 complete waves in 100 μs. Hence:

time for one complete carrier wave,

$T = 10\,\mu s = 1 \times 10^{-5}\,s$

carrier wave frequency $f_c = \frac{1}{T} = 100\,000\,Hz$

Step 2 For the signal, there is one complete wave in 100 μs. Hence:

frequency of the signal $f_s = \frac{1}{100 \times 10^{-6}}$

$= 10\,000\,Hz$

A radio wave of 100 000 Hz is in the long-wavelength, low-frequency region of the radio electromagnetic spectrum. A sound frequency of 10 000 Hz is a very high frequency note but is audible.

QUESTIONS

1 Imagine that all the numbers on the time axis in Figure 20.3 are doubled, so that 100 becomes 200 and 200 becomes 400. Calculate the frequency of the carrier wave and the frequency of the signal.

2 Draw an amplitude-modulated wave with a carrier wave of frequency 1.0 MHz and a signal frequency of 100 kHz. The time axis on your graph should be from 0 to 10 μs. On your graph, mark the time for one complete wave of the signal and for one complete wave of the carrier.

3 Explain how an amplitude-modulated wave changes when the input signal:

a increases in loudness

b increases in frequency.

Frequency modulation

In **frequency modulation** (FM) the frequency of the modulated wave varies with time. Without any signal, the frequency of the modulated wave is equal to the frequency of the carrier wave. The size of the input signal at any instant causes the frequency of the modulated wave to change. When the input signal is positive, the frequency of the modulated wave is increased so that it is larger than the frequency of the carrier wave. The larger the signal, the greater is the increase in the frequency. When the signal is negative, the frequency of the modulated wave is less

than the frequency of the carrier wave. Figure 20.4 shows frequency modulation.

In frequency modulation (FM), the modulated wave has a constant amplitude. It is the frequency of the modulated wave that changes as the signal changes.

The frequency deviation of the carrier wave may be given a value, for example, 25 kHz V^{-1}. This means that for every 1.0 V change in the voltage of the signal, the frequency of the carrier wave changes by 25 kHz. In some countries, the maximum allowed change in frequency during FM is set as 75 kHz, in order that the frequency of one station does not overlap the frequency of the next station. If the frequency deviation is 25 kHz V^{-1} and the maximum change is 75 kHz, then the maximum signal producing the FM is 3.0 V.

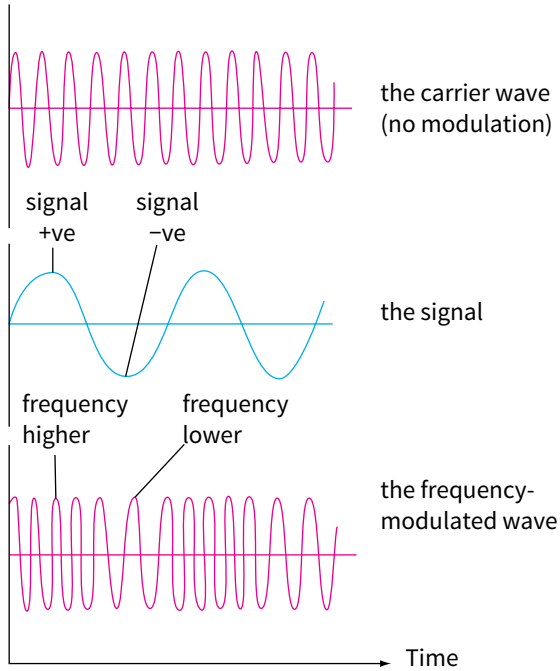


Figure 20.4 Frequency modulation.

WORKED EXAMPLE

2 A carrier wave of frequency 300 kHz and amplitude 5.0 V is frequency modulated by a sinusoidal signal of frequency 6 kHz and amplitude 2.0 V. The frequency deviation of the carrier wave is 30 kHz V^{-1}. Describe the modulated carrier wave produced.

Step 1 Consider the amplitude of the modulated signal. The amplitude of the carrier wave is unchanged at 5.0 V during frequency modulation. The signal alters the frequency of the carrier wave, not its amplitude.

Step 2 Now consider how the signal will modify the carrier frequency. The frequency shift produced by the signal is $\pm 2 \times 30 = \pm 60$ kHz, so the carrier wave varies in frequency between 240 and 360 kHz. This variation in frequency occurs 6000 times every second as the signal varies at this frequency.

QUESTIONS

4 Explain how a frequency-modulated wave changes, when the input signal:

a increases in loudness

b increases in frequency.

5 A signal of frequency 16 kHz and amplitude 3.0 V is used for frequency modulation of a carrier wave of frequency 500 kHz. The frequency deviation of the carrier wave is 8.0 kHz V^{-1}.

a What is the maximum frequency shift produced?

b What is the maximum frequency of the modulated carrier wave?

c How many times per second does the modulated carrier wave increase and decrease in frequency?

Sidebands and bandwidth

A carrier wave contains only one frequency, the carrier wave frequency f_c. When the carrier wave is modulated in amplitude by a single frequency f_m, then the carrier wave is found to contain two more frequencies, known as sideband frequencies, one at a frequency $(f_c - f_m)$ and the other at $(f_c + f_m)$. Figure 20.5 shows these frequencies.

When music or speech is transmitted, the carrier is modulated by a range of frequencies which change with time. Each frequency f_m present in the signal gives rise to an extra pair of frequencies in the modulated wave. The result is a band of frequencies, called the **upper** and **lower sidebands**, stretching above and below the carrier frequency by the value of the highest modulating frequency.

Figure 20.6 shows the frequency spectrum for a carrier wave of frequency 1 MHz modulated with frequencies between 0 and f_m = 15 kHz = 0.015 MHz. The highest

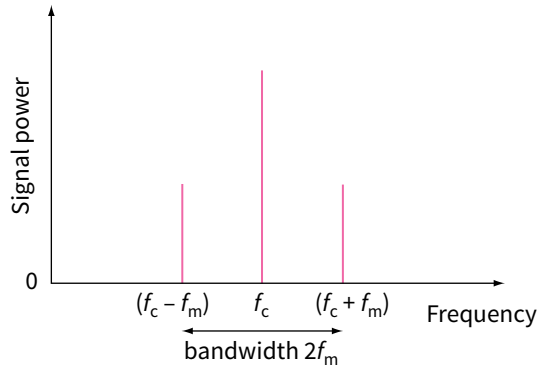


Figure 20.5 The frequency spectrum of a carrier wave modulated in amplitude by a signal of one frequency.

frequency present in the spectrum is $(f_c + f_m) = 1.015\,\text{MHz}$ and the lowest frequency is $(f_c - f_m) = 0.985\,\text{MHz}$.

The actual shape of the sidebands in Figure 20.6 will vary at any instant as the signal changes. The maximum and minimum values are important, as these must not overlap the sidebands from any other radio station.

The value of f_m needed depends on the quality required in the signal. High-quality music only needs frequencies up to 15 kHz, even though the ear can hear frequencies up to 20 kHz. Speech only needs frequencies up to 3.4 kHz for people to understand one another.

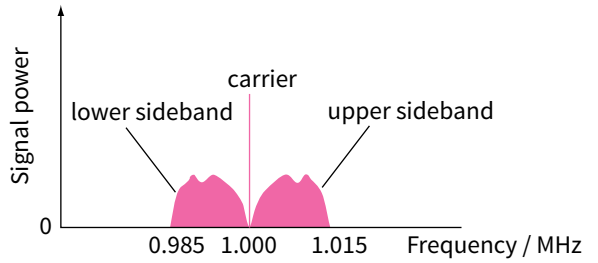


Figure 20.6 The frequency spectrum for an amplitude-modulated wave.

You can see that the modulated carrier wave occupies a region of the spectrum from 0.985 MHz to 1.015 MHz. The **bandwidth** of a signal is the range of frequencies that the signal occupies. In other words, it is the difference between the highest-frequency signal component and the lowest-frequency signal component.

In Figure 20.6 the bandwidth is

$1.015 - 0.985 = 0.030\,\text{MHz}$

In Figure 20.5 the bandwidth is $(f_c + f_m) - (f_c - f_m) = 2f_m$.

The frequency spectrum of a frequency-modulated (FM) carrier wave is more complex. In particular, there are often more than two sideband frequencies for each signal frequency. This means that frequency modulation requires a greater bandwidth for each radio station.

WORKED EXAMPLE

3 Radio stations, which broadcast in the long-wave (LW) region of the electromagnetic spectrum, use carrier frequencies between 140 kHz and 280 kHz. The sidebands are within 4.5 kHz on either side of the carrier frequency. State the bandwidth of each radio station in the LW region of the spectrum and calculate the maximum number of radio stations which can transmit in the LW region.

Step 1 The bandwidth of an individual station is twice the width of an individual sideband:

bandwidth = 2 × 4.5 = 9.0 kHz

Step 2 The LW region is divided into regions of width 9.0 kHz. Hence:

$$\text{number of stations} = \frac{\text{allowed frequency range}}{\text{bandwidth}}$$

$$= \frac{(280 - 140)}{9.0}$$

$$= 15.5 = 15 \text{ stations}$$

Suppose a country decides to increase the quality of music transmitted by each radio station. What happens to the bandwidth and the maximum number of stations in the LW region? Better sound quality requires an increase in the maximum frequency of the signal that modulates the carrier wave, and so the bandwidth needed increases. This decreases the number of available stations in the LW region of the spectrum.

Comparing AM and FM transmissions

You may have noticed crackle on a radio when you switch lights in your house on and off or when there is a lightning strike nearby. The lightning strike or switching a current on or off creates a burst of radio waves. These radio waves produce unwanted electrical interference and change the amplitude of the radio wave received by a radio. Since the **amplitude** of the wave carries the signal, when amplitude modulation is used the output of the radio is affected. Most electrical interference does not affect the **frequency** of the radio wave received by a radio and thus electrical interference affects FM less than AM. FM radio was actually invented to overcome the electrical interference and noise problems of AM radio and this remains an important advantage today.

FM came later than AM and had to use higher frequencies than AM. Although there was extra cost in developing the electronics to work at these higher

frequencies, it was still an advantage. The greater range of frequencies available means that each station can use a higher bandwidth (about 200 kHz, compared to 9 kHz for AM). FM signals typically contain frequencies of 15 kHz or higher and the quality of sound produced is much higher when using FM transmission.

However, AM transmission has a number of advantages.

- The bandwidth needed for each AM transmission is less than for FM transmission. This means that more stations can be included in any given frequency range of the electromagnetic spectrum.
- The actual receiver and transmitter used for AM are less complicated and cheaper than for FM transmission.
- AM transmissions use lower frequencies and have longer wavelengths than FM. This means that these radio waves can diffract some way around the Earth, whereas FM is line-of-sight only. Thus AM can cover a larger area than FM transmissions, for the same power output.

The relative advantages of FM and AM are summarised in Table 20.1.

Advantages of FM	Advantages of AM
less electrical interference and noise	greater area covered by one transmitter
greater bandwidth produces a better quality of sound	smaller bandwidth means more stations available in any frequency range
	cheaper radio sets

Table 20.1 The relative advantages of FM and AM broadcasting.

QUESTIONS

6 Figure 20.7 shows the frequency spectrum of the signal from a radio transmitter. A carrier and two sideband frequencies are present.

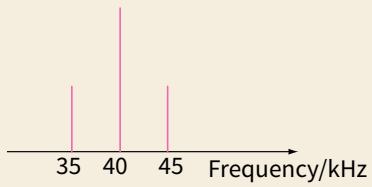


Figure 20.7 For Question 6.

a What is the name of the type of modulation that produces two sideband frequencies?

b What is the carrier frequency?

c What is the frequency of the signal used to modulate the carrier wave?

d What is the bandwidth of the transmitted signal?

7 **a** Calculate the number of separate AM radio stations of bandwidth 9 kHz that are possible in the frequency spectrum available for AM between 530 and 1700 kHz.

b Suggest why FM stations of bandwidth 200 kHz are not used for this range of frequencies.

8 Is the greater bandwidth available on FM an advantage or a disadvantage?

9 FM is used largely in towns and AM in rural settings. Suggest why.

Analogue and digital signals

An analogue quantity is one that can have any value, for example the height of a person. A digital quantity has only a few values, usually just two; for example a person is either male or female.

So far, the signals we have dealt with in this book have been analogue signals. For example, the voltage signal generated by a microphone is an **analogue signal**; the output voltage from the microphone can have any value, within limits, and is an exact representation of the pressure variation in a sound wave.

A **digital signal**, on the other hand, looks completely different and consists of a series of zeroes (0) and ones (1). A 1 in a digital signal is just the presence of a voltage pulse, usually a voltage value of a few volts. A 0 in a digital signal is the absence of a pulse and is a voltage close to 0 V. A typical digital electronic system will interpret any voltage below about 0.3 V as a 0 and any voltage above about 1.5 V as a 1. Small fluctuations in voltage will not be noticed.

Figure 20.8 shows an analogue and a digital signal. The digital signal is the number 0101001101, which is actually a pulse of 0 V followed by a pulse of 3 V and so on.

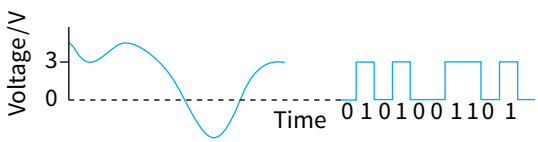


Figure 20.8 Analogue and digital signals.

Advantages of digital signals

Most devices such as microphones or thermistors produce analogue voltage signals. However, digital signals have advantages and it is often worthwhile to change an analogue signal into a digital signal.

The major advantage is that digital signals can deal with 'noise' produced over long distances. All signals, both analogue and digital, become weaker as they travel and they pick up electrical noise. The decrease in strength is known as attenuation, and can be corrected by amplifying the signal at regular intervals during long-distance transmission.

Noise is electrical interference, caused in a number of different ways: by the spark from a car ignition, by induced voltages from the magnetic fields caused by currents around the home, by the radio signals emitted by a mobile phone nearby, and even by the random thermal motion of electrons in a wire or by vibrating atoms. You may have noticed a background hiss on telephone conversations. This also is an example of noise.

Noise is the random, unwanted signal that adds to and distorts a transmitted signal. Amplification of a signal amplifies the noise at the same time as the signal.

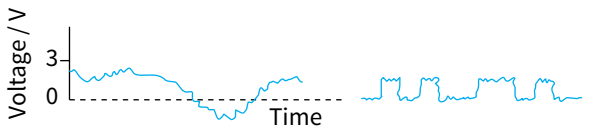


Figure 20.9 Weakened and noisy signals.

Figure 20.9 shows the signals from Figure 20.8 after they have travelled a long distance. You will see that they are lower in amplitude and have unwanted variations, or noise.

There is little improvement possible for an analogue signal; amplification will not remove the noise. However, **regeneration** will remove the noise from a digital signal. The signal is 'cleaned' of the noise and returned to its initial shape.

At the end of a long-distance transmission, an electronic circuit, the regeneration amplifier, receives the digital signal. This electronic circuit expects to receive a pulse of a few volts or no pulse at all; any small variations added to the pulse or the 0 V make no difference. The regeneration amplifier can only give a 0 or a 1 as an output. As long as the noise does not completely change the shape of the signal, then the regeneration amplifier returns the digital signal shown in Figure 20.9 to the perfect pulses shown in Figure 20.8.

Other advantages of using digital signals are:

- Digital signals are compatible with modern technology and can be stored and processed more easily, for example in a computer or on a compact disc (CD).
- Digital electronic systems are, in general, more reliable and easier to design and build.
- Digital signals build in safeguards so that if there is an error in reception it is noticed and parts of the signal can be sent again.

Analogue-to-digital conversion

The key to the digital revolution has been the ability to change speech and music from analogue into digital form in **analogue-to-digital conversion (ADC)** and then convert them back again into analogue form in **digital-to-analogue conversion (DAC)**.

In order to understand this process, you need to be able to count using binary numbers as well as ordinary decimal numbers. The decimal system has base 10 and the number of digits increases by one when going from the number 9 to the number 10. The number 9 has only one digit whereas the number 10 has two digits. The binary number system has base 2 and the number of digits increases when going up from the number 1; so the next number above 1 is 10. The binary number 10 is not the same as the decimal number 10. Table 20.2 compares counting in the decimal system and in the binary system.

Decimal number	Binary number	Decimal number	Binary number
0	0	6	110
1	1	7	111
2	10	8	1000
3	11	9	1001
4	100	10	1010
5	101	11	1011

Table 20.2 Binary and decimal numbers.

For example:

- in the decimal system, the number 243 is a combination of 2×100, 4×10 and 3×1
- in the binary system, the number 111 is a combination of 1×4, 1×2 and 1×1.

Counting in the binary system is very similar to counting in the decimal system except that there are only the digits 0 and 1.

Each digit in the binary number is known as a **bit**. The bit on the left-hand side of a binary number is the **most significant bit (MSB)** and has the highest value.

Large numbers require more bits. Table 20.2 shows numbers containing 4 bits, although 0011 is the correct way of writing a 4-bit number, rather than writing 11. A digital telephone system commonly transmits numbers containing 8 bits and there are $2^8 = 256$ different 8-bit binary numbers.

Changing an analogue signal into a digital signal involves sampling. In analogue-to digital conversion (ADC), **sampling** is the measurement of the analogue signal at regular time intervals.

The value of the sampled signal is used to produce a binary number. Each time that the signal is sampled the ADC produces a binary number of a certain number of bits. Since the sample is taken many times per second, many binary numbers are created, one after the other, and this series of 0s and 1s becomes the digital signal that is transmitted.

The process is illustrated in Figure 20.10 where 4-bit binary numbers are produced.

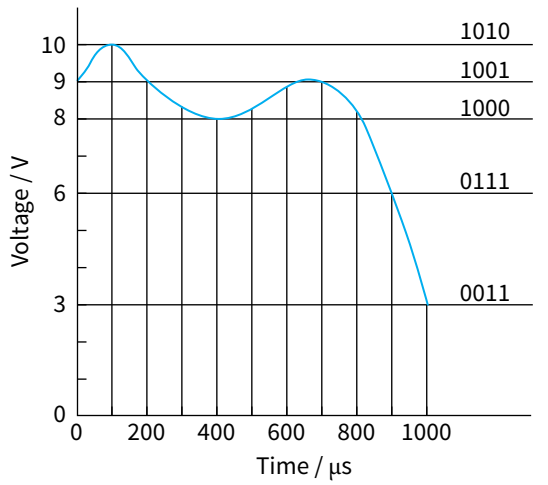


Figure 20.10 Analogue-to-digital conversion.

When the time $t = 0$, the numerical value of the voltage signal is 9 as a decimal number. When converted into binary, this number is 1001. When $t = 100\,\mu s$, the voltage is 10 as a decimal number and 1010 as a binary number.

You will notice that at some values of t the signal is not a whole number on the voltage axis. The nearest number is chosen.

When the output is sampled every 100 μs, a set of binary numbers is produced: (1001), then (1010), then (1001), then (1000), then (1000) and so on. These sets of 4-bit numbers are transmitted one after the other. If they are transmitted a long distance, a regeneration amplifier is used along the way to keep the same pattern of pulses. A digital-to-analogue converter changes the digital signal back into analogue form at the end of the transmission.

Figure 20.11 shows the result of this conversion back into analogue form. The blue circles show the values of the voltage, which are each a decimal number formed from a 4-bit binary number. The black line drawn through the circles is the output signal. Some electronic systems contain extra filter circuits that are able to smooth the output, and they produce the blue line as the final output.

The black line, the output, is clearly not exactly the same as the original signal. There are two reasons for this.

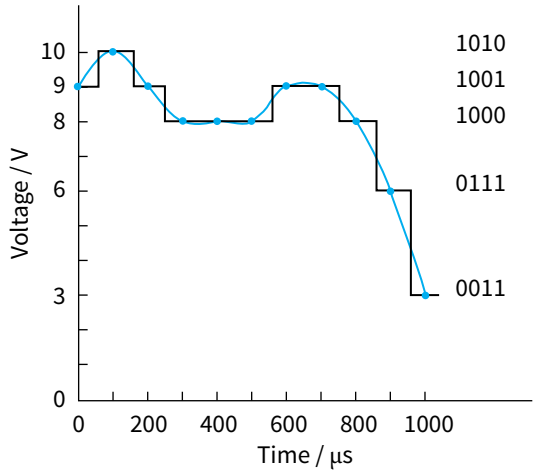


Figure 20.11 Digital-to-analogue conversion.

Firstly, the sampled signal is not always a whole number. For example, in Figure 20.10 when $t = 300\,\mu s$ the actual voltage is 8.3 V but only the number 8 can be sent, not 8.3.

To improve the sampling, the voltage that corresponds to the difference between 0 and 1 must be decreased. In the above example, the difference between 0 and 1 in binary is 1 V and so the signal is 'accurate' to only 1 V. If the difference between 0 and 1 is made to be 0.1 V then 'accuracy' is improved. Adding an extra bit is similar to having an extra significant figure when measuring a voltage as 8.3 V rather than 8 V. Integers up to 10 need a 4-bit binary number. If the system handles numbers to within 0.1 then 10.0 requires an 8-bit number, which has 128 different possible levels.

The other problem is that the sampling rate is not high enough. The sampling rate is the number of samples made per second. In the example in Figure 20.10, the sample is taken every 100 μs and so the sampling rate is to another value and back $\frac{1}{0.0001} = 10\,000$ times a second. If the signal changes to another value and back between one sample and the next then no record is made of that change. Obviously the higher the sampling rate, the closer the final signal will be to the original signal.

The maximum sampling rate required is only twice the highest frequency present in the signal (this is known as Nyquist's theorem). The human ear can hear up to 20 kHz and so the maximum sampling rate needed is 40 kHz for music. For every second of music on a compact disc storing 8-bit numbers sampled 40 000 times a second, 320 000 separate binary digits are stored. The sampling rate required for a telephone system will be lower, since only frequencies up to about 3400 Hz are required for basic speech recognition. Many digital telephone systems sample the input signal 8000 times a second and so only transmit frequencies below 4 kHz.

QUESTIONS

10 Convert the following decimal numbers into binary numbers:

a 14

b 16

11 Convert the following binary numbers into decimal numbers:

a 1111

b 0001011

12 The diagrams in Figure 20.12 show a digital signal at the start of a long cable and at the end of the cable. Both diagrams are drawn to the same horizontal scale (time) and vertical scale (voltage).

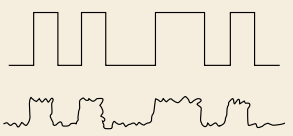


Figure 20.12 For Question 12.

a Explain what feature of the top diagram shows that the signal is digital.

b State and explain **two** advantages of digital transmission over analogue transmission of data.

c State and explain **two** reasons why the signal at the end of the long cable differs from the signal at the start.

Channels of communication

The term **channel of communication** refers to the medium, the path or even the actual frequency range used to convey information from a transmitter to a receiver.

When you listen to a radio, the radio signal may have travelled through the air by a number of different routes. When you talk to someone on a telephone in a different country then the signal may have passed along a **wire-pair**, a **coaxial cable**, through the air by a **microwave link** or been converted into pulses of light and then transmitted down an **optic fibre**. These are all different channels of communication.

Before comparing the different channels you will need to understand another technical term, crosslinking, and be able to calculate signal attenuation.

Crosstalk

You may have experienced crosstalk or **crosslinking** when using a radio or a telephone. If you tune your radio set to one radio station, sometimes you can also hear another station. When talking on the telephone you can suddenly find yourself listening to a completely different telephone conversation; the telephone network has connected your telephone to someone you have not dialled. Crosslinking occurs when a signal, transmitted on one circuit or channel, creates an undesired effect in another circuit or channel.

Signal attenuation

Attenuation is the gradual decrease in the power of a signal the further it travels. The causes of energy loss depend on the type of signal:

- As an electrical signal passes along a wire, there is a voltage drop across the resistance of the wire itself. This reduces the voltage of the signal that arrives at the end of the wire. The energy loss in the wire causes electrical heating in the resistance of the wire (I^2Rt).
- A radio wave spreads out from a transmitter. On its own this spreading causes a decrease in intensity, but there is also a loss in signal strength because of the absorption of energy by the material through which the wave travels.
- Light travelling through an optic fibre may be scattered or absorbed by irregularities in the glass structure.

The decrease in signal power from the transmitted value P_1 to that received P_2 can be very high. The ratio P_2 to P_1 is measured using a logarithmic scale rather than by the simple ratio of the two powers.

The logarithm to base 10 of the ratio gives us the number of bels (B). When multiplied by 10 we obtain the number of **decibels** (dB). Your calculator may show logarithms to base 10 as $\log_{10}$.

They are written here as lg and must not be confused with logarithms to base e, which are usually written as ln.

$$\text{number of B} = \lg\left(\frac{P_2}{P_1}\right)$$

$$\text{number of dB} = 10\lg\left(\frac{P_2}{P_1}\right)$$

For example, suppose P_2 is 1000 times greater than P_1:

$$\text{number of dB} = 10\lg\left(\frac{1000}{1}\right) = 30$$

The number is positive because there is an increase in power – the signal is amplified. Attenuation produces a negative number of decibels; for example, an attenuation of –30 dB means that the received signal is 1000 times smaller than the signal transmitted.

You may be much more familiar with logarithms to base e than with logarithms to base 10. All logarithms obey the same rules; some, which you should know, are:

log of a product	$\log(ab) = \log(a) + \log(b)$
log of a ratio	$\log\left(\frac{a}{b}\right) = \log(a) - \log(b)$
log of a power	$\log(a^n) = n\log(a)$

Overcoming attenuation

In long-distance cables, the attenuation is given as attenuation per unit length, with units such as dB km^{-1}. The attenuation is found from the equation:

$$\text{attenuation per unit length (dB km}^{-1}\text{)} = \frac{\text{attenuation (dB)}}{\text{length of cable (km)}}$$

When a signal travels along a cable, the level of the noise is important. The signal must be distinguishable above the level of the noise. The signal-to-noise ratio, measured in decibels, is given by the expression:

$$\text{signal-to-noise ratio} = 10\lg\left(\frac{\text{signal power}}{\text{noise power}}\right)$$

At regular intervals along a cable, **repeaters** amplify the signal. If the signal is analogue then repeaters also amplify the noise. Multiplying both signal and noise by the same amount keeps the signal-to-noise ratio the same. Regeneration of a digital signal at the same time as amplification removes most of the noise. This ensures that the signal-to-noise ratio remains high.

WORKED EXAMPLE

4 A signal of power 18.0 mW passes along one cable, where the attenuation is 20 dB. It then passes along another cable, where the attenuation is 30 dB. What is the power at the end of the two cables?

Step 1 Apply the decibel equation to each cable in turn.

In the first cable, if the input is P_1 and the output P_2, then:

$$20 = 10\lg\left(\frac{P_1}{P_2}\right)$$

Hint: *Notice that both sides of the equation produce a positive number since* $P_1 > P_2$.

In the second cable, the input is P_2, the output of the first channel. If the output is P_3, then:

$$30 = 10\lg\left(\frac{P_2}{P_3}\right)$$

Step 2 Add the two equations; this gives:

$$50 = 10\left[\lg\left(\frac{P_1}{P_2}\right) + \lg\left(\frac{P_2}{P_3}\right)\right]$$

Applying the 'log of a product rule' gives:

$$50 = 10\lg\left(\frac{P_1}{P_2} \times \frac{P_2}{P_3}\right) = 10\lg\left(\frac{P_1}{P_3}\right)$$

This shows that the total attenuation of the two cables is 50 dB, equal to the sum of the attenuations of the consecutive channels. Hence you can add attenuations to find the total attenuation (but be careful if a signal is being both amplified and attenuated).

Step 3 We have $P_1 = 18$ mW and we need to find P_3. Substituting gives:

$$50 = 10\lg\left(\frac{18}{P_3}\right)$$

so:

$$\lg\left(\frac{18}{P_3}\right) = \frac{50}{10} = 5$$

Taking inverse logs, or pressing the inverse lg button on your calculator, gives:

$$\left(\frac{18}{P_3}\right) = 10^5$$

$$P_3 = 1.8 \times 10^{-4}\ \text{mW}$$

You could apply the decibel equation to each cable in turn and use the output of the first cable as the input to the second cable. You should find that the result is the same.

WORKED EXAMPLE

5 The input signal to a cable has power 1.2×10^{-3} W. The signal attenuation per unit length in the cable is 14 dB km^{-1} and the average noise level along the cable is constant at 1.0×10^{-10} W. An acceptable signal-to-noise ratio is at least 30 dB.

Calculate the minimum acceptable power for the signal and the maximum length of the cable that can be used without a repeater.

Step 1 The signal-to-noise ratio must be at least 30 dB. Hence, using:

$$\text{signal-to-noise ratio} = 10 \lg\left(\frac{\text{signal power}}{\text{noise power}}\right)$$

we have:

$$30 = 10 \lg\left(\frac{P}{1 \times 10^{-10}}\right)$$

where P is the minimum acceptable power. Solving for P gives:

$$= 1.0 \times 10^{-7} \text{ W}$$

Step 2 A repeater is needed to regenerate the signal when the signal-to-noise ratio falls to 30 dB, i.e. its power is 10^3 times the noise level, and this is 1.0×10^{-7} W. We can calculate the attenuation needed to reduce the signal to this level:

$$\text{attenuation} = 10 \lg\left(\frac{1.2 \times 10^{-3}}{1.0 \times 10^{-7}}\right)$$

$$= 41 \text{ dB}$$

Hence the length of cable is $\frac{41}{14} = 2.9$ km.

If the cable is 10 km in length, the total attenuation is:

$14 \text{ dB km}^{-1} \times 10 \text{ km} = 140 \text{ dB}$.

The signal of power 1.2×10^{-3} W is attenuated to a power P where:

$$140 = 10 \lg\left(\frac{0.0012}{P}\right)$$

$$P = 12 \times 10^{-17} \text{ W}$$

You can see that the power in the signal is much smaller than the minimum acceptable power – it is even smaller than the noise level. The signal-to-noise ratio is now $10 \lg (12 \times 10^{-17} / 1.0 \times 10^{-8}) = -79$ dB, smaller than the acceptable +30 dB. A repeater is needed well before the end of the 10 km of cable.

QUESTIONS

13 A signal has an input power 5.0 mW and an output power of 0.000 2 mW. What is the attenuation in dB?

14 The attenuation of a 6.0 mW signal is 30 dB. What is the final power?

15 What is the signal-to-noise ratio when the signal and the noise have equal power?

16 A signal of 1.0 mW passes through an amplifier of gain 30 dB and then along a cable where the attenuation is 18 dB.

a What is the overall gain of the signal in dB?

b What is the output power at the end of the cable?

Comparison of different channels

Each type of signal channel has its good points and its disadvantages, which we will now consider.

Wire-pairs and coaxial cables

The earliest telephones used a pair of wires strung on either side of a pole (Figure 20.13). As the use of electricity became more common, the amount of electrical interference increased, causing crackle and hiss on the line. The potential difference between the two wires is the signal. Each wire acts as an aerial, picks up unwanted electromagnetic waves and distorts the signal.



Figure 20.13 An early telegraph pole.

When two wires are close together, each wire picks up an equal amount of electrical interference. There is no additional potential difference between the two wires and so having the wires close together reduces the interference. Figure 20.14 shows a twisted wire-pair with the wires close together. The connection from your telephone to a socket nearby is likely to use two insulated copper wires placed close together or, more likely, a twisted pair of wires.



Figure 20.14 Twisted wire-pairs in a computer network.

A wire-pair is by far the cheapest transmission medium but it does have some disadvantages.

The changing currents in the wires themselves produce electromagnetic (EM) fields and this makes the wires act as aerials, radiating EM waves. The energy to emit these waves reduces the strength of the signal sent along the wire. As the frequency of the changing current increases, the emission of EM waves increases and so the bandwidth of wire-pairs is low. Energy is also lost in the wires due to electrical heating in the resistance of the wires.

Wire-pairs are often close together in a telephone system and EM waves pass from one wire-pair to another. This is the origin of some forms of crosstalk, where you can hear another telephone conversation.

Wire-pairs are easily 'tapped'. A connection made to each wire allows an unwanted person to hear a telephone conversation. The security of a wire-pair is low.

Coaxial cable, as shown in Figure 20.15, reduces the amount of crosstalk in wire-pairs when transmission occurs at high speed. The copper core and the finely woven copper wire or braid are the two conductors that transmit the signal. The braid is usually connected to earth, so, ideally, the potential of this wire does not change. Electromagnetic (EM) waves do not pass easily through metal and so the braid provides a screen or barrier that reduces the interference that reaches the copper core. An ideal coaxial cable also prevents any emission of EM waves at radio frequencies and has less attenuation than a wire-pair. Although coaxial cable is more expensive than a wire-pair, it can transmit data faster, over longer distances, and with less electrical interference. Coaxial cable often connects a radio transmitter to an aerial, as coaxial cable has a high bandwidth, which can exceed 100 MHz with a cable 30 m in length. It is also slightly more difficult to 'tap' into a coaxial cable than into a wire-pair.

Table 20.3 summarises the advantages and disadvantages of wire-pairs and coaxial cable.

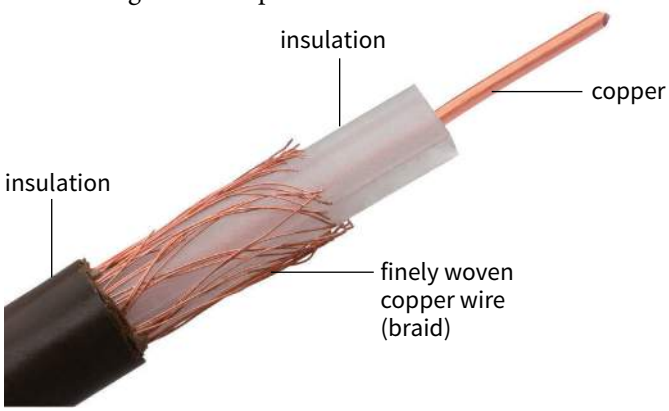


Figure 20.15 Coaxial cable.

Wire-pairs	Coaxial cable
are cheap and convenient	is more expensive
strongly attenuate a signal	is less attenuating
have low bandwidth	has higher bandwidth
pick up some noise and interference	has less electrical interference and noise
suffer from crosstalk	has little crosstalk
have low security	is more secure

Table 20.3 Comparison of wire-pairs and coaxial cable.

Radio waves and microwave links

Radio waves can travel by a number of different paths from a transmitter to a receiver, as shown in Figure 20.16.

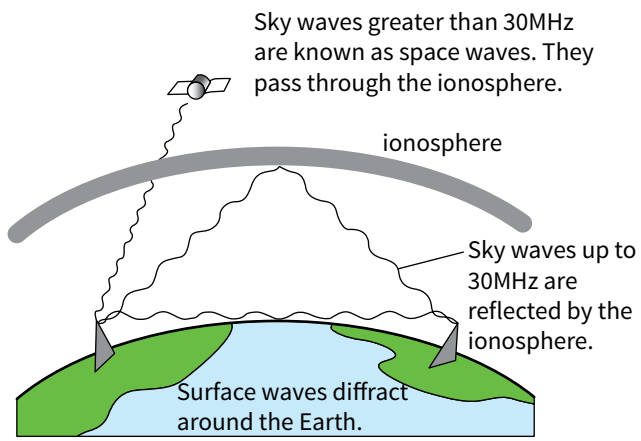


Figure 20.16 Radio wave transmissions.

Surface (ground) waves travel close to the surface of the Earth. Lower frequencies, up to 3 MHz, diffract around the surface of the Earth because of their long wavelengths; this gives them a long range of up to 1000 km. AM broadcasts in the medium-wave (MW) and long-wave (LW) bands travel efficiently as surface waves.

Sky waves are waves above 3 MHz in frequency, which are diffracted only a little by the Earth and travel almost in straight lines. Some waves at these frequencies travel for a short distance as surface waves, but they can only travel about 100 km in this way. Sky waves that travel in the atmosphere may be reflected from a layer of charged particles known as the ionosphere. When these reflected waves reach the Earth's surface they may be reflected back to the ionosphere. Through multiple reflections by the ionosphere and the ground, sky waves can travel for large distances around the Earth. Because the density of charged particles in the ionosphere is constantly changing, the reflection of sky waves is not reliable. Short-wave (SW) radio uses frequencies in the high-frequency (HF) band that reflect from the ionosphere. Your radio set can receive distant SW radio stations at night, when the frequencies reflected by the ionosphere are not absorbed by other regions in the atmosphere.

Space waves are sky waves with a frequency greater than 30 MHz, which pass through the ionosphere. The transmission is line-of-sight, so, if the receiver and transmitter are on the Earth's surface, there must be a clear line between the receiver and the transmitter (not blocked by hills or mountains). Some police and emergency services, as well as television transmissions, use space waves with frequencies above 30 MHz in the very-high-frequency (VHF) and ultra-high-frequency (UHF) bands. The boundary between radio waves and microwaves is not clearly defined, although frequencies above 1 GHz are generally described as microwaves. Microwaves are able to pass through the ionosphere to reach satellites in space. Bluetooth® technology and Wi-Fi® use microwaves for communication around the home. Using Wi-Fi®, microwaves link your laptop to the main computer in your home.

Table 20.4 is a summary of the frequencies used for the different radio transmissions and the distances travelled. The wavelength of any radio wave can be found using the formula $c = f\lambda$ where $c = 3.0 \times 10^8\ \text{m s}^{-1}$, which is the speed of the radio wave and the speed of light. So, for example, a wave of frequency 1 GHz has a wavelength of:

$$\lambda = \frac{3.0 \times 10^8}{1 \times 10^9} = 0.3\ \text{m}$$

	Frequency range	Communication method and waveband	Distance travelled
surface wave	up to 3 MHz	LW and MW radio in the LF band	up to 1000 km
sky wave	3–30 MHz	SW radio in the HF band	worldwide by reflection
space wave	30–300+ MHz	FM radio in the VHF band, TV and mobile phones in the UHF band	line-of-sight
microwave	1–300 GHz	microwave, satellite links and Wi-Fi in the super-high-frequency (SHF) and extra-high-frequency (EHF) bands.	line-of-sight except when retransmitted by satellite

Table 20.4 Data for radio and microwaves.

The radio wave given out by a transmitting aerial travels in all directions. The atmosphere absorbs the wave, the amount of absorption depending on the radio frequency. The distance travelled by a radio wave therefore varies with its frequency. In some cases, the aerial focuses the radio waves towards the receiver by using an aerial shaped as a dish, as shown in Figure 20.17. Without such a dish, the strength of the signal decreases strongly with distance.

The microwave tower shown in Figure 20.17 holds a number of parabolic reflectors or dishes. Each dish points towards a dish on another tower some miles away and they transmit microwaves back and forth between them. The transmission is line-of-sight and the height of the tower increases the distance of transmission.



Figure 20.17 A microwave tower.

The bandwidth available increases as the frequency of the wave increases. As microwaves have a high bandwidth, they can carry many telephone conversations at once. They are also very secure and difficult to tap into, as the beam of microwaves that travels between the two dishes is narrow and does not spread out. Until fibre optic cable was available, microwave links carried the majority of long-distance telephone conversations.

To summarise, high-frequency radio waves and microwaves:

- have high bandwidth and can carry a large amount of information
- can be transmitted as narrow beams which are more secure
- are line-of-sight and often use a satellite or microwave link.

Satellites and optic fibres

Transmissions in the LW and MW bands use a surface wave and do not travel further than about 1000 km. For long-distance communication, it is possible to transmit using a sky wave or using a space wave and a satellite.

Figures 20.1 and 20.18 both show a communications satellite in space. The satellite receives a space wave from a transmitter on Earth, the uplink, with a carrier frequency in the microwave region. Because the satellite can only reflect a tiny fraction of the signal sent from Earth, the reflected signal received back on Earth would be far too small. Instead, the satellite re-transmits the signal it receives as the downlink back to Earth, on another frequency and with more power than it received. If the downlink and uplink frequencies were the same, then the much larger signal sent from Earth would swamp the signal sent from the satellite, so different frequencies are used. The satellite transmits the signal back to an individual satellite dish back on Earth or to many dishes, over a wide area, particularly when broadcasting television programmes.

The first communications satellites used a frequency of 6 GHz for the uplink and 4 GHz for the downlink, but now even higher frequencies are used.

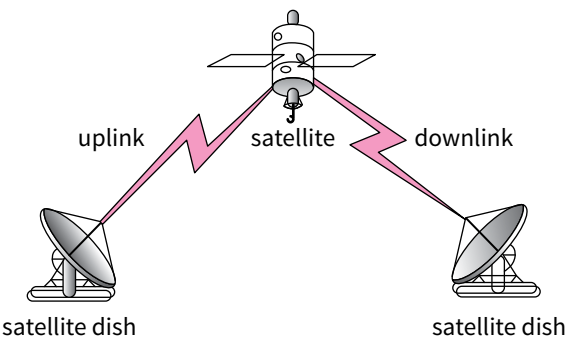


Figure 20.18 A satellite system.

Here are some advantages of communication by satellite rather than by sky wave:

- The concentration of ions in the ionosphere is constantly changing and reflection of the sky wave is not always possible; sometimes layers in the ionosphere even absorb radio frequencies.
- The satellite boosts the signal for its return to Earth and provides a stronger signal than is obtained by reflection from the ionosphere.
- Satellite communication uses higher frequencies, which have higher bandwidth and can carry more information per second.
- Only a few frequencies in the MW and SW bands are available. More frequencies are available for communicating if a satellite uses higher frequencies.

In order to obtain a constant link between the satellite and the satellite dish on Earth, it is essential that the satellite dish always points towards the satellite. If the satellite moves across the sky, the dish must move to track the movement of the satellite. To avoid this problem many communications satellites rotate around the Earth in a geostationary orbit. They orbit the Earth in the same direction as the Earth rotates, at a height of 36 000 km above the Earth's surface. At this height, each satellite has a period of rotation of 24 hours. This means that each satellite naturally takes 1 day to orbit the Earth, exactly the time that it takes a point on the surface of the Earth to make one complete rotation. A geostationary satellite is in orbit above the equator and it never appears to move when viewed from any point on the Earth. The satellite does not rotate with the same **speed** as a point on the Earth because its orbit is far larger than the circumference of the Earth.

There are many other satellites around the Earth. Some of these are in polar orbit. These satellites commonly travel above the North and South Poles in a time much shorter than a day. They are usually closer to the Earth than geostationary satellites and are used for surface observation and as weather satellites. At a commonly used height of 1000 km above the Earth's surface, the period of rotation around the Earth is only 100 minutes. Being closer to the Earth, polar-orbit satellites can see smaller detail when used for observation and espionage. As they orbit from the North to the South Pole, the Earth rotates underneath them and so they pass over the whole Earth in a 24-hour period.

There is always a delay in sending a message to a satellite because the satellite is high above the Earth. This delay can be annoying when talking by telephone. For example, if the satellite is directly overhead and the signal travels a distance of 72 000 km up to the satellite and then down to the other person, the time delay is 0.24 s. The reply

from the other person also takes 0.24 s and so there always seems to be a delay of at least 0.48 s in the conversation.

Geostationary satellites used for communication can transmit to each other around the world but they cannot always receive from regions close to the poles, as the curvature of the Earth blocks the signal. The delay when communicating with a polar satellite is much smaller but you may have to wait until the satellite is overhead to transmit or receive.

The features of a geostationary satellite are:

- the satellite rotates with the same period as the Earth
- the satellite is in orbit above the equator with a period of 1 day
- the satellite appears to remain fixed in position above a point on the equator and so satellite dishes do not need to be moved.

Compared to a geostationary satellite, a satellite in polar orbit:

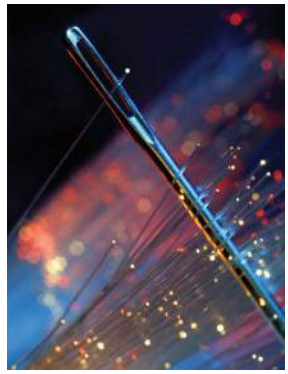
- travels from pole to pole, with an shorter period of orbit
- is at a smaller height above the Earth and can detect objects of smaller detail
- is not always in the same position relative to the Earth and so dishes must be moved
- has smaller delay times.

An alternative for long-distance communications is the optic fibre, a very thin glass or plastic fibre that carries light or infrared. Optic fibres use glass and infrared for long distances, rather than plastic or light, as the glass can be very pure and does not absorb or scatter infrared. As optic fibres have very low signal attenuation, the distance between repeater amplifiers can be high.

Figure 20.19 shows a photograph of a number of optic fibres and the internal structure of one fibre. The three rays of light are totally internally reflected from one end of the fibre to the other.

When used for communication, an electrical signal causes a laser or a light-emitting diode (LED) to emit pulses of light or infrared, with a frequency of the order of 2×10^{14} Hz (2×10^{8} MHz). Because the frequency is so high, the potential bandwidth available is also very high. The pulses of light or infrared provide the digital signal that passes along the fibre. With a cable containing more than a hundred fibres and each fibre carrying a large number of pulses per second, the whole cable can carry ten million telephone conversations at the same time.

Fibre optic cables have replaced the use of satellites for long-distance transmission. Just a few fibre optic cables running across the oceans and from city to city link almost the whole world and make the internet possible – the 'information superhighway'. The delay between



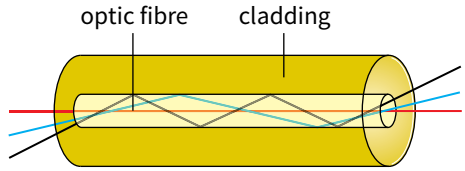


Figure 20.19 An optic fibre passing through the eye of a needle, and its internal structure.

transmission and reception is less than with a satellite as the distances travelled round the world by the signal are less than up to a satellite and back down again. The disadvantages of fibre optic cables are that an electrical signal must first be converted to pulses of light and the optic fibres are difficult to connect to one another as two fibres cannot just be glued together.

Compared to a metal cable, a fibre optic cable:

- has much greater bandwidth and can carry more information per second
- has less signal attenuation, so repeater and regeneration amplifiers can be further apart
- is more difficult to tap, making the data it carries more secure
- does not suffer from electrical interference and crosstalk
- weighs less and so large lengths can be handled more easily
- is immune to lightning and the effects of nearby power lines
- can be used in flammable situations as no sparks are produced
- is cheaper than the same length of copper wire.

QUESTION

17 A television signal can be transmitted using a **coaxial cable** to pass an **analogue signal**, using a **space wave** in the UHF band, using a **sky wave** linked to a satellite or using fibre optic cable and the internet.

a Explain what is meant by the terms in **bold** type.

b State approximate values for the frequencies and wavelengths used by the carrier wave in each case.

c Give **one** advantage and **one** disadvantage of each of these four methods.

Summary

- Modulation varies either the amplitude or frequency of a carrier wave to carry information in the signal.
- Bandwidth is the range of frequencies present in a broadcast signal.
- FM broadcasts have higher bandwidth and sound quality than AM.
- Analogue and digital signals transmit information.
- Analogue-to-digital conversion (ADC) turns analogue signals into digital signals and digital-to-analogue conversion (DAC) turns them back again.
- Regeneration removes noise and interference from digital signals.
- For good reproduction of a signal, the sampling rate and number of bits should be high.
- Wire-pairs, coaxial cables, radio waves, microwaves and optic fibres transmit signals.
- Attenuation of a signal is measured in dB, where number of dB = $10 \lg\left(\frac{P_1}{P_2}\right)$.
- Satellites for communication are often in geostationary or polar orbits.

End-of-chapter questions

1 **a** Draw a sketch graph of an AM wave and use your graph to explain how the AM wave has been formed from the carrier wave and a signal. Refer to the frequencies involved. [4]

b Draw a sketch graph of an FM wave and use your graph to explain how the FM wave has been formed from a carrier wave and a signal. Refer to the frequencies involved. [4]

2 The output of a microphone is an analogue signal with a bandwidth of 3.4 kHz.

a Explain what is meant by

i **analogue** [2]

ii **bandwidth**. [1]

b Compare the bandwidth of the microphone with the typical range of frequencies that can be heard by the human ear. Comment on the difference between the two values. [2]

3 An LED provides input power of 1.26 mW to an optic fibre of length 60 m. The output at the other end of the fibre is 1.12 mW.

a Calculate the attenuation in the optic fibre. [2]

b Calculate the attenuation per unit length in the optic fibre. [2]

4 The signal attenuation per unit length of an optic fibre is 0.30 dB km^{-1}. An input signal to the optic fibre is 100 dB above the noise level. The level of the noise remains constant along the optic fibre at 6.0 pW.

Calculate:

a the power of the input signal to the optic fibre [3]

b the maximum length of the optic fibre used, if the signal at the end of the fibre is to remain at least 30 dB above the noise level. [2]

5 **a** Describe the orbit of a geostationary satellite. [2]

b State a typical wavelength for communication between the Earth's surface and a geostationary satellite. [1]

c State **one** advantage and **one** disadvantage of the use for telephone communication of a geostationary satellite compared with a satellite in polar orbit. [4]

6 Figure 20.15 shows a coaxial cable.

a State and explain the purpose of the fine woven copper wire or braid. [2]

b Optic fibre has a larger bandwidth than a coaxial cable. Explain why increased bandwidth has reduced the cost of telephone calls to distant countries. [1]

c **i** Explain what is meant by **electrical noise**. [1]

ii State **two** causes of noise in a copper cable. [2]

7 Radio signals may be transmitted by surface (ground) waves, by space waves or by sky waves. State a typical value for the frequency and the maximum distance of transmission for each type of wave. [6]

8 Figure 20.20 shows an amplitude-modulated radio wave carrying a signal.

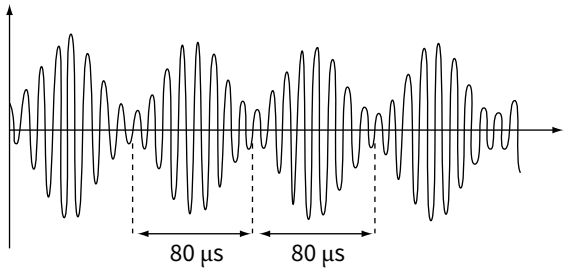


Figure 20.20 For End-of-chapter Question 8.

a Determine the radio frequency of the carrier wave. [2]

b Calculate the frequency of the signal. [2]

c Draw the frequency spectrum of the modulated radio wave [3]

9 Figure 20.21 shows the variation with time of a signal voltage, *V*, over a 5-hour period.

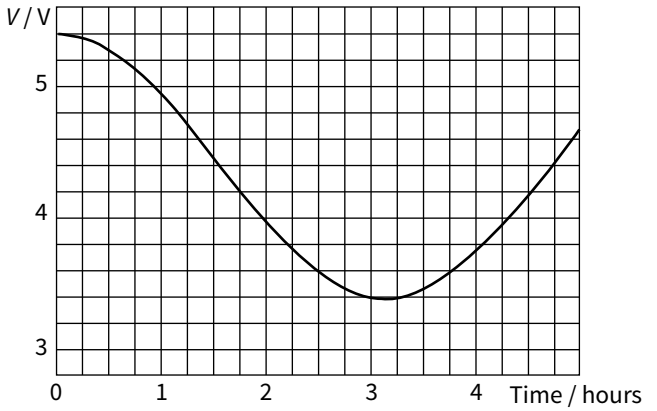


Figure 20.21 For End-of-chapter Question 9.

a State the name of the type of signal shown in the diagram. [1]

b The signal is turned into a digital signal with 4-bit binary numbers. The value of *V* is (0000) when *V* is 3 V and each subsequent binary unit represents an extra 0.2 V. Copy and complete the table to give the value of *V* as a decimal number and as a binary number for times of 0, 2 and 4 hours. [3]

Time / hours	*V* / V (decimal)	*V* / V (binary)
0		
2		
4		

c The digital signal that is produced cannot be used to produce a perfect reproduction of the original signal.

i Explain why this is the case. [2]

ii Explain how the sampling process can be improved. [2]

10 A cable of length 20 km signal transmits a signal from one end to the other. The attenuation per unit length of the cable is 10 dB km^{-1}. Four repeater amplifiers along the cable each have a gain of 40 dB.

a Calculate:

i the total attenuation caused by the cable [1]

ii the total gain caused by the amplifiers. [1]

b The input signal has power 200 mW. Calculate the output power of the signal from the cable. [2]

c Calculate the output power of the signal from the cable if repeater amplifiers are not used. [1]

11 A communications satellite in geostationary orbit receives an uplink signal at a frequency of 14 GHz and transmits a downlink signal back to Earth at a frequency of 11 GHz.

a Describe the key features of a **geostationary** satellite. [2]

b Suggest a reason why different frequencies are used for the uplink and the downlink. [1]

c The uplink signal has a power of 2 kW and the downlink signal 20 W. Suggest why the power in the two signals is so different. [1]



Chapter 21: Thermal physics

Learning outcomes

You should be able to:

- use a simple kinetic model of matter to explain phenomena associated with changes of state
- explain what is meant by temperature and outline how it can be measured
- outline the advantage of the thermodynamic scale of temperature
- relate the internal energy of an object to the energy of its particles
- use the first law of thermodynamics
- define and use specific heat capacity and specific latent heat, and outline how these quantities can be measured

From water to steam

When water boils, it changes state – it turns to steam. A liquid has become a gas. This is a familiar process, but Figure 21.1 shows a dramatic example of such a change of state. This is a geyser in New Zealand, formed when water is trapped underground where it is in contact with hot rocks. The temperature and pressure of the water build up until it suddenly erupts above the surface to form a tall plume of scalding water and steam.



Figure 21.1 At regular intervals of time, the White Lady Geyser, near Rotorua in New Zealand, throws up a plume of water and steam.

Changes of state

The kinetic model of matter can be used to describe the structure of solids, liquids and gases. You should recall that the kinetic model describes the behaviour of matter in terms of moving particles (atoms, molecules, etc.). Figure 21.2 should remind you of how we picture the three states of matter at the atomic scale:

- In a solid, the particles are close together, tightly bonded to their neighbours, and vibrating about fixed positions.
- In a gas, the particles have broken free from their neighbours; they are widely separated and are free to move around within their container.

QUESTION

1 Describe a liquid in terms of the arrangement of its particles, the bonding between them and their motion.

In this chapter, we will extend these ideas to look at the energy changes involved when materials are heated and cooled.

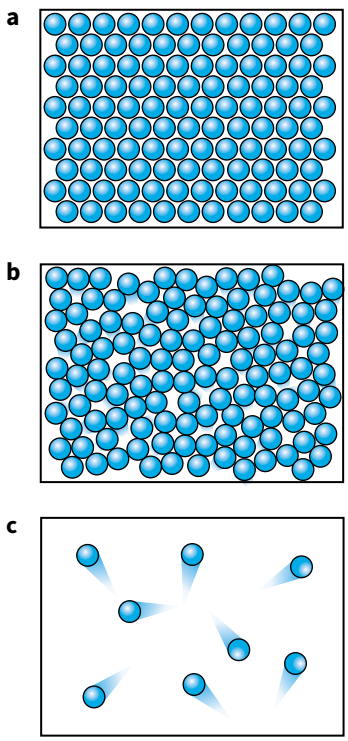


Figure 21.2 Typical arrangements of atoms in **a** a solid, **b** a liquid and **c** a gas.

Energy changes

Energy must be supplied to raise the temperature of a solid, to melt it, to heat the liquid and to boil it. Where does this energy go? It is worth taking a close look at a single change of state and thinking about what is happening on the atomic scale. Figure 21.3a shows a suitable arrangement. A test tube containing octadecanoic acid (a white, waxy substance at room temperature) is warmed in a water bath. At 80 °C, the substance is a clear liquid. The tube is then placed in a rack and allowed to cool. Its temperature is monitored, either with a thermometer or with a temperature probe and datalogger. Figure 21.3b shows typical results.

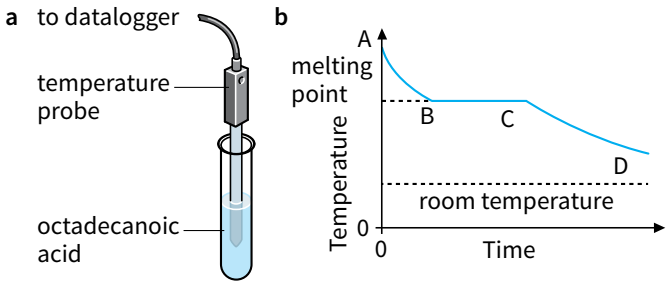


Figure 21.3 **a** Apparatus for obtaining a cooling curve, and **b** typical results.

The temperature drops rapidly at first, then more slowly as it approaches room temperature. The important section of the graph is the region BC. The temperature remains steady for some time. The clear liquid is gradually returning to its white, waxy solid state. It is essential to note that energy is still being lost even though the temperature is not decreasing. When no liquid remains, the temperature starts to drop again.

From the graph, we can deduce the melting point of octadecanoic acid. This is a technique used to help identify substances by finding their melting points.

Heating ice

In some ways, it is easier to think of the experiment above in reverse. What happens when we heat a substance?

Imagine taking some ice from the deep freeze. Put the ice in a well-insulated container and heat it at a steady rate. Its temperature will rise; eventually we will have a container of water vapour. (Note that water vapour is an invisible gas; the 'steam' that you see when a kettle boils is not a gas but a cloud of tiny droplets of liquid water.)

Figure 21.4 shows the results we might expect if we could carry out this idealised experiment. Energy is supplied to the ice at a constant rate. We will consider the different sections of this graph in some detail, in order to describe where the energy is going at each stage.

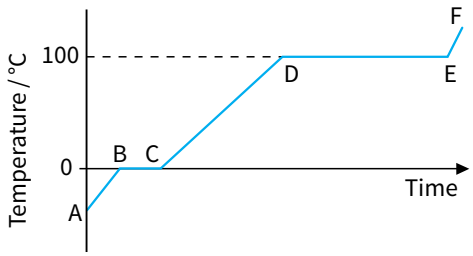


Figure 21.4 A graph of temperature against time for water, heated at a steady rate.

We need to think about the kinetic and potential energies of the molecules. If they move around more freely and faster, their kinetic energy has increased. If they break free of their neighbours and become more disordered, their electrical potential energy has increased.

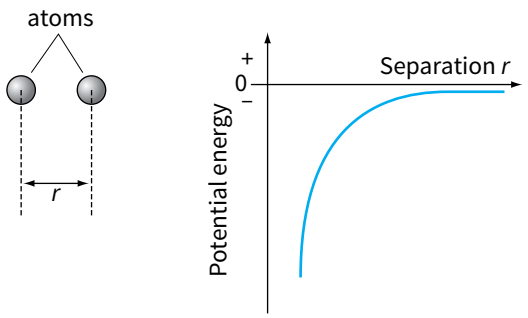


Figure 21.5 The electrical potential energy of atoms is negative and increases as they get further apart.

You know that the kinetic energy of a particle is the energy it has due to its motion. Figure 21.5 shows how the electrical potential energy of two isolated atoms depends on their separation. Work must be done (energy must be put in) to separate neighbouring atoms – think about the work you must do to snap a piece of plastic or to tear a sheet of paper. The graph shows that:

- the electrical potential energy of two atoms very close together is large and negative
- as the separation of the atoms increases, their potential energy also increases
- when the atoms are completely separated, their potential energy is maximum and has a value of zero.

Now look at the graph shown in Figure 21.4.

Section AB

The ice starts below 0 °C; its temperature rises. The molecules gain energy and vibrate more and more. Their vibrational kinetic energy is increasing. There is very little change in the mean separation between the molecules and hence very little change in their electrical potential energy.

Section BC

The ice melts at 0 °C. The molecules become more disordered. There is a modest increase in their electrical potential energy.

Section CD

The ice has become water. Its temperature rises towards 100 °C. The molecules move increasingly rapidly. Their kinetic energy is increasing. There is very little change in the mean separation between the molecules and therefore very little change in their electrical potential energy.

Section DE

The water is boiling. The molecules are becoming completely separate from one another. There is a large increase in the separation between the molecules and hence their electrical potential energy has increased greatly. Their movement becomes very disorderly.

Section EF

The steam is being heated above 100 °C. The molecules move even faster. Their kinetic energy is increasing. The molecules have maximum electrical potential energy of zero.

You should see that, when water is heated, each change of state (melting, boiling) involves the following:

- there must be an input of energy
- the temperature does not change
- the molecules are breaking free of one another
- their potential energy is increasing.

In between the changes of state:

- the input of energy raises the temperature of the substance
- the molecules move faster
- their kinetic energy is increasing.

The hardest point to appreciate is that you can put energy into the system without its temperature rising. This happens during any change of state; the energy goes to breaking the bonds between neighbouring molecules. The energy which must be supplied to cause a change of state is sometimes called 'latent heat'. The word 'latent' means 'hidden' and refers to the fact that, when you melt something, its temperature does not rise and the energy that you have put in seems to have disappeared.

It may help to think of temperature as a measure of the average kinetic energy of the molecules. When you put a thermometer in some water to measure its temperature, the water molecules collide with the thermometer and share their kinetic energy with it. At a change of state, there is no change in kinetic energy, so there is no change in temperature.

Notice that melting the ice (section BC) takes much less energy than boiling the same amount of water (section DE). This is because, when a solid melts, the molecules are still bonded to most of their immediate neighbours. When a liquid boils, each molecule breaks free of all of its neighbours. Melting may involve the breaking of one or two bonds per molecule, whereas boiling involves breaking eight or nine.

Evaporation

A liquid does not have to boil to change into a gas. A puddle of rain-water dries up without having to be heated to 100 °C. When a liquid changes to a gas without boiling, we call this **evaporation**.

Any liquid has some vapour associated with it. If we think about the microscopic picture of this, we can see why (Figure 21.6). Within the liquid, molecules are moving about. Some move faster than others, and can break free from the bulk of the liquid. They form the vapour above the liquid. Some molecules from the vapour may come back into contact with the surface of the liquid, and return to the liquid. However, there is a net outflow of energetic molecules from the liquid, and eventually it will evaporate away completely.

You may have had your skin swabbed with alcohol or ether before an injection. You will have noticed how cold your skin becomes as the volatile liquid evaporates. Similarly, you can become very cold if you get wet and

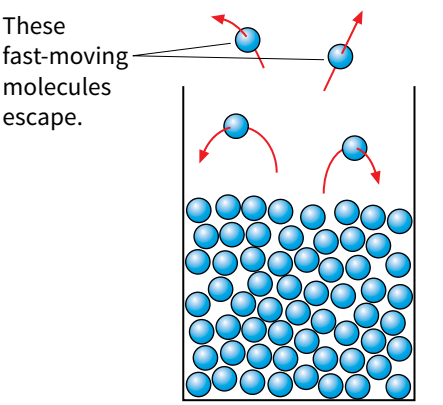


Figure 21.6 Fast-moving molecules leave the surface of a liquid – this is evaporation.

stand around in a windy place. This cooling of a liquid is a very important aspect of evaporation.

When a liquid evaporates, it is the most energetic molecules that are most likely to escape. This leaves molecules with a below-average kinetic energy. Since temperature is a measure of the average kinetic energy of the molecules, it follows that the temperature of the evaporating liquid must fall.

QUESTION

2 Use the kinetic model of matter to explain the following:

- **a** If you leave a pan of water on the hob for a long time, it does not all boil away as soon as the temperature reaches 100 °C.
- **b** It takes less energy to melt a 1.0 kg block of ice at 0 °C than to boil away 1.0 kg of water at 100 °C.
- **c** When a dog is overheated, it pants.

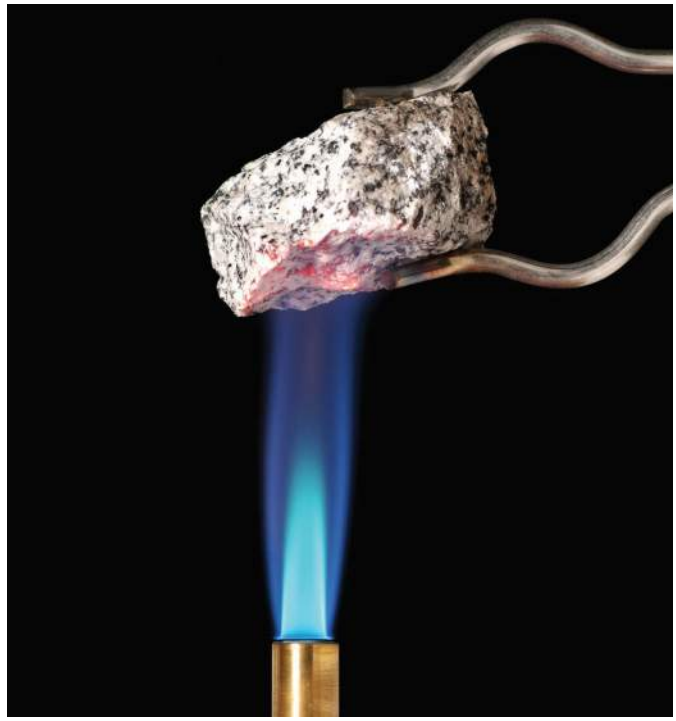


Figure 21.7 Increasing the internal energy of a stone.

Internal energy

All matter is made up of particles, which we will refer to here as 'molecules'. Matter can have energy. For example, if we lift up a stone, it has gravitational potential energy. If we throw it, it has kinetic energy. Kinetic and potential energies are the two general forms of energy. We consider the stone's potential and kinetic energies to be properties or attributes of the stone itself; we calculate their values (mgh and $\frac{1}{2}mv^2$) using the mass and speed of the stone.

Now think about another way in which we could increase the energy of the stone: we could heat it (Figure 21.7). Now where does the energy from the heater go? The stone's gravitational potential and kinetic energies do not increase; it is not higher or faster than before. The energy seems to have disappeared into the stone.

Of course, you already know the answer to this. The stone gets hotter, and that means that the molecules which make up the stone have more energy, both kinetic and electrical potential. They vibrate more and faster, and they move a little further apart. This energy of the molecules is known as the **internal energy** of the stone. The internal energy of a system (e.g. the heated stone) is defined as follows:

> The internal energy of a system is the sum of the random distribution of kinetic and potential energies of its atoms or molecules.

Molecular energy

Earlier in this chapter, where we studied the phases of matter, we saw how solids, liquids and gases could be characterised by differences in the arrangement, order and motion of their molecules. We could equally have said that, in the three phases, the molecules have different amounts of kinetic and potential energy.

Now, it is a simple problem to find the internal energy of an amount of matter. We add up the kinetic and potential energies associated with all the molecules in that matter. For example, consider the gas shown in Figure 21.8. There are ten molecules in the box, each having kinetic and potential energy. We can work out what all of these are and add them together, to get the total internal energy of the gas in the box.

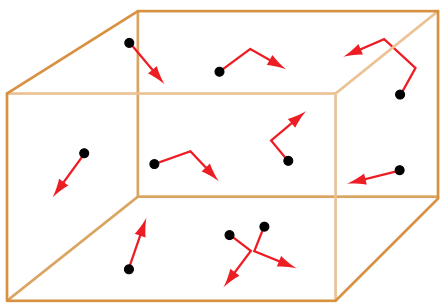


Figure 21.8 The molecules of a gas have both kinetic and potential energy.

Changing internal energy

There are two obvious ways in which we can increase the internal energy of some gas: we can heat it (Figure 21.9a), or we can do work on it by compressing it (Figure 21.9b).

Heating a gas

The walls of the container become hot and so its molecules vibrate more vigorously. The molecules of the cool gas strike the walls and bounce off faster. They have gained kinetic energy, and we say the temperature has risen.

Doing work on a gas

In this case, a wall of the container is being pushed inwards. The molecules of the cool gas strike a moving wall and bounce off faster. They have gained kinetic energy and again the temperature has risen. This explains why a gas gets hotter when it is compressed.

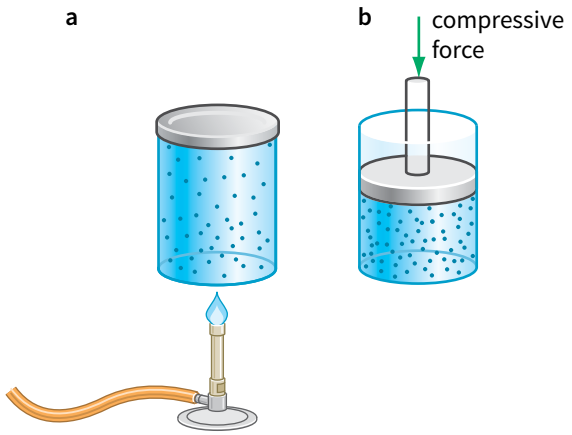


Figure 21.9 Two ways to increase the internal energy of a gas: **a** by heating it, and **b** by compressing it.

There are other ways in which the internal energy of a system can be increased: by passing an electric current through it, for example. However, doing work and heating are all we need to consider here.

The internal energy of a gas can also decrease; for example, if it loses heat to its surroundings, or if it expands so that it does work on its surroundings.

First law of thermodynamics

You will be familiar with the idea that energy is **conserved**; that is, energy cannot simply disappear, or appear from nowhere. This means that, for example, all the energy put into a gas by heating it and by doing work on it must end up in the gas; it increases the internal energy of the gas.

We can write this as an equation:

increase in internal energy = energy supplied by heating + energy supplied by doing work

In symbols:

$$\Delta U = q + w$$

This is known as the **first law of thermodynamics** and is a formal statement of the principle of conservation of energy. (It applies to all situations, not simply to a mass of gas.) Since you have learned previously that energy is conserved, it may seem to be a simple idea, but it took scientists a good many decades to understand the nature of energy and to appreciate that it is conserved.

QUESTION

3 Use the first law of thermodynamics to answer the following.

a A gas is heated by supplying it with 250 kJ of energy; at the same time, it is compressed so that 500 kJ of work is done on the gas. Calculate the change in the internal energy of the gas.

b The same gas is heated as before with 250 kJ of energy. This time the gas is allowed to expand so that it does 200 kJ of work on its surroundings. Calculate the change in the internal energy of the gas.

The meaning of temperature

Picture a beaker of boiling water. You want to measure its temperature, so you pick up a thermometer which is lying on the bench. The thermometer reads 20 °C. You place the thermometer in the water and the reading goes up … 30 °C, 40 °C, 50 °C. This tells you that the thermometer is getting hotter; energy is being transferred from the water to the thermometer.

Eventually, the thermometer reading reaches 100 °C and it stops rising. Because the reading is steady, you can deduce that energy is no longer being transferred to the thermometer and so its scale tells you the temperature of the water.

This simple, everyday activity illustrates several points:

- We are used to the idea that a thermometer shows the temperature of something with which it is in contact. In fact, it tells you **its own temperature**. As the reading on the scale was rising, it wasn't showing the temperature of the water. It was showing that the temperature of the thermometer was rising.
- Energy is transferred from a hotter object to a cooler one. The temperature of the water was greater than the temperature of the thermometer, so energy transferred from one to the other.
- When two objects are at the same temperature, there is no transfer of energy between them. That is what happened when the thermometer reached the same temperature as the water, so it was safe to say that the reading on the thermometer was the same as the temperature of the water.

From this, you can see that temperature tells us about the direction in which energy flows. If two objects are placed in contact (so that energy can flow between them), it will flow from the hotter to the cooler. Energy flowing from a region of higher temperature to a region of lower temperature is called **thermal energy**. (Here, we are not concerned with the mechanism by which the energy is transferred. It may be by conduction, convection or radiation.)

When two objects, in contact with each other, are at the same temperature, there will be no net transfer of thermal energy between them. We say that they are in **thermal equilibrium** with each other – see Figure 21.10.

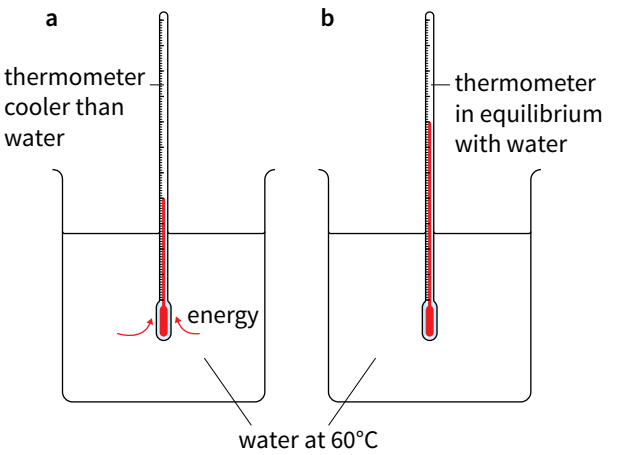


Figure 21.10 **a** Thermal energy is transferred from the hot water to the cooler thermometer because of the temperature difference between them. **b** When they are at the same temperature, there is no transfer of thermal energy and they are in thermal equilibrium.

The thermodynamic (Kelvin) scale

The Celsius scale of temperature is a familiar, everyday scale of temperature. It is based on the properties of water. It takes two fixed points, the melting point of pure ice and the boiling point of pure water, and divides the range between them into 100 equal intervals.

There is nothing special about these two fixed points. In fact, both change if the pressure changes or if the water is impure. The **thermodynamic scale**, also known as the Kelvin scale, is a better scale in that one of its fixed points, **absolute zero**, has a greater significance than either of the Celsius fixed points.

It is not possible to have a temperature lower than 0 K. Sometimes it is suggested that, at this temperature, matter has no energy left in it. This is not strictly true; it is more correct to say that, for any matter at absolute zero, it is impossible to **remove** any more energy from it. Hence absolute zero is the temperature at which all substances have the minimum internal energy. (The kinetic energy of the atoms or molecules is zero and their electrical potential energy is minimum.)

We use different symbols to represent temperatures on these two scales: θ for the Celsius scale, and T for the thermodynamic (Kelvin) scale. To convert between the two scales, we use these relationships:

$$\theta\,(°\text{C}) = T\,(\text{K}) - 273.15$$

$$T\,(\text{K}) = \theta\,(°\text{C}) + 273.15$$

For most practical purposes, we round off the conversion factor to 273 as shown in the conversion chart (Figure 21.11).

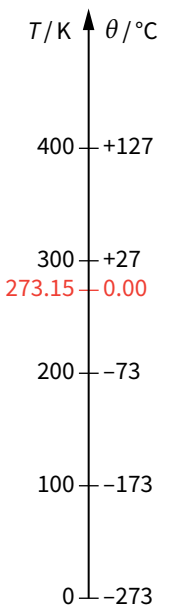


Figure 21.11 A conversion chart relating temperatures on the thermodynamic (Kelvin) and Celsius scales.

The thermodynamic scale of temperature is designed to overcome a problem with scales of temperature such as the Celsius scale, which depends on the melting point and boiling point of pure water. To measure a temperature on this scale, you might use a liquid-in-glass thermometer. However, the expansion of a liquid may be non-linear. This means that if you compare the readings from two different types of liquid-in-glass thermometer, for example a mercury thermometer and an alcohol thermometer, you can only be sure that they will agree at the two fixed points on the Celsius scale. At other temperatures, their readings may differ.

The thermodynamic scale is said to be an absolute scale as it is not defined in terms of a property of any particular substance. It is based on the idea that the average kinetic energy of the particles of a substance increases with temperature. The average kinetic energy is the same for all substances at a particular thermodynamic temperature; it does not depend on the material itself. In fact, as you will see in Chapter 22, the average kinetic energy of a gas molecule is proportional to the thermodynamic temperature. So, if we can measure the average kinetic energy of the particles of a substance, we can deduce the temperature of that substance.

The thermodynamic scale has two fixed points:

- absolute zero, which is defined as 0 K
- the triple point of water, the temperature at which ice, water and water vapour can co-exist, which is defined as 273.16 K (equal to 0.01 °C).

So the gap between absolute zero and the triple point of water is divided into 273.16 equal divisions. Each division is 1 K. The scale is defined in this slightly odd way so that the scale divisions on the thermodynamic scale are equal in size to the divisions on the Celsius scale, making conversions between the two scales relatively easy.

A **change** in temperature of 1 K is thus equal to a **change** in temperature of 1 °C.

Thermometers

A thermometer is any device which can be used to measure temperature. Each type of thermometer makes use of some physical property of a material which changes with temperature. The most familiar is the length of a column of liquid in a tube, which gets longer as the temperature increases because the liquid expands – this is how a liquid-in-glass thermometer works. Other properties which can be used as the basis of thermometers include:

- the resistance of an electrical resistor or thermistor
- the voltage produced by a thermocouple
- the colour of an electrically heated wire
- the volume of a fixed mass of gas at constant pressure.

In each case, the thermometer must be calibrated at two or more known temperatures (such as the melting and boiling points of water, which correspond to 0 °C and 100 °C), and the scale between divided into equal divisions. There is no guarantee that two thermometers will agree with each other except at these fixed points. Now we will look in detail at two types of electrical thermometer.

QUESTIONS

4 **a** Convert each of the following temperatures from the Celsius scale to the thermodynamic scale: 0 °C, 20 °C, 120 °C, 500 °C, −23 °C, −200 °C.

b Convert each of the following temperatures from the thermodynamic scale to the Celsius scale: 0 K, 20 K, 100 K, 300 K, 373 K, 500 K.

5 The electrical resistance of a pure copper wire is mostly due to the vibrations of the copper atoms. Table 21.1 shows how the resistance of a length of copper wire is found to change as it is heated. Copy the table and add a column showing the temperatures in K. Draw a graph to show these data. (Start the temperature scale of your graph at 0 K.) Explain why you might expect the resistance of copper to be zero at this temperature.

Temperature / °C	Resistance / Ω
10	3120
50	3600
75	3900
100	4200
150	4800
220	5640
260	6120

Table 21.1 The variation of resistance with temperature for a length of copper wire.

In Chapter 11, we saw that electrical resistance changes with temperature. For metals, resistance increases with temperature at a fairly steady rate. However, for a thermistor, the resistance changes rapidly over a relatively narrow range of temperatures. A small change in temperature results in a large change in resistance, so a thermometer based on a thermistor will be sensitive over that range of temperatures.

A **thermocouple** is another electrical device which can be used as the sensor of a thermometer. Figure 21.12 shows the principle. Wires of two different metals, X and Y, are required. A length of metal X has a length of metal Y soldered to it at each end. This produces two **junctions**, which are the important parts of the thermocouple. If the two junctions are at different temperatures, an e.m.f. will be produced between the two free ends of the thermocouple, and can be measured using a voltmeter. The greater the difference in temperatures, the greater the voltage produced; however, this e.m.f. may not vary linearly with temperature, i.e. a graph of e.m.f. against temperature is not usually a straight line.

Electrical thermometers can measure across a great **range** of temperatures, from 0 K to hundreds or even thousands of kelvin.

Table 21.2 compares resistance and thermocouple thermometers.

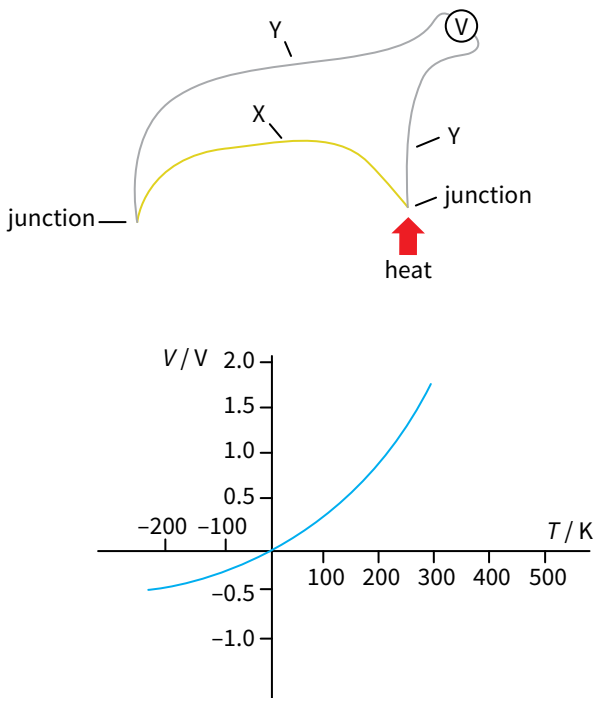


Figure 21.12 The construction of a thermocouple thermometer; the voltage produced depends on the temperature (as shown in the calibration graph) and on the metals chosen.

Feature	Resistance thermometer	Thermocouple thermometer
robustness	very robust	robust
range	thermistor: narrow range resistance wire: wide range	can be very wide
size	larger than thermocouple, has greater thermal capacity therefore slower acting	smaller than resistance thermometers, has smaller thermal capacity, so quicker acting and can measure temperature at a point
sensitivity	thermistor: high sensitivity over narrow range resistance wire: less sensitive	can be sensitive if appropriate metals chosen
linearity	thermistor: fairly linear over narrow range resistance wire: good linearity	non-linear so requires calibration
remote operation	long conducting wires allow the operator to be at a distance from the thermometer	long conducting wires allow the operator to be at a distance from the thermometer

Table 21.2 Comparing resistance and thermocouple thermometers.

QUESTION

6 Give **one** word for each of the following:

- **a** adding a scale to a thermometer
- **b** all the temperatures, from lowest to highest, which a thermometer can measure
- **c** the extent to which equal rises in temperature give equal changes in the thermometer's output
- **d** how big a change in output is produced by a given change in temperature.

Calculating energy changes

So far, we have considered the effects of heating a substance in qualitative terms, and we have given an explanation in terms of a kinetic model of matter. Now we will look at the amount of energy needed to change the temperature of something, and to produce a change of state.

Specific heat capacity

If we heat some material so that its temperature rises, the amount of energy we must supply depends on three things:

- the mass m of the material we are heating
- the temperature change $\Delta\theta$ we wish to achieve
- the material itself.

Some materials are easier to heat than others. It takes more energy to raise the temperature of 1 kg of water by 1 °C than to raise the temperature of 1 kg of alcohol by the same amount.

We can represent this in an equation. The amount of energy E that must be supplied is given by:

$$E = mc\Delta\theta$$

where c is the **specific heat capacity** of the material. Rearranging this equation gives:

$$c = \frac{E}{m\Delta\theta}$$

The specific heat capacity of a material can be defined as a word equation as follows:

$$\text{specific heat capacity} = \frac{\text{energy supplied}}{\text{mass} \times \text{temperature change}}$$

Alternatively, specific heat capacity can be defined in words as follows:

> The specific heat capacity of a substance is the energy required per unit mass of the substance to raise the temperature by 1 K (or 1 °C).

The word 'specific' here means 'per unit mass', i.e. per kg. From this form of the equation, you should be able to see that the units of c are $\text{J kg}^{-1}\text{ K}^{-1}$ (or $\text{J kg}^{-1}\text{ °C}^{-1}$). Table 21.3 shows some values of specific heat capacity measured at 0 °C.

Specific heat capacity is related to the gradient of the sloping sections of the graph shown earlier in Figure 21.4. The steeper the gradient, the faster the substance heats up, and hence the lower its specific heat capacity must be. Worked example 1 shows how to calculate the specific heat capacity of a substance.

Substance	c / J kg^{-1} K^{-1}
aluminium	880
copper	380
lead	126
glass	500–680
ice	2100
water	4180
seawater	3950
ethanol	2500
mercury	140

Table 21.3 Values of specific heat capacity.

WORKED EXAMPLE

1 When 26 400 J of energy is supplied to a 2.0 kg block of aluminium, its temperature rises from 20 °C to 35 °C. The block is well insulated so that there is no energy loss to the surroundings. Determine the specific heat capacity of aluminium.

Step 1 We are going to use the equation:

$$c = \frac{E}{m\Delta\theta}$$

We need to write down the quantities that we know:

$E = 26\,400$ J $m = 2.0$ kg

$\Delta\theta = (35 - 20)$ °C = 15 °C (or 15 K)

Step 2 Now substitute these values and solve the equation:

$$c = \frac{E}{m\Delta\theta}$$

$$c = \frac{26\,400}{(2.0 \times 15)} = 880 \text{ J kg}^{-1}\text{K}^{-1}$$

QUESTIONS

You will need to use data from Table 21.3 to answer these questions.

7 Calculate the energy which must be supplied to raise the temperature of 5.0 kg of water from 20 °C to 100 °C.

8 Which requires more energy, heating a 2.0 kg block of lead by 30 K, or heating a 4.0 kg block of copper by 5.0 K?

9 A well-insulated 1.2 kg block of iron is heated using a 50 W heater for 4.0 min. The temperature of the block rises from 22 °C to 45 °C. Find the experimental value for the specific heat capacity of iron.

BOX 21.1: Determining specific heat capacity *c*

How can we determine the specific heat capacity of a material? The principle is simple: supply a known amount of energy to a known mass of the material and measure the rise in its temperature. Figure 21.13 shows one practical way of doing this for a metal.

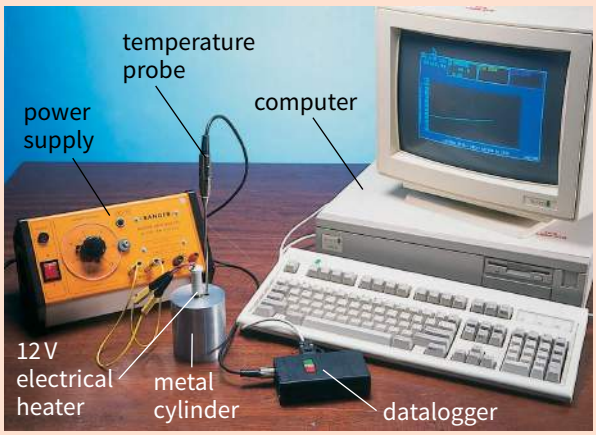


Figure 21.13 A practical arrangement for determining the specific heat capacity of a metal.

The metal is in the form of a cylindrical block of mass 1.00 kg. An electrical heater is used to supply the energy. This type of heater is used because we can easily determine the amount of energy supplied – more easily than if we heated the metal with a Bunsen flame, for example. An ammeter and voltmeter are used to make the necessary measurements.

A thermometer or temperature sensor is used to monitor the block's temperature as it is heated. The block must not be heated too quickly; we want to be sure that the energy has time to spread throughout the metal.

The block should be insulated by wrapping it in a suitable material – this is not shown in the illustration. It would be possible in principle to determine *c* by making just one measurement of temperature change, but it is better to record values of the temperature as it rises and plot a graph of temperature θ against time *t*. The method of calculating *c* is illustrated in Worked example 2.

Sources of error

This experiment can give reasonably good measurements of specific heat capacities. As noted earlier, it is desirable to have a relatively low rate of heating, so that energy spreads throughout the block. If the block is heated rapidly, different parts may be at different temperatures.

Thermal insulation of the material is also vital. Inevitably, some energy will escape to the surroundings. This means that more energy must be supplied to the block for each degree rise in temperature and so the experimental value for the specific heat capacity will be too high. One way around this is to cool the block below room temperature before beginning to heat it. Then, as its temperature rises past room temperature, heat losses will be zero in principle, because there is no temperature difference between the block and its surroundings.

WORKED EXAMPLE

2 An experiment to determine the specific heat capacity *c* of a 1.00 kg aluminium block is carried out; the block is heated using an electrical heater. The current in the heater is 4.17 A and the p.d. across it is 12 V. Measurements of the rising temperature of the block are represented by the graph shown in Figure 21.14. Determine a value for the specific heat capacity *c* of aluminium.

Step 1 Write down the equation that relates energy change to specific heat capacity:

$E = mc\Delta\theta$

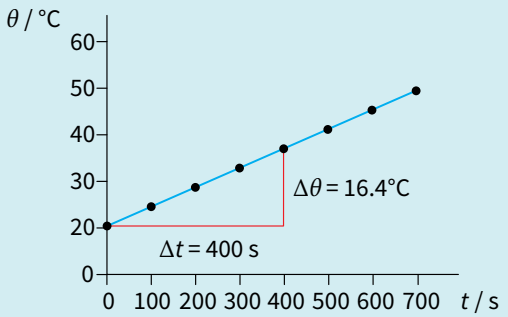


Figure 21.14 Graph of temperature against time for an aluminium block as it is heated.

WORKED EXAMPLE (continued)

Step 2 Divide both sides by a time interval Δt:

$$\frac{E}{\Delta t} = mc\left(\frac{\Delta\theta}{\Delta t}\right)$$

The quantity $\frac{E}{\Delta t}$ is the rate at which energy is supplied, i.e. the power P of the heater. The quantity $\frac{\Delta\theta}{\Delta t}$ is the rate of rise of temperature of the block, i.e. the gradient of the graph of θ against t. Hence:

$P = m \times c \times$ gradient

Step 3 Calculate the power of the heater and the gradient of the graph.

power = p.d. × current

$P = VI = 12 \times 4.17 \approx 50\,\text{W}$

$$\text{gradient} = \frac{\Delta\theta}{\Delta t} = \frac{16.4}{400} = 0.041\,^\circ\text{C}\,\text{s}^{-1}$$

Step 4 Substitute values, rearrange and solve.

$50 = 1.00 \times c \times 0.041$

$$c = \frac{50}{1.00 \times 0.041} = 1220\,\text{J}\,\text{kg}^{-1}\,\text{K}^{-1}$$

QUESTIONS

10 At higher temperature, the graph shown in Figure 21.14 deviates increasingly from a straight line. Suggest an explanation for this.

11 In measurements of the specific heat capacity of a metal, energy losses to the surroundings are a source of error. Is this a systematic error or a random error? Justify your answer.

12 In an experiment to measure the specific heat capacity of water, a student used an electrical heater to heat some water. His results are shown below. Calculate a value for the heat capacity of water. Comment on any likely sources of error.

mass of beaker	=	150 g
mass of beaker + water	=	672 g
current in the heater	=	3.9 A
p.d. across the heater	=	11.4 V
initial temperature	=	18.5 °C
final temperature	=	30.2 °C
time taken	=	13.0 min

QUESTION

13 A block of paraffin wax was heated gently, at a steady rate. Heating was continued after the wax had completely melted. The graph of Figure 21.15 shows how the material's temperature varied during the experiment.

a For each section of the graph (AB, BC and CD), describe the state of the material.

b For each section, explain whether the material's internal energy was increasing, decreasing or remaining constant.

c Consider the two sloping sections of the graph. State whether the material's specific heat capacity is greater when it is a solid or when it is a liquid. Justify your answer.

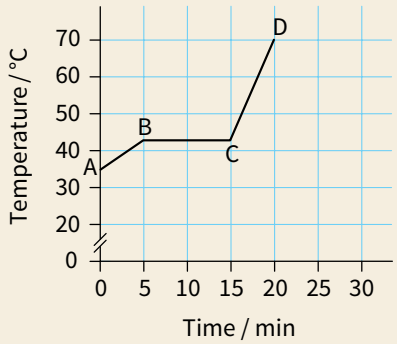


Figure 21.15 Temperature variation of a sample of wax, heated at a constant rate.

Specific latent heat

Energy must be supplied to melt or boil a substance. (In this case, there is no temperature rise to consider since the temperature stays constant during a change of state.) This energy is called latent heat.

The specific latent heat of a substance is the energy required per kilogram of the substance to change its state without any change in temperature.

When a substance melts, this quantity is called the **specific latent heat of fusion**; for boiling, it is the **specific latent heat of vaporisation**.

To calculate the amount of energy E required to melt or vaporise a mass m of a substance, we simply need to know its specific latent heat L:

$$E = mL$$

L is measured in J kg^{-1}. (Note that there is no 'per °C' since there is no change in temperature.) For water the values are:

- specific latent heat of fusion of water, 330 kJ kg^{-1}
- specific latent heat of vaporisation of water, 2.26 MJ kg^{-1}

You can see that L for boiling water to form steam is roughly seven times the value for melting ice to form water. As we saw on page 329, this is because, when ice melts, only one or two bonds are broken for each molecule; when water boils, several bonds are broken per molecule. Worked example 3 shows how to calculate these amounts of energy.

WORKED EXAMPLE

3 The specific latent heat of vaporisation of water is 2.26 MJ kg^{-1}. Calculate the energy needed to change 2.0 g of water into steam at 100 °C.

Step 1 We have been given the following quantities:

$m = 2.0\,\text{g} = 0.002\,\text{kg}$ and $L = 2.26\,\text{MJ kg}^{-1}$

Step 2 Substituting these values in the equation $E = mL$, we have:

energy $= 0.002 \times 2.26 \times 10^6 = 4520\,\text{J}$

QUESTIONS

14 The specific latent heat of fusion of water is 330 kJ kg^{-1}. Calculate the energy needed to change 2.0 g of ice into water at 0 °C. Suggest why the answer is much smaller than the amount of energy calculated in Worked example 3.

15 A sample of alcohol is heated with a 40 W heater until it boils. As it boils, the mass of the liquid decreases at a rate of 2.25 g per minute. Assuming that 80% of the energy supplied by the heater is transferred to the alcohol, estimate the specific latent heat of vaporisation of the alcohol. Give your answer in kJ kg^{-1}.

BOX 21.2: Determining specific latent heat L

The principle of determining the specific latent heat of a material is similar to determining the specific heat capacity (but remember that there is no change in temperature).

Figure 21.16 shows how to measure the specific latent heat of vaporisation of water. A beaker containing water is heated using an electrical heater. A wattmeter (or an ammeter and a voltmeter) determines the rate at which energy is supplied to the heater. The beaker is insulated to minimise energy loss, and it stands on a balance. A thermometer is included to ensure that the temperature of the water remains at 100 °C.

The water is heated at a steady rate and its mass recorded at equal intervals of time. Its mass decreases as it boils.

A graph of mass against time should be a straight line whose gradient is the rate of mass loss. The wattmeter shows the rate at which energy is supplied to the water via the heater. We thus have:

$$\text{specific latent head} = \frac{\text{rate of supply of energy}}{\text{rate of loss of mass}}$$

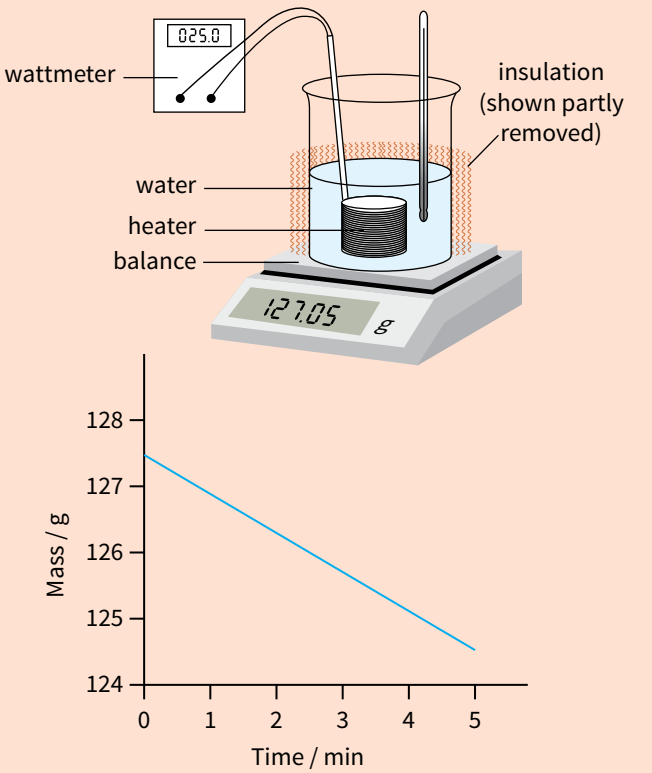


Figure 21.16 Determining the specific latent heat of vaporisation of water.

BOX 21.2: Determining specific latent heat *L* (continued)

A similar approach can be used to determine the specific latent heat of fusion of ice. In this case, the ice is heated electrically in a funnel; water runs out of the funnel and is collected in a beaker on a balance.

As with any experiment, we should consider sources of error in measuring *L* and their effects on the final result. When water is heated to produce steam, some energy may escape to the surroundings so that the measured energy is greater than that supplied to the water. This systematic error gives a value of *L* which is greater than the true value. When ice is melted, energy from the surroundings will conduct into the ice, so that the measured value of *L* will be an underestimate.

Summary

- The kinetic model of matter allows us to explain behaviour (e.g. changes of state) and macroscopic properties (e.g. specific heat capacity and specific latent heat) in terms of the behaviour of molecules.
- The internal energy of a system is the sum of the random distribution of kinetic and potential energies associated with the atoms or molecules that make up the system.
- If the temperature of an object increases, there is an increase in its internal energy.
- Internal energy also increases during a change of state, but there is no change in temperature.
- The first law of thermodynamics expresses the conservation of energy:

  increase in internal energy = energy supplied by heating + work done on the system
- Temperatures on the thermodynamic (Kelvin) and Celsius scales of temperature are related by:

  $T\,(\mathrm{K}) = \theta\,(^\circ\mathrm{C}) + 273.15$

  $\theta\,(^\circ\mathrm{C}) = T\,(\mathrm{K}) - 273.15$
- At absolute zero, all substances have a minimum internal energy.
- A thermometer makes use of a physical property of a material that varies with temperature.
- The word equation for the specific heat capacity of a substance is:

  $$\text{specific heat capacity} = \frac{\text{energy supplied}}{\text{mass} \times \text{temperature change}}$$

  The specific heat capacity of a substance is the energy required per unit mass of the substance to raise the temperature by 1 K (or 1 °C).
- The energy transferred in raising the temperature of a substance is given by $E = mc\Delta\theta$.
- The specific latent heat of a substance is the energy required per kilogram of the substance to change its state without any change in temperature: $E = mL$.

End-of-chapter questions

1 Describe the changes to the kinetic energy, the potential energy and the total internal energy of the molecules of a block of ice as:

a it melts at 0 °C [3]

b the temperature of the water rises from 0 °C to room temperature. [3]

2 Explain, in terms of kinetic energy, why the temperature of a stone increases when it falls from a cliff and lands on the beach below. [3]

3 Explain why the barrel of a bicycle pump gets very hot as it is used to pump up a bicycle tyre. (Hint: the work done against friction is not large enough to explain the rise in temperature.) [3]

4 The so-called 'zeroth law of thermodynamics' states that if the temperature of body A is equal to the temperature of body B and the temperature of body B is the same as body C, then the temperature of body C equals the temperature of body A.

Explain, in terms of energy flow, why the concept of temperature would be meaningless if this law was not obeyed. [2]

5 Copy and complete the table, which shows the melting and boiling points (at standard atmospheric pressure) of different materials in both degrees Celsius and kelvin. [8]

Substance	Melting point		Boiling point	
	°C	K	°C	K
oxygen	−223			90
hydrogen		14	−253	
lead	327			2023
mercury		234	357	

6 When a thermocouple has one junction in melting ice and the other junction in boiling water it produces an e.m.f. of 63 μV.

a What e.m.f. would be produced if the second junction was also placed in melting ice? [1]

b When the second junction is placed in a cup of coffee, the e.m.f. produced is 49 μV. Calculate the temperature of the coffee. [2]

c The second junction is now placed in a beaker of melting lead at 327 °C.

i Calculate the e.m.f. which would be produced. [2]

ii State the assumption you make. [1]

7 The following list gives four different types of thermometer:

thermistor

thermocouple

constant pressure gas thermometer

liquid-in-glass thermometer.

For each of the jobs below, state which type of thermometer you would use and justify your choice.

a A gardener measuring the temperature in a greenhouse [2]

b An engineer mapping the temperature at different points on the cylinder head of a car engine [2]

c A technician monitoring the temperatures in the core of a nuclear reactor. [2]

8 **a** A 500 W kettle contains 300 g of water at 20 °C. Calculate the minimum time it would take to raise the temperature of the water to boiling point. [5]

b The kettle is allowed to boil for 2 minutes. Calculate the mass of water that remains in the kettle. State any assumptions that you make. [4]

(Specific heat capacity of water = 4.18×10^3 J kg^{-1} °C^{-1}, specific latent heat of vaporisation of water = 2.26×10^6 J kg^{-1}.)

9 **a** Explain what is meant by the **specific heat capacity** of a substance. [2]

b Figure 21.17 shows an apparatus used to measure the specific heat capacity of a copper block of mass 850 g.

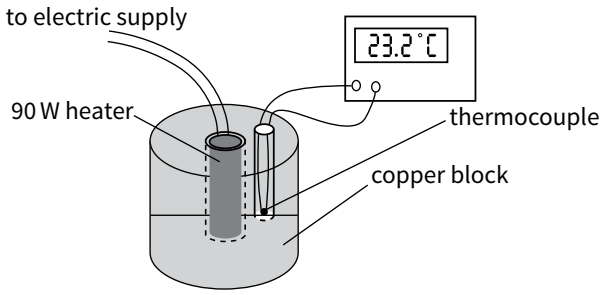


Figure 21.17 For End-of-chapter Question 9.

The block is initially at 12 °C. The heater is switched on and the time taken for the temperature to rise to 30 °C is recorded. The block is cooled to the original temperature, the potential difference across the heater is increased and the experiment is repeated.

The results are shown in the table.

Power output of heater / W	Time taken / s
40	190
60	114

i Explain why a thermocouple thermometer is suitable for this experiment. [2]

ii Explain why the experiment is repeated using a different power input. [2]

iii Calculate the specific heat capacity of copper. [5]

10 **a** A cylinder of carbon dioxide is at room temperature and at a pressure of 20 atmospheres. A cloth is placed over the outlet and the tap opened. Solid carbon dioxide is formed on the cloth. Explain, using the first law of thermodynamics, why the carbon dioxide cools sufficiently for the solid to form. [3]

b Solid carbon dioxide sublimes to form carbon dioxide gas; that is, it changes directly from a solid to a gas. This change is called **sublimation**. Figure 21.18 shows the apparatus used to measure the specific latent heat of sublimation of carbon dioxide.

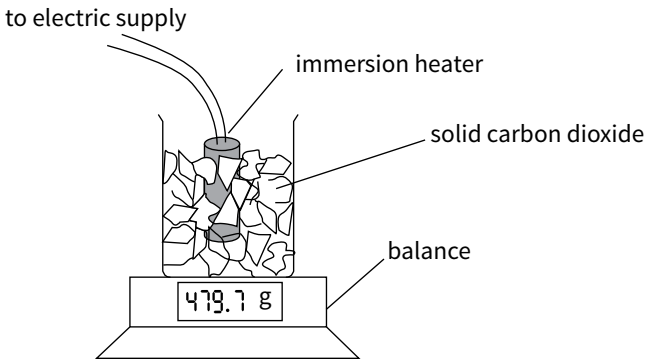


Figure 21.18 For End-of-chapter Question 10.

Explain what is meant by the **latent heat of sublimation**. [2]

c The change in mass of the solid carbon dioxide in a time interval of 5 minutes is measured with the heater switched off. The change of mass, in the same time interval, is found with the heater switched on. The energy to supplied to the heater is also measured.

The results are shown in the table below.

	Initial reading on balance / g	**Final reading on balance / g**	**Energy supplied to the heater / J**
heater off	484.3	479.7	0
heater on	479.7	454.2	12 000

i Explain why readings are taken with the heater switched off. [1]

ii Calculate the specific latent heat of sublimation of carbon dioxide. [3]

d The specific latent heat of sublimation is greater than either the specific latent heat of fusion or the specific latent heat of vaporisation.

Suggest a reason for this. [2]

11 a Explain why energy is needed for boiling even though the temperature of the liquid remains constant. [2]

Figure 21.19 shows an apparatus that can be used to measure the specific latent heat of vaporisation of nitrogen.

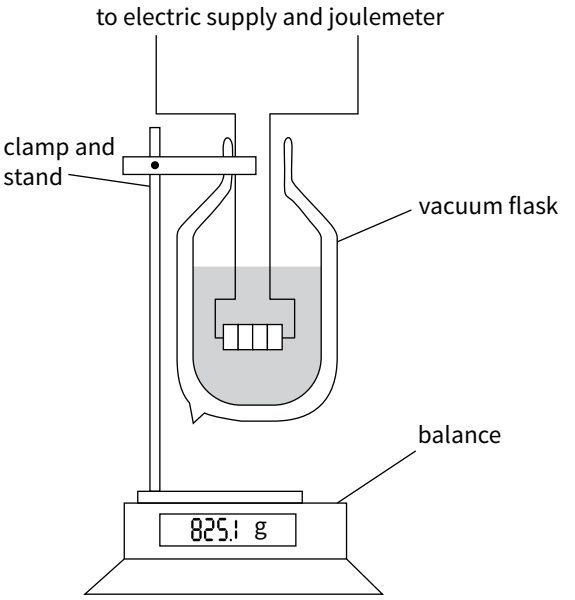


Figure 21.19 For End-of-chapter Question 11.

b Suggest why the nitrogen is contained in a vacuum flask. [1]

c The change in mass of the nitrogen is measured over a specific time interval with the heater switched off. The heater is switched on, transferring energy at 40 W, and the change of mass is found once more.

The results are shown in the table.

	Initial reading on balance / g	**Final reading on balance / g**	**Time / minutes**
heater off	834.7	825.5	4
heater on	825.5	797.1	2

Calculate the specific latent heat of vaporisation of liquid nitrogen. [4]

12 a i Explain what is meant by **internal energy**. [2]

ii Explain what is meant by the **absolute zero of temperature**. [2]

b An electric hot water heater has a power rating of 9.0 kW. The water is heated as it passes through the heater. Water flows through the heater at a speed of 1.2 m s^{-1} through pipes which have a total cross-sectional area of 4.8×10^{-5} m^2. The temperature of the water entering the heater is 15 °C.

i Calculate the mass of water flowing through the heater each second. [2]

ii Calculate the temperature at which the water leaves the heater. [3]

iii State any assumptions you have made when doing your calculation. [1]

iv It is possible to adjust the temperature of the water from the heater. Suggest how the temperature of the water could be increased. [1]

(Density of water = 1000 kg m^{-3}, specific heat capacity of water = 4200 J kg^{-1} $°C^{-1}$.)



Chapter 22: Ideal gases

Learning outcomes

You should be able to:

- solve problems using the equation of state for an ideal gas
- interpret Brownian motion in terms of the movement of molecules
- state the assumptions of the kinetic theory of gases
- deduce a relationship between pressure, volume and the microscopic properties of the molecules of a gas
- relate the kinetic energy of the molecules of a gas to its temperature

The idea of a gas

Figure 22.1 shows a weather balloon being launched. Balloons like the one in Figure 22.1 carry instruments high into the atmosphere, to measure pressure, temperature, wind speed and other variables .

The balloon is filled with helium so that its overall density is less than that of the surrounding air. The result is an upthrust on the balloon, greater than its weight, so that it rises upwards. As it moves upwards, the pressure of the surrounding atmosphere decreases so that the balloon expands. The temperature drops, which tends to make the gas in the balloon shrink. In this chapter we will look at the behaviour of gases as their pressure, temperature and volume change.



Figure 22.1 A weather balloon being launched.

Particles of a gas

We picture the particles of a gas as being fast-moving. They bounce off the walls of their container (and off each other) as they travel around at high speed (see Figure 22.2). How do we know that these particles are moving like this?

It is much harder to visualise the particles of a gas than those of a solid, because they move about in such a disordered way, and most of a gas is empty space. The movement of gas particles was investigated in the 1820s by a Scottish botanist, Robert Brown. He was using a microscope to look at pollen grains suspended in water, and saw very small particles moving around inside the water. He then saw the same motion in particles of dust in the air. It is easier in the laboratory to look at the movement of tiny particles of smoke in air.

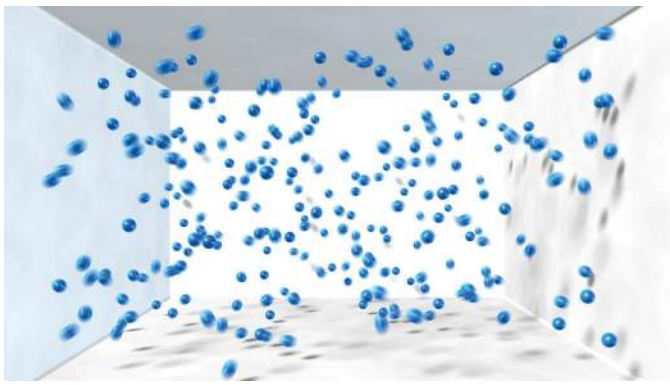


Figure 22.2 Particles of a gas – collisions with the walls of the container cause the gas's pressure on the container. (Particles do not have shadows like this. The shadows are added here to show depth.)

Fast molecules

For air at standard temperature and pressure (STP, −0 °C and 100 kPa), the average speed of the molecules is about 400 m s^{-1}. At any moment, some are moving faster than this and others more slowly. If we could follow the movement of a single air molecule, we would find that, some of the time, its speed was greater than this average; at other times it would be less. The velocity (magnitude and direction) of an individual molecule changes every time it collides with anything else.

This value for molecular speed is reasonable. It is comparable to (but greater than) the speed of sound in air (approximately 330 m s^{-1} at STP). Very fast-moving particles can easily escape from the Earth's gravitational field. The required escape velocity is about 11 km s^{-1}. Since we still have an atmosphere, on average the air molecules must be moving much more slowly than this value.

Path of a particle

The erratic motion of particles in water that Brown observed comes about because the particles are constantly bombarded by the much smaller, faster water molecules. This motion came to be known as **Brownian motion**, and it can be observed in both liquids and gases.

Figure 22.3 shows the sort of path followed by a particle showing Brownian motion. In fact, this is from a scientific paper by the French physicist Jean Perrin, published in 1911. He was looking at the movement of a single pollen grain suspended in water.

He recorded its position every 30 s; the grid spacing is approximately 3 μm. From this he could deduce the average speed of the grain and hence work out details of the movement of water molecules.

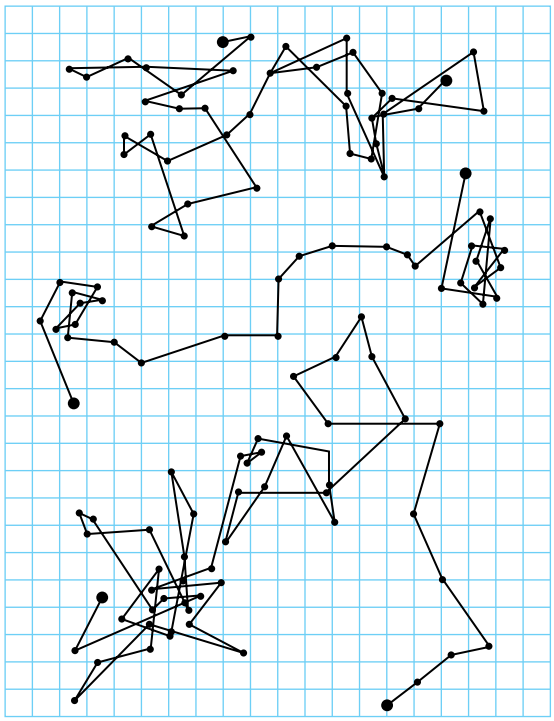


Figure 22.3 Brownian motion of pollen grains, as drawn by the French scientist Jean Perrin.

QUESTIONS

1. Consider a smoke particle of mass M and speed V. It is constantly buffeted by air molecules. The mass of a single air molecule is m and it has speed v. It is reasonable to assume that, on average, the smoke particle will have kinetic energy approximately equal to the kinetic energy of a single air molecule. Show that, since $M >> m$ (M is much greater than m), it follows that the air molecules must be moving much faster than the smoke grain ($v >> V$).
2. If equal numbers of air molecules hit a smoke particle from all directions, the smoke particle does not move. State three ways in which the random movement of molecules of air cause the smoke particles to move.
3. Describe and explain what you would expect to see through a microscope observing Brownian motion when the temperature increases.
4. An oxygen molecule is moving around inside a spherical container of diameter 0.10 m. The molecule's speed is 400 m s^{-1}. Estimate the number of times each second the molecule collides with the walls of the container. (You can assume that the molecule's speed is constant.)

BOX 22.1: Observing Brownian motion

The oxygen and nitrogen molecules that make up most of the air are far too small to see; they are much smaller than the wavelength of light. To observe the effect of the air molecules we have to look at something bigger. In this experiment (Figure 22.4), the smoke cell contains air into which a small amount of smoke has been introduced. The cell is lit from the side, and the microscope is used to view the smoke particles.

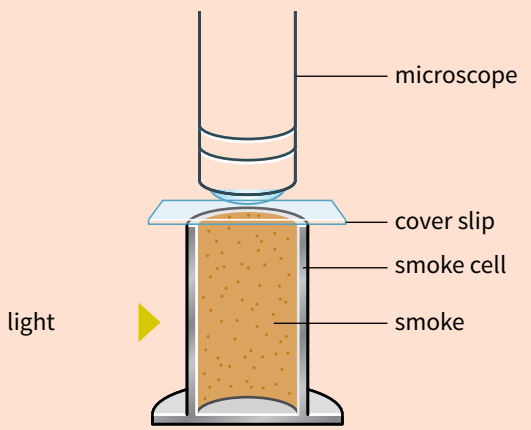


Figure 22.4 Experimental arrangement for observing Brownian motion.

The smoke particles show up as tiny specks of reflected light, although they are too small for any detail of their shape to be seen. What is noticeable is the way they move. If you can concentrate on a single particle, you will see that it follows a somewhat jerky and erratic path. This is a consequence of the repeated collisions between the smoke particles and air molecules. The erratic motion of the smoke particles provides direct evidence that the air molecules must:

- be moving
- also have haphazard motion.

Since the air molecules are much smaller than the smoke grain, we can deduce that they must be moving much faster than the smoke grain if they are to affect it in this way.

(You may observe that all of the smoke particles in your field of view have a tendency to travel in one particular direction. This is a consequence of convection currents in the air. Also, you may have to adjust the focus of the microscope to keep track of an individual particle, as it moves up or down in the cell.)

Explaining pressure

A gas exerts pressure on any surface with which it comes into contact. Pressure is a macroscopic property, defined as the force exerted per unit area of the surface.

The pressure of the atmosphere at sea level is approximately 100 000 Pa. The surface area of a typical person is 2.0 m^2. Hence the force exerted on a person by the atmosphere is about 200 000 N. This is equivalent to the weight of about 200 000 apples!

Fortunately, air inside the body presses outwards with an equal and opposite force, so we do not collapse under the influence of this large force. We can explain the macroscopic phenomenon of pressure by thinking about the behaviour of the microscopic particles that make up the atmosphere.

Figure 22.5 shows the movement of a single molecule of air in a box. It bounces around inside, colliding with the various surfaces of the box. At each collision, it exerts a small force on the box. The pressure on the inside of the box is a result of the forces exerted by the vast number of molecules in the box. Two factors affect the force, and hence the pressure, that the gas exerts on the box:

- the number of molecules that hit each side of the box in one second
- the force with which a molecule collides with the wall.

If a molecule of mass m hits the wall head-on with a speed v it will rebound with a speed v in the opposite direction. The change in momentum of the molecule is $2mv$. Since force is equal to rate of change of momentum, the higher the speed of the molecule the greater the force that it exerts as it collides with the wall. Hence the pressure on the wall will increase if the molecules move faster.

If the piston in a bicycle pump is pushed inwards but the temperature of the gas inside is kept constant, then more molecules will hit the piston in each second but each collision will produce the same force, because the temperature and therefore the average speed of the molecules is the same. The increased rate of collisions alone means that the force on the piston increases and thus the pressure rises. If the temperature of the gas in a container rises then the molecules move faster and hit the sides faster and more often; both of these factors cause the pressure to rise.

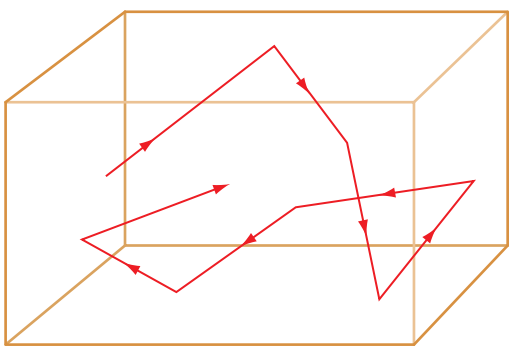


Figure 22.5 The path of a single molecule in an otherwise empty box.

QUESTIONS

5 State and explain in terms of the kinetic model what happens to the pressure inside a tyre when more molecules at the same temperature are pumped into the tyre.

6 Explain, using the kinetic model, why a can containing air may explode if the temperature rises.

Measuring gases

We are going to picture a container of gas, such as the box shown in Figure 22.6. There are four properties of this gas that we might measure: pressure, temperature, volume and mass. In this chapter, you will learn how these quantities are related to one another.

Pressure

This is the normal force exerted per unit area by the gas on the walls of the container. We saw in Chapter 6 that this pressure is the result of molecular collisions with the walls of the container. Pressure is measured in pascals, Pa ($1\,Pa = 1\,N\,m^{-2}$).

Temperature

This might be measured in °C, but in practice it is more useful to use the thermodynamic (Kelvin) scale of temperature. You should recall how these two scales are related:

$$T(K) = \theta(°C) + 273.15$$

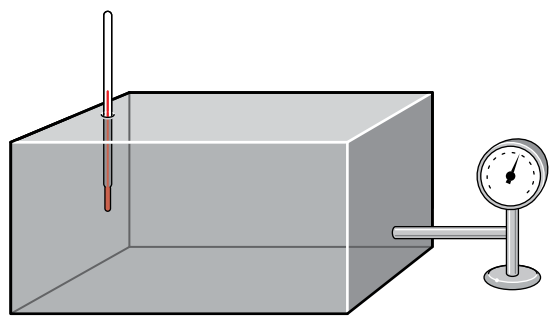


Figure 22.6 A gas has four measurable properties, which are all related to one another: pressure, temperature, volume and mass.

Volume

This is a measure of the space occupied by the gas. Volume is measured in m^3.

Mass

This is measured in g or kg. In practice, it is more useful to consider the **amount** of gas, measured in moles.

The mole is defined as follows:

> One mole of any substance is the amount of that substance which contains the same number of particles as there are in 0.012 kg of carbon-12.

(In this definition, 'particles' may be atoms, molecules, ions, etc.)

One mole of any substance has a mass in grams which is numerically equal to the relative atomic or molecular mass of the substance. For example, one mole of oxygen (O_2) has a mass of about 32 g.

A mole of any substance (solid, liquid or gas) contains a standard number of particles (molecules or atoms). This number is known as the **Avogadro constant**, N_A. The experimental value for N_A is $6.02 \times 10^{23}\,mol^{-1}$. We can easily determine the number of atoms in a sample if we know how many moles it contains. For example:

2.0 mol of helium contains

$2.0 \times 6.02 \times 10^{23} = 1.20 \times 10^{24}$ atoms

10 mol of carbon contains

$10 \times 6.02 \times 10^{23} = 6.02 \times 10^{24}$ atoms

We will see later that, if we consider equal numbers of moles of two different gases under the same conditions, their physical properties are the same.

QUESTIONS

7 The mass of one mole of carbon is 12 g. Determine:
- **a** the number of atoms in one mole of carbon
- **b** the number of moles and the number of atoms in 54 g of carbon
- **c** the number of atoms in 1.0 kg of carbon.

8 The molar mass of uranium is about $235\,g\,mol^{-1}$.
- **a** Calculate the mass of a single atom of uranium.
- **b** A small pellet of uranium has a mass of 20 mg. For this pellet, calculate:
  - **i** the number of moles
  - **ii** the number of uranium atoms.

QUESTION

9 'It can be useful to recall that 1.0 kg of matter contains of the order of 10^{26} atoms.' Making suitable estimates, test this statement.

Boyle's law

This law relates the pressure p and volume V of a gas. It was discovered in 1662 by Robert Boyle. If a gas is compressed, its pressure increases and its volume decreases. Pressure and volume are inversely related.

We can write **Boyle's law** as:

> The pressure exerted by a fixed mass of gas is inversely proportional to its volume, provided the temperature of the gas remains constant.

Note that this law relates two variables, pressure and volume, and it requires that the other two, mass and temperature, remain constant.

Boyle's law can be written as:

$$p \propto \frac{1}{V}$$

or simply:

$$pV = \text{constant}$$

We can also represent Boyle's law as a graph, as shown in Figure 22.7. A graph of p against $\frac{1}{V}$ is a straight line passing through the origin, showing direct proportionality. For solving problems, you may find it more useful to use the equation in this form:

$$p_1V_1 = p_2V_2$$

Here, p_1 and V_1 represent the pressure and volume of the gas before a change, and p_2 and V_2 represent the pressure and volume of the gas after the change. Worked example 1 shows how to use this equation.

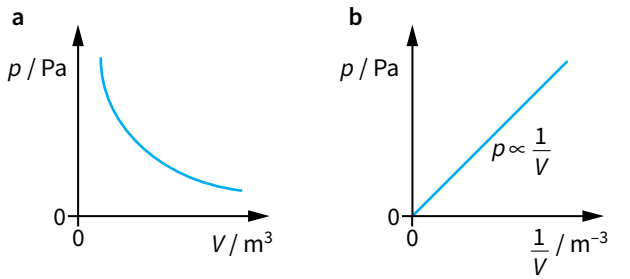


Figure 22.7 Graphical representations of the relationship between pressure and volume of a gas (Boyle's law).

WORKED EXAMPLE

1 A cylinder contains 0.80 dm^3 of nitrogen gas at a pressure of 1.2 atmosphere (1 atm = 1.01×10^5 Pa). A piston slowly compresses the gas to a pressure of 6.0 atm. The temperature of the gas remains constant. Calculate the final volume of the gas.

Note from the question that the temperature of the gas is constant, and that its mass is fixed (because it is contained in a cylinder). This means that we can apply Boyle's law.

Step 1 We are going to use Boyle's law in the form $p_1V_1 = p_2V_2$. Write down the quantities that you know, and that you want to find out.

$p_1 = 1.2$ atm $\quad V_1 = 0.80\ dm^3$

$p_2 = 6.0$ atm $\quad V_2 = ?$

Note that we don't need to worry about the particular units of pressure and volume being used here, so long as they are the same on both sides of the equation. The final value of V_2 will be in dm^3 because V_1 is in dm^3.

Step 2 Substitute the values in the equation, rearrange and find V_2.

$$p_1V_1 = p_2V_2$$

$$1.2 \times 0.8 = 6.0 \times V_2$$

$$V_2 = \frac{1.2 \times 0.8}{6.0} = 0.16\ dm^3$$

So the volume of the gas is reduced to 0.16 dm^3.

The pressure increases by a factor of 5, so the volume decreases by a factor of 5.

QUESTION

10 A balloon contains 0.04 m^3 of air at a pressure of 120 kPa. Calculate the pressure required to reduce its volume to 0.025 m^3 at constant temperature.

Changing temperature

Boyle's law requires that the temperature of a gas is fixed. What happens if the temperature of the gas is allowed to change? Figure 22.8 shows the results of an experiment in which a fixed mass of gas is cooled at constant pressure. The gas contracts; its volume decreases.

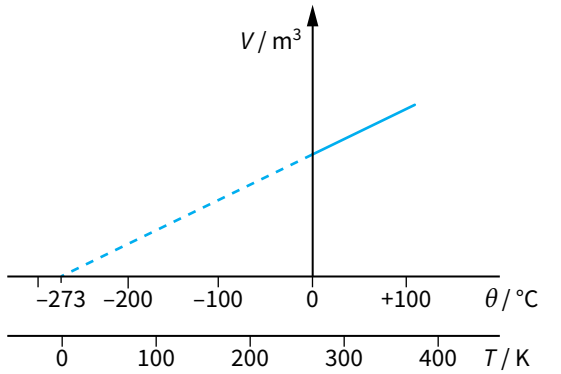


Figure 22.8 The volume of a gas decreases as its temperature decreases.

This graph does not show that the volume of a gas is proportional to its temperature on the Celsius scale. If a gas contracted to zero volume at 0 °C, the atmosphere would condense on a cold day and we would have a great deal of difficulty in breathing! However, the graph **does** show that there is a temperature at which the volume of a gas does, in principle, shrink to zero. Looking at the lower temperature scale on the graph, where temperatures are shown in kelvin (K), we can see that this temperature is 0 K, or absolute zero. (Historically, this is how the idea of absolute zero first arose.)

We can represent the relationship between volume V and thermodynamic temperature T as:

$$V \propto T$$

or simply:

$$\frac{V}{T} = \text{constant}$$

Note that this relationship only applies to a fixed mass of gas and to constant pressure.

The relationship above is an expression of **Charles's law**, named after the French physicist Jacques Charles, who in 1787 experimented with different gases kept at constant pressure.

If we combine Boyle's law and Charles's law, we can arrive at a single equation for a fixed mass of gas:

$$\frac{pV}{T} = \text{constant}$$

Shortly, we will look at the constant quantity which appears in this equation, but first we will consider the extent to which this equation applies to real gases.

Real and ideal gases

The relationships between p, V and T that we have considered above are based on experimental observations of gases such as air, helium, nitrogen, etc., at temperatures

and pressures around room temperature and pressure. In practice, if we change to more extreme conditions, such as low temperatures or high pressures, gases start to deviate from these laws as the gas atoms exert significant electrical forces on each other. For example, Figure 22.9 shows what happens when nitrogen is cooled down towards absolute zero. At first, the graph of volume against temperature follows a good straight line. However, as it approaches the temperature at which it condenses, it deviates from ideal behaviour, and at 77 K it condenses to become liquid nitrogen.

Thus we have to attach a condition to the relationships discussed above. We say that they apply to an **ideal gas**. When we are dealing with real gases, we have to be aware that their behaviour may be significantly different from the ideal equation:

$$\frac{pV}{T} = \text{constant}$$

An ideal gas is thus defined as one for which we can apply the equation:

$$\frac{pV}{T} = \text{constant}$$

for a fixed mass of gas.

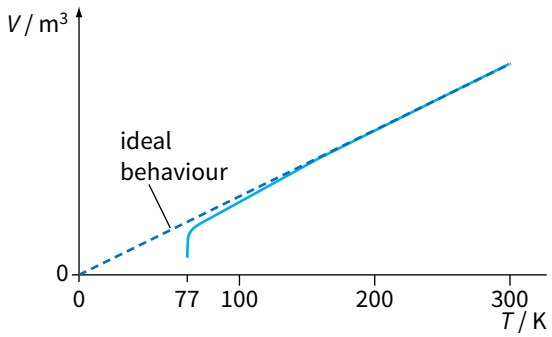


Figure 22.9 A real gas (in this case, nitrogen) deviates from the behaviour predicted by Charles's law at low temperatures.

Ideal gas equation

So far, we have seen how p, V and T are related. It is possible to write a single equation relating these quantities which takes into account the amount of gas being considered.

If we consider n moles of an ideal gas, we can write the equation in the following form:

$$pV = nRT$$

This equation is called the **ideal gas equation** or the **equation of state** for an ideal gas. It relates all four of the variable quantities discussed at the beginning of this chapter. The constant of proportionality R is called the universal molar gas constant. Its experimental value is:

$$R = 8.31\ \text{J mol}^{-1}\ \text{K}^{-1}$$

Note that it doesn't matter what gas we are considering – it could be a very 'light' gas like hydrogen, or a much 'heavier' one like carbon dioxide. So long as it is behaving as an ideal gas, we can use the same equation of state with the same constant R.

Calculating the number *n* of moles

Sometimes we know the mass of gas we are concerned with, and then we must find how many moles this represents. To do this, we use the relationship:

$$\text{number of moles} = \frac{\text{mass (g)}}{\text{molar mass (g mol}^{-1})}$$

For example: How many moles are there in 1.6 kg of oxygen?

$$\text{molar mass of oxygen} = 32\ \text{g mol}^{-1}$$

$$\text{number of moles} = \frac{1600\ \text{g}}{32\ \text{g mol}^{-1}} = 50\ \text{mol}$$

(Note that this tells us that there are 50 moles of oxygen **molecules** in 1.6 kg of oxygen. An oxygen molecule consists of two oxygen atoms – its formula is O_2 – so 1.6 kg of oxygen contains 100 moles of oxygen **atoms**.)

Now look at Worked examples 2 and 3.

WORKED EXAMPLE

2 Calculate the volume occupied by one mole of an ideal gas at room temperature (20 °C) and pressure (1.013×10^5 Pa).

Step 1 Write down the quantities given.

$p = 1.013 \times 10^5$ Pa $n = 1.0$

$T = 293$ K

Hint: *Note that the temperature is converted to kelvin.*

Step 2 Substituting these values in the equation of state gives:

$$V = \frac{nRT}{P} = \frac{1 \times 8.31 \times 293}{1.103 \times 10^5}$$

$V = 0.0240\ \text{m}^3 = 2.40 \times 10^{-2}\ \text{m}^3$

$= 24.0\ \text{dm}^3$

Hint: *1 dm = 0.1 m; hence 1 dm³ = 10⁻³ m³.*

This value, the volume of one mole of gas at room temperature and pressure, is well worth remembering. It is certainly known by most chemists.

WORKED EXAMPLE

3 A car tyre contains 0.020 m^3 of air at 27 °C and at a pressure of 3.0×10^5 Pa. Calculate the mass of the air in the tyre. (Molar mass of air = 28.8 g mol^{-1}.)

Step 1 Here, we need first to calculate the number of moles of air using the equation of state. We have:

$p = 3.0 \times 10^5$ Pa $\quad V = 0.02\ m^3$

$T = 27\ °C = 300\ K$

Hint: *Don't forget to convert the temperature to kelvin.*

So, from the equation of state:

$$n = \frac{pV}{RT} = \frac{30 \times 10^5 \times 0.02}{8.31 \times 300}$$

$n = 2.41$ mol

Step 2 Now we can calculate the mass of air:

mass = number of moles × molar mass

mass = 2.41 × 28.8 = 69.4 g ≈ 69 g

QUESTIONS

For the questions which follow, you will need the following value:

$R = 8.31\ J\ mol^{-1}\ K^{-1}$

11 At what temperature (in K) will 1.0 mol of a gas occupy 1.0 m^3 at a pressure of 1.0×10^4 Pa?

12 Nitrogen consists of molecules N_2. The molar mass of nitrogen is 28 g mol^{-1}. For 100 g of nitrogen, calculate:

a the number of moles

b the volume occupied at room temperature and pressure. (r.t.p. = 20 °C, 1.01×10^5 Pa.)

13 Calculate the volume of 5.0 mol of an ideal gas at a pressure of 1.0×10^5 Pa and a temperature of 200 °C.

14 A sample of gas contains 3.0×10^{24} atoms. Calculate the volume of the gas at a temperature of 300 K and a pressure of 120 kPa.

15 At what temperature would 1.0 kg of oxygen occupy 1.0 m^3 at a pressure of 1.0×10^5 Pa? (Molar mass of O_2 = 32 g mol^{-1}.)

QUESTION

16 A cylinder of hydrogen has a volume of 0.10 m^3. Its pressure is found to be 20 atmospheres at 20 °C.

a Calculate the mass of hydrogen in the cylinder.

b If it were instead filled with oxygen to the same pressure, how much oxygen would it contain?

(Molar mass of H_2 = 2.0 g mol^{-1}, molar mass of O_2 = 32 g mol^{-1}; 1 atmosphere = 1.01×10^5 Pa.)

Modelling gases – the kinetic model

In this chapter, we are concentrating on the macroscopic properties of gases (pressure, volume, temperature). These can all be readily measured in the laboratory. The equation:

$$\frac{pV}{T} = \text{constant}$$

is an empirical relationship. In other words, it has been deduced from the results of experiments. It gives a good description of gases in many different situations. However, an empirical equation does not **explain** why gases behave in this way. An explanation requires us to think about the underlying nature of a gas and how this gives rise to our observations.

A gas is made of particles (atoms or molecules). Its pressure arises from collisions of the particles with the walls of the container; more frequent or harder collisions give rise to greater pressure. Its temperature indicates the average kinetic energy of its particles; the faster they move, the greater their average kinetic energy and the higher the temperature.

The **kinetic theory of gases** is a theory which links these microscopic properties of particles (atoms or molecules) to the macroscopic properties of a gas. Table 22.1 shows the assumptions on which the theory is based.

On the basis of these assumptions, it is possible to use Newtonian mechanics to show that pressure is inversely proportional to volume (Boyle's law), volume is directly proportional to thermodynamic (kelvin) temperature (Charles's law), and so on. The theory also shows that the particles of a gas have a range of speeds – some move faster than others.

Assumption	Explanation/comment
A gas contains a very large number of particles (atoms or molecules).	A small 'cube' of air can have as many as 10^{20} molecules.
The forces between particles are negligible, except during collisions.	If the particles attracted each other strongly over long distances, they would all tend to clump together in the middle of the container. The particles travel in straight lines between collisions.
The volume of the particles is negligible compared to the volume occupied by the gas.	When a liquid boils to become a gas, its particles become much farther apart.
Most of the time, a particle moves in a straight line at a constant velocity. The time of collision with another particle or with the container walls is negligible compared with the time between collisions.	The particles collide with the walls of the container and with each other, but for most of the time they are moving with constant velocity.
The collisions of particles with each other and with the container are perfectly elastic, so that no kinetic energy is lost.	Kinetic energy cannot be lost. The internal energy of the gas is the total kinetic energy of the particles.

Table 22.1 The basic assumptions of the kinetic theory of gases.

Things are different when a gas is close to condensing. At temperatures a little above the boiling point, the molecules of a gas are moving more slowly and they tend to stick together – a liquid is forming. So we cannot consider them to be moving about freely, and the kinetic theory of gases must be modified. This is often how physics progresses. A theory is developed which explains a simple situation. Then the theory is modified to explain more complex situations.

The kinetic theory has proved to be a very powerful model. It convinced many physicists of the existence of particles long before it was ever possible to visualise them.

Molecules in a box

We can use the kinetic model to deduce an equation which relates the macroscopic properties of a gas (pressure, volume) to the microscopic properties of its molecules (mass and speed). We start by picturing a single molecule in a cube-shaped box of side l (Figure 22.10). This molecule has mass m, and is moving with speed c parallel to one side of the box (c is not the speed of light in this case). It rattles back and forth, colliding at regular intervals with the ends of the box and thereby contributing to the pressure of the gas. We are going to work out the pressure this one

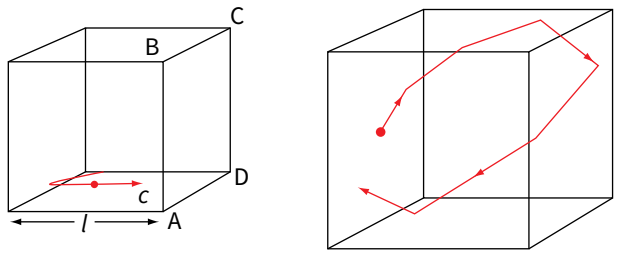


Figure 22.10 A single molecule of a gas, moving in a box.

molecule exerts on one end of the box and then deduce the total pressure produced by all the molecules.

Consider a collision in which the molecule strikes side ABCD of the cube. It rebounds elastically in the opposite direction, so that its velocity is $-c$. its momentum changes from mc to $-mc$. The change in momentum arising from this single collision is thus:

$$\text{change in momentum} = -mc - (+mc)$$
$$= -mc - mc = -2mc$$

Between consecutive collisions with side ABCD, the molecule travels a distance of $2l$ at speed c. Hence:

$$\text{time between collisions with side ABCD} = \frac{2l}{c}$$

Now we can find the force that this one molecule exerts on side ABCD, using Newton's second law of motion. This says that the force produced is equal to the rate of change of momentum:

$$\text{force} = \frac{\text{change in momentum}}{\text{time taken}} = \frac{2mc}{2l/c} = \frac{mc^2}{l}$$

(We use $+2mc$ because now we are considering the force of the molecule on side ABCD, which is in the opposite direction to the change in momentum of the molecule.)

The area of side ABCD is l^2. From the definition of pressure, we have:

$$\text{pressure} = \frac{\text{force}}{\text{area}} = \frac{mc^2/l}{l^2} = \frac{mc^2}{l^3}$$

This is for one molecule, but there is a large number N of molecules in the box. Each has a different velocity, and each contributes to the pressure. We write the average value of c^2 as $<c^2>$, and multiply by N to find the total pressure:

$$\text{pressure } p = \frac{Nm<c^2>}{l^3}$$

But this assumes that all the molecules are travelling in the same direction and colliding with the same pair of opposite faces of the cube. In fact they will be moving in all three dimensions equally, so we need to divide by 3 to find the pressure exerted.

$$\text{pressure } p = \tfrac{1}{3}\frac{Nm<c^2>}{l^3}$$

Here, l^3 is equal to the volume V of the cube, so we can write:

$$p = \frac{1}{3}\frac{Nm}{V}\langle c^2\rangle \quad \text{or} \quad pV = \frac{1}{3}Nm\langle c^2\rangle$$

(Notice that, in the second form of the equation, we have the macroscopic properties of the gas – pressure and volume – on one side of the equation and the microscopic properties of the molecules on the other side.)

Finally, the quantity Nm is the mass of all the molecules of the gas, and this is simply equal to the mass M of the gas. So $\frac{Nm}{V}$ is equal to the density ρ of the gas, and we can write:

$$p = \frac{1}{3}\rho\langle c^2\rangle$$

So the pressure of a gas depends only on its density and the mean square speed of its molecules.

A plausible equation?

It is worth thinking a little about whether the equation $p = \frac{1}{3}\frac{Nm}{V}\langle c^2\rangle$ seems to make sense. It should be clear to you that the pressure is proportional to the number of molecules, N. More molecules mean greater pressure. Also, the greater the mass of each molecule, the greater the force it will exert during a collision.

The equation also suggests that pressure p is proportional to the average value of the speed squared. This is because, if a molecule is moving faster, not only does it strike the container harder, but it also strikes the container more often.

The equation suggests that the pressure p is inversely proportional to the volume occupied by the gas. Here, we have deduced Boyle's law. If we think in terms of the kinetic model, we can see that if a mass of gas occupies a larger volume, the molecules will spend more time in the bulk of the gas, and less time colliding with the walls. So, the pressure will be lower.

These arguments should serve to convince you that the equation is plausible; this sort of argument cannot prove the equation.

Temperature and molecular kinetic energy

Now we can compare the equation $pV = \frac{1}{3}Nm\langle c^2\rangle$ with the ideal gas equation $pV = nRT$. The left-hand sides are the same, so the two right-hand sides must also be equal:

$$\frac{1}{3}Nm\langle c^2\rangle = nRT$$

We can use this equation to tell us how the absolute temperature of a gas (a macroscopic property) is related to the mass and speed of its molecules. If we focus on the quantities of interest, we can see the following relationship:

$$m\langle c^2\rangle = \frac{3nRT}{N}$$

The quantity $\frac{N}{n} = N_A$ is the Avogadro constant, the number of particles in 1 mole. So:

$$m\langle c^2\rangle = \frac{3RT}{N_A}$$

It is easier to make sense of this if we divide both sides by 2, to get the familiar expression for kinetic energy:

$$\frac{1}{2}m\langle c^2\rangle = \frac{3RT}{2N_A}$$

The quantity R/N_A is defined as the **Boltzmann constant**, k. Its value is $1.38 \times 10^{-23}\,\text{J K}^{-1}$. Substituting k in place of R/N_A gives

$$\frac{1}{2}m\langle c^2\rangle = \frac{3kT}{2}$$

The quantity $\frac{1}{2}m\langle c^2\rangle$ is the average kinetic energy E of a molecule in the gas, and k is a constant. Hence the thermodynamic temperature T is proportional to the average kinetic energy of a molecule.

The mean translational kinetic energy of an atom (or molecule) of an ideal gas is proportional to the thermodynamic temperature.

QUESTIONS

17 Check that the units on the left-hand side of the equation $p = \frac{1}{3}\frac{Nm}{V}\langle c^2\rangle$ are the same as those on the right-hand side.

18 The quantity Nm is the total mass of the molecules of the gas, i.e. the mass of the gas. At room temperature, the density of air is about $1.29\,\text{kg m}^{-3}$ at a pressure of $10^5\,\text{Pa}$.

a Use these figures to deduce the value of $\langle c^2\rangle$ for air molecules at room temperature.

b Find a typical value for the speed of a molecule in the air by calculating $\sqrt{\langle c^2\rangle}$. How does this compare with the speed of sound in air, approximately $330\,\text{m s}^{-1}$?

It is easier to recall this as:

mean translational kinetic energy of atom $\propto T$

We need to consider two of the terms in this statement. Firstly, we talk about **translational** kinetic energy. This is the energy that the molecule has because it is moving along; a molecule made of two or more atoms may also spin or tumble around, and is then said to have rotational kinetic energy – see Figure 22.11.

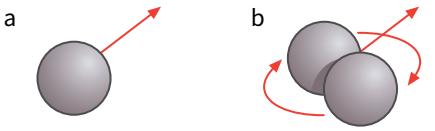


Figure 22.11 **a** A monatomic molecule has only translational kinetic energy. **b** A diatomic molecule can have both translational and rotational kinetic energy.

Secondly, we talk about **mean** (or average) translational kinetic energy. There are two ways to find the average translational kinetic energy (k.e.) of a molecule of a gas. Add up all the kinetic energies of the individual molecules of the gas and then calculate the average k.e. per molecule. Alternatively, watch an individual molecule over a period of time as it moves about, colliding with other molecules and the walls of the container and calculate its average k.e. over this time. Both should give the same answer.

The Boltzmann constant is an important constant in physics because it tells us how a property of microscopic particles (the kinetic energy of gas molecules) is related to a macroscopic property of the gas (its absolute temperature). That is why its units are joules per kelvin (J K^{-1}). Its value is very small ($1.38 \times 10^{-23}\,\text{J K}^{-1}$) because the increase in kinetic energy in J of a molecule is very small for each kelvin increase in temperature.

It is useful to remember the equation linking kinetic energy with temperature as 'average k.e. is three-halves kT'.

QUESTION

19 **a** The Boltzmann constant k is equal to R/N_A. From values of R and N_A, show that k has the value $1.38 \times 10^{-23}\,\text{J K}^{-1}$.

b Write down an equation linking the Boltzmann constant, the thermodynamic temperature and the average kinetic energy of a molecule.

Mass, kinetic energy and temperature

Since mean k.e. $\propto T$, it follows that if we double the thermodynamic temperature of an ideal gas (e.g. from 300 K to 600 K), we double the mean k.e. of its molecules. It doesn't follow that we have doubled their speed; because k.e. $\propto v^2$, their mean speed has increased by a factor of $\sqrt{2}$.

Air is a mixture of several gases: nitrogen, oxygen, carbon dioxide, etc. In a sample of air, the mean k.e. of the nitrogen molecules is the same as that of the oxygen molecules and that of the carbon dioxide molecules. This comes about because they are all repeatedly colliding with one another, sharing their energy. Carbon dioxide molecules have greater mass than oxygen molecules; since their mean translational k.e. is the same, it follows that the carbon dioxide molecules move more slowly than the oxygen molecules.

QUESTIONS

20 Calculate the mean translational k.e. of atoms in an ideal gas at 27 °C.

21 The atoms in a gas have a mean translational k.e. equal to 5.0×10^{-21} J. Calculate the temperature of the gas in K and in °C.

22 Show that, if the mean speed of the molecules in an ideal gas is doubled, the thermodynamic temperature of the gas increases by a factor of 4.

23 A fixed mass of gas expands to twice its original volume at a constant temperature. How do the following change?

a the pressure of the gas

b the mean translational kinetic energy of its molecules.

24 Air consists of molecules of:
oxygen (molar mass = $32\,\text{g mol}^{-1}$) and
nitrogen (molar mass = $28\,\text{g mol}^{-1}$).
Calculate the mean translational k.e. of these molecules in air at 20 °C. Use your answer to estimate a typical speed for each type of molecule.

25 Show that the change in the internal energy of one mole of an ideal gas per unit change in temperature is always a constant. What is this constant?

Summary

- Brownian motion provides evidence for the fast, random movement of molecules in a gas.
- For an ideal gas:

  $\frac{pV}{T}$ = constant
- One mole of any substance contains N_A particles (atoms or molecules).

  N_A = Avogadro constant = $6.02 \times 10^{23}\,\text{mol}^{-1}$
- The equation of state for an ideal gas is:

  $pV = nRT$ for n moles
- From the kinetic model of a gas, we can deduce the relationship:

  $p = \frac{1}{3}\frac{Nm\langle c^2\rangle}{V} = \frac{1}{3}\rho\langle c^2\rangle$

  where $\langle c^2\rangle$ is the mean square molecular speed and ρ is the density of the gas.
- The mean translational kinetic energy E of a particle (atom or molecule) of an ideal gas is proportional to the thermodynamic temperature T.

  $E = \frac{1}{2}m\langle c^2\rangle = \frac{3kT}{2}$

End-of-chapter questions

1 **a** State how many atoms there are in:

i a mole of helium gas [1]

ii a mole of chlorine gas [1]

iii a kilomole of neon gas. [1]

b A container holds four moles of carbon dioxide. Calculate:

i the number of carbon dioxide molecules there are in the container [1]

ii the number of carbon atoms there are in the container [1]

iii the number of oxygen atoms there are in the container. [1]

2 A bar of gold has a mass of 1.0 kg. Calculate:

a the number of moles of gold in the bar [2]

b the number of gold atoms in the bar [1]

c the mass of one gold atom. [1]

(Relative atomic mass of gold = 197.)

3 A cylinder holds 140 dm^3 of nitrogen at room temperature and pressure. Moving slowly so that there is no change in temperature, a piston is pushed to reduce the volume of the nitrogen to 42 dm^3.

a Calculate the pressure of the nitrogen after compression. [2]

b Explain the effect on the temperature and pressure of the nitrogen if the piston were pushed in very quickly. [1]

4 The atmospheric pressure is 100 kPa, equivalent to the pressure exerted by a column of water 10 m high. A bubble of oxygen of volume 0.42 cm^3 is released by a water plant at a depth of 25 m. Calculate the volume of the bubble when it reaches the surface. State any assumptions you make. [4]

5 A cylinder contains 40 dm^3 of carbon dioxide at a pressure of 4.8×10^5 Pa at room temperature.
Calculate:

a the number of moles of carbon dioxide [2]

b the mass of carbon dioxide. [2]

(Relative molecular mass of carbon dioxide = 44.)

6 Calculate the volume of 1 mole of ideal gas at a pressure of 1.01×10^5 Pa and at a temperature of 0 °C. [2]

7 A vessel of volume 200 dm^3 contains 3.0×10^{26} molecules of gas at a temperature of 127 °C. Calculate the pressure exerted by the gas on the vessel walls. [3]

8 **a** Calculate the average speed of helium molecules at room temperature and pressure. (Density of helium at room temperature and pressure = 0.179 kg m^{-3}.) [3]

b Comment on how this speed compares with the average speed of air molecules at the same temperature and pressure. [2]

9 A sample of neon is contained in a cylinder at 27 °C. Its temperature is raised to 243 °C.

a Calculate the kinetic energy of the neon atoms at:

i 27 °C [3]

ii 243 °C. [2]

b Compare the speeds of the molecules at the two temperatures. [2]

10 A truck is to cross the Sahara desert. The journey begins just before dawn when the temperature is 3 °C. The volume of air held in each tyre is 1.50 m^3 and the pressure in the tyres is 3.42×10^5 Pa.

a Explain how the air molecules in the tyre exert a pressure on the tyre walls. [3]

b Calculate the number of moles of air in the tyre. [3]

c By midday the temperature has risen to 42 °C.

i Calculate the pressure in the tyre at this new temperature. You may assume that no air escapes and the volume of the tyre is unchanged. [2]

ii Calculate the increase in the average translational kinetic energy of an air molecule due to this temperature rise. [2]

11 **a** Explain what is meant by **Brownian motion** and how it provides evidence for the existence of molecules. [3]

b The density of air at room temperature and pressure, r.t.p. (20 °C and 1.03×10^5 Pa), is 1.21 kg m^{-3}. Calculate the average speed of air molecules at r.t.p. [4]

c State and explain the effect on the average speed of the air molecules of:

i raising the temperature of the air [2]

ii going to a higher altitude (but keeping the temperature constant). [1]

12 **a** Explain what is meant by an **ideal gas**. [2]

b A cylinder contains 500 g of helium at a pressure of 5.0×10^5 Pa and at a temperature of 27 °C.
You may assume helium acts as an ideal gas.
Calculate:

i the number of moles of helium the cylinder holds [1]

ii the number of atoms of helium the cylinder holds. [1]

c Calculate the volume of the cylinder. [3]

d When the tap of the cylinder is opened for a short time a small amount of the helium escapes into the atmosphere. As it does so, the temperature of the helium drops significantly.

i Explain why the temperature drops. [2]

ii Describe what happens to the average speed of the atoms of the escaped helium. [1]

(Relative atomic mass of helium = 4.)

13 A hot air balloon, its basket and passengers has a total mass of 450 kg. The inflated balloon holds 3000 m^3 of air.

a Calculate the minimum force needed to lift the balloon off the ground. [1]

b Atmospheric pressure is 1.03×10^5 Pa and the density of air at this pressure is 1.23 kg m^{-3}.

i Calculate the mass of 3000 m^3 of air. [1]

ii Calculate the number of moles of air in the balloon. You may consider air at this pressure to act as an ideal gas with a relative molecular mass of 29. [1]

c As the air is heated it expands and some of it is expelled through a vent at the top of the balloon.
Calculate the maximum mass of air that can remain in the balloon to give sufficient upthrust for the balloon to just leave the ground. [2]

d Calculate the minimum temperature that the air inside the balloon must be heated to for the balloon to just lift off the ground. [4]



Chapter 23: Coulomb's law

Learning outcomes

You should be able to:

- recall and use Coulomb's law
- calculate the field strength and potential due to a point charge
- define electric potential and relate field strength to the potential gradient
- compare and contrast electric and gravitational fields

Living in a field

The scientist in the photograph (Figure 23.1) is using a detector to measure the electric field produced by a mobile phone mast. People often worry that the electric field produced by a mobile phone transmitter may be harmful, but detailed studies have yet to show any evidence for this. If you hold a mobile phone close to your ear, the field strength will be far greater than that produced by a nearby mast.



Figure 23.1 Mobile phone masts produce weak electric fields – this scientist is using a small antenna to detect and measure the field of a nearby mast to ensure that it is within safe limits.

Electric fields

In Chapter 8, we presented some fundamental ideas about electric fields:

- An electric field is a field of force and can be represented by field lines.
- The electric field strength at a point is the force per unit positive charge that acts on a stationary charge:

$$\text{field strength} = \frac{\text{force}}{\text{charge}} \qquad E = \frac{F}{Q}$$

- There is a uniform field between charged parallel plates:

$$\text{field strength} = \frac{\text{potential difference}}{\text{separation}} \qquad E = \frac{V}{d}$$

In this chapter, we will extend these ideas to consider how electric fields arise from electric charges. We will also compare electric fields with gravitational fields (Chapter 18).

Coulomb's law

Any electrically charged object produces an electric field in the space around it. It could be something as small as an electron or a proton, or as large as a planet or star. To say that it produces an electric field means that it will exert a force on any other charged object which is in the field. How can we determine the size of such a force?

The answer to this was first discovered by Charles Coulomb, a French physicist. He realised that it was important to think in terms of **point charges**; that is, electrical charges which are infinitesimally small so that we need not worry about their shapes. In 1785 Coulomb proposed a law that describes the force that one charged particle exerts on another. This law is remarkably similar in form to Newton's law of gravitation (page 273).

A statement of **Coulomb's law** is as follows:

> Any two point charges exert an electrical force on each other that is proportional to the product of their charges and inversely proportional to the square of the distance between them.

We consider two point charges Q_1 and Q_2 separated by a distance r (Figure 23.2). The force each charge exerts on the other is F. According to Newton's third law of motion, the point charges interact with each other and therefore exert equal but opposite forces on each other.

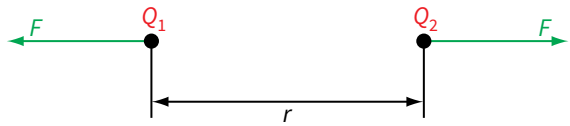


Figure 23.2 The variables involved in Coulomb's law.

According to Coulomb's law, we have:

$$\text{force} \propto \text{product of the charges} \qquad F \propto Q_1 Q_2$$

$$\text{force} \propto \frac{1}{\text{distance}^2} \qquad F \propto \frac{1}{r^2}$$

Therefore:

$$F \propto \frac{Q_1 Q_2}{r^2}$$

We can write this in a mathematical form:

$$F = \frac{kQ_1 Q_2}{r^2}$$

The constant of proportionality is:

$$k = \frac{1}{4\pi\varepsilon_0}$$

where ε_0 is known as the **permittivity of free space** (ε is the Greek letter epsilon). The value of ε_0 is approximately $8.85 \times 10^{-12}\,\text{F m}^{-1}$. An equation for Coulomb's law is thus:

$$F = \frac{Q_1 Q_2}{4\pi\varepsilon_0 r^2}$$

By substituting for π and ε_0, we can show that the force F can also be given by the equation:

$$F \approx 9.0 \times 10^9 \frac{Q_1 Q_2}{r^2}$$

i.e. the constant k has the approximate numerical value of $9.0 \times 10^9\,\text{N m}^2\,\text{C}^{-2}$.

This approximation can be useful for making rough calculations, but more precise calculations require that the value of ε_0 given above be used.

Following your earlier study of Newton's law of gravitation, you should not be surprised by this relationship. The force depends on each of the properties producing it (in this case, the charges), and it is an inverse square law with distance – if the particles are twice as far apart, the electrical force is a quarter of its previous value (Figure 23.3).

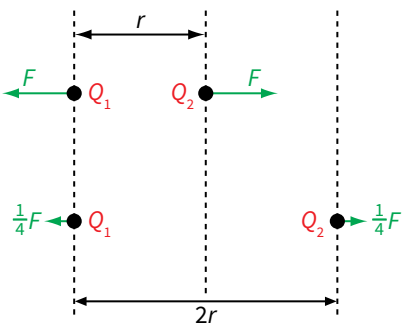


Figure 23.3 Doubling the separation results in one-quarter of the force, a direct consequence of Coulomb's law.

Note also that, if we have a positive and a negative charge, then the force F is negative. We interpret this as an attraction. Positive forces, as between two like charges, are repulsive. In gravity, we only have attraction.

So far we have considered point charges. If we are considering uniformly charged spheres we measure the distance from the centre of one to the centre of the other – they behave as if their charge was all concentrated at the centre. Hence we can apply the equation for Coulomb's law for both point charges (e.g. protons, electrons, etc.) and uniformly charged spheres, as long as we use the **centre-to-centre** distance between the objects.

BOX 23.1: Investigating Coulomb's law

It is quite tricky to investigate the force between charged objects, because charge tends to leak away into the air or to the Earth during the course of any experiment. The amount of charge we can investigate is difficult to measure, and usually small, giving rise to tiny forces.

Figure 23.4 shows one method for investigating the inverse square law for two charged metal balls (polystyrene balls coated with conducting silver paint). As one charged ball is lowered down towards the other, their separation decreases and so the force increases, giving an increased reading on the balance.

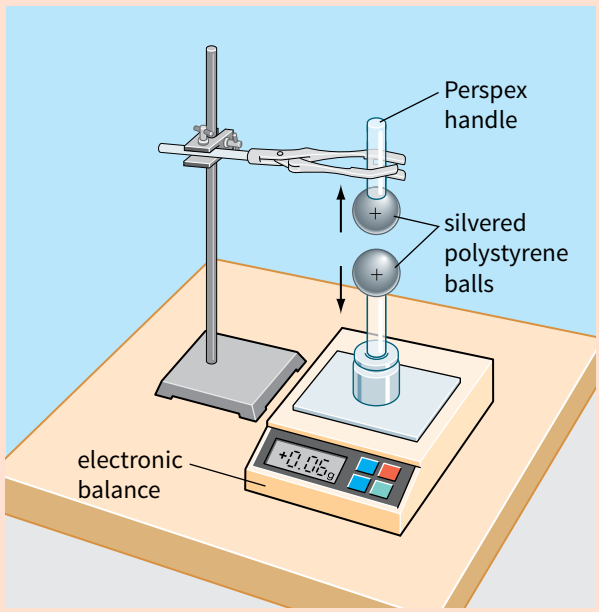


Figure 23.4 Investigating Coulomb's law.

Electric field strength for a radial field

In Chapter 8, we saw that the electric field strength at a point is defined as the force per unit charge exerted on a positive charge placed at that point, $E = \frac{F}{Q}$.

So, to find the field strength near a point charge Q_1 (or outside a uniformly charged sphere), we have to imagine a small positive test charge Q_2 placed in the field, and determine the force per unit charge on it. We can then use the definition above to determine the electric field strength for a point (or spherical) charge.

The force between the two point charges is given by:

$$F = \frac{Q_1 Q_2}{4\pi\varepsilon_0 r^2}$$

The electric field strength E due to the charge Q_1 at a distance of r from its centre is thus:

$$E = \frac{\text{force}}{\text{test charge}} = \frac{Q_1 Q_2}{4\pi\varepsilon_0 r^2 Q_2}$$

or:

$$E = \frac{Q}{4\pi\varepsilon_0 r^2}$$

The field strength E is not a constant; it decreases as the distance r increases. The field strength obeys an inverse square law with distance – just like the gravitational field strength for a point mass. The field strength will decrease by a factor of four when the distance from the centre is doubled.

Note also that, since force is a vector quantity, it follows that electric field strength is also a vector. We need to give its direction as well as its magnitude in order to specify it completely. Worked example 1 shows how to use the equation for field strength near a charged sphere.

QUESTIONS

You will need the data below to answer the following questions. (You may take the charge of each sphere to be situated at its centre.)

$\varepsilon_0 = 8.85 \times 10^{-12}\,\text{F m}^{-1}$

1 A metal sphere of radius 20 cm carries a positive charge of +2.0 μC.

a What is the electric field strength at a distance of 25 cm from the centre of the sphere?

b An identical metal sphere carrying a negative charge of −1.0 μC is placed next to the first sphere. There is a gap of 10 cm between them. Calculate the electric force that each sphere exerts on the other.

Remember to calculate the centre-to-centre distance between the two spheres.

c Determine the electric field strength midway along a line joining the centres of the spheres.

2 A Van de Graaff generator produces sparks when the field strength at its surface is $4.0 \times 10^4\,\text{V cm}^{-1}$. If the diameter of the sphere is 40 cm, what is the charge on it?

WORKED EXAMPLE

1 A metal sphere of diameter 12 cm is positively charged. The electric field strength at the surface of the sphere is $4.0 \times 10^5\,\text{V m}^{-1}$. Draw the electric field pattern for the sphere and determine the total surface charge.

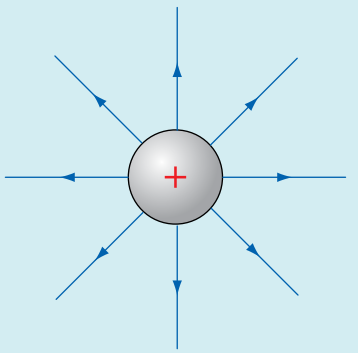


Figure 23.5 The electric field around a charged sphere.

Step 1 Draw the electric field pattern (Figure 23.5). The electric field lines must be normal to the surface and radial.

Step 2 Write down the quantities given:

electric field strength $E = 4.0 \times 10^5\,\text{V m}^{-1}$

radius $r = \frac{0.12}{2} = 0.06\,\text{m}$

Step 3 Use the equation for the electric field strength to determine the surface charge:

$$E = \frac{Q}{4\pi\varepsilon_0 r^2} = 0.06\,\text{m}$$

$$Q = 4\pi\varepsilon_0 r^2 \times E$$

$$= 4\pi \times 8.85 \times 10^{-12} \times (0.06)^2 \times 4.0 \times 10^5$$

$$= 1.6 \times 10^{-7}\,\text{C}\ (0.16\,\mu\text{C})$$

Electric potential

When we discussed gravitational potential (page 276), we started from the idea of potential energy. The potential at a point is then the potential energy of unit mass at the point. We will approach the idea of electrical potential in the same way. However, you may be relieved to find that you already know something about the idea of electrical potential, because you know about voltage and potential difference. This section shows how we formalise the idea of voltage, and why we use the expression 'potential difference' for some kinds of voltage.

Electric potential energy

When an electric charge moves through an electric field, its potential energy changes. Think about this concrete example: if you want to move one positive charge closer to another positive charge, you have to push it (Figure 23.6). This is simply because there is a force of repulsion between the charges. You have to do work in order to move one charge closer to the other.



Figure 23.6 Work must be done to push one positive charge towards another.

In the process of doing work, energy is transferred from you to the charge that you are pushing. Its potential energy increases. If you let go of the charge, it will move away from the repelling charge. This is analogous to lifting up a mass; it gains gravitational potential energy as you lift it, and it falls if you let go.

Energy changes in a uniform field

We can also think about moving a positive charge in a uniform electric field between two charged parallel plates. If we move the charge towards the positive plate, we have to do work. The potential energy of the charge is therefore increasing. If we move it towards the negative plate, its potential energy is decreasing (Figure 23.7a).

Since the force is the same at all points in a uniform electric field, it follows that the energy of the charge

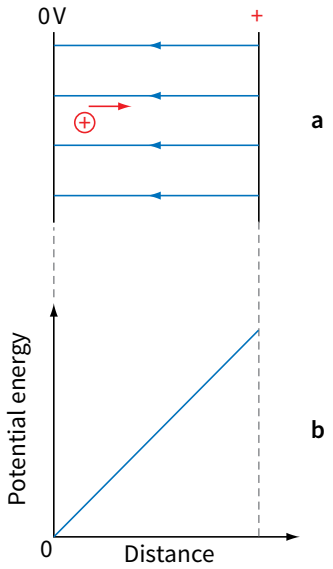


Figure 23.7 Electrostatic potential energy changes in a uniform field.

increases steadily as we push it from the negative plate to the positive plate. The graph of potential energy against distance is a straight line, as shown in Figure 23.7b.

We can calculate the change in potential energy of a charge Q as it is moved from the negative plate to the positive plate very simply. Potential difference is defined as the energy change per coulomb between two points (recall from Chapter 9 that one volt is one joule per coulomb). Hence, for charge Q, the work done in moving it from the negative plate to the positive plate is:

$$W = QV$$

We can rearrange this equation as:

$$V = \frac{W}{Q}$$

This is really how voltage V is defined. It is the energy per unit positive charge at a point in an electric field. By analogy with gravitational potential, we call this the electric potential at a point. Now you should be able to see that what we regard as the familiar idea of voltage should more correctly be referred to as electric potential. The difference in potential between two points is the potential difference (p.d.) between them.

Just as with gravitational fields, we must define the zero of potential (this is the point where we consider a charge to have zero potential energy). Usually, in a laboratory situation, we define the Earth as being at a potential of zero volts. If we draw two parallel charged plates arranged horizontally, with the lower one earthed (Figure 23.8), you can see immediately how similar this is to our idea of gravitational fields. The diagram also shows how we can include equipotential lines in a representation of an electric field.

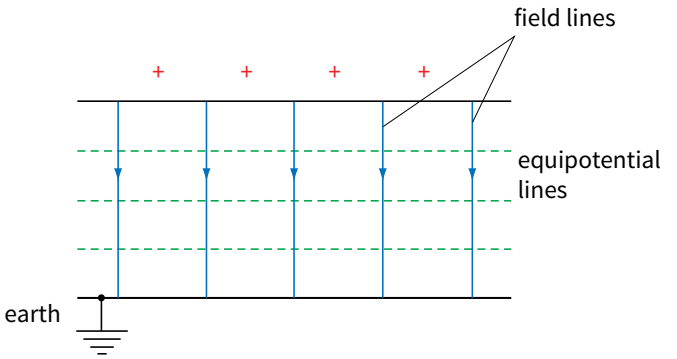


Figure 23.8 Equipotential lines in a uniform electric field.

We can extend the idea of electric potential to measurements in electric fields. In Figure 23.9, the power supply provides a potential difference of 10 V. The value of the potential at various points is shown. You can see that the middle resistor has a potential difference across it of (8 – 2) V = 6 V.

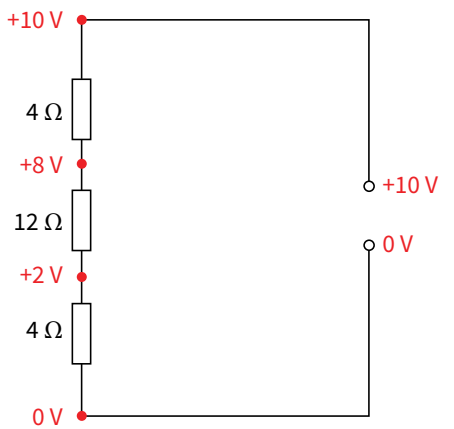


Figure 23.9 Changes in potential (shown in red) around an electric circuit.

Energy in a radial field

Imagine again pushing a small positive test charge towards a large positive charge. At first, the repulsive force is weak, and you have only to do a small amount of work. As you get closer, however, the force increases (Coulomb's law), and you have to work harder and harder.

The potential energy of the test charge increases as you push it. It increases more and more rapidly the closer you get to the repelling charge. This is shown by the graph in Figure 23.10. We can write an equation for the potential V at a distance r from a charge Q:

$$V = \frac{Q}{4\pi\varepsilon_0 r}$$

(This comes from the calculus process of integration, applied to the Coulomb's law equation.) You should be able to see how this relationship parallels the equivalent formula for gravitational potential in a radial field:

$$\phi = -\frac{GM}{r}$$

Note that we do not need the minus sign in the electric equation as it is included in the charge. A negative charge gives an attractive (negative) field whereas a positive charge gives a repulsive (positive) field.

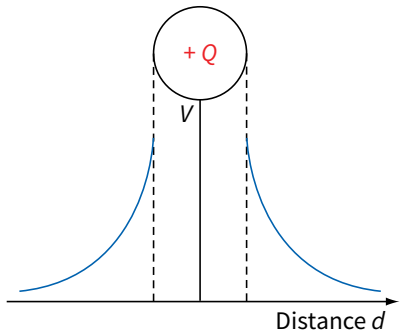


Figure 23.10 The potential changes according to an inverse law near a charged sphere.

We can show these same ideas by drawing field lines and equipotential lines. The equipotentials get closer together as we get closer to the charge (Figure 23.11).

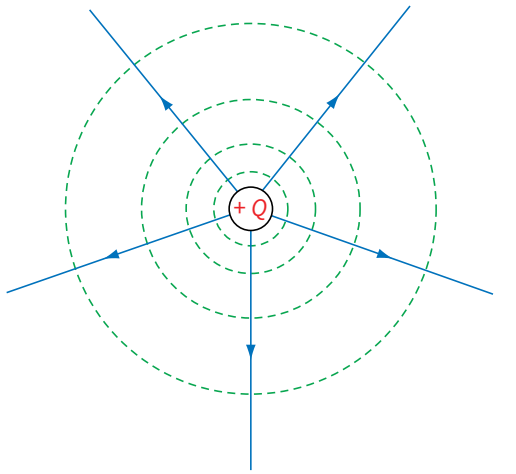


Figure 23.11 The electric field around a positive charge. The dashed equipotential lines are like the contour lines on a map; they are spaced at equal intervals of potential.

To arrive at the result above, we must again define our zero of potential. Again, we say that a charge has zero potential energy when it is at infinity (some place where it is beyond the influence of any other charges). If we move towards a positive charge, the potential is positive. If we move towards a negative charge, the potential is negative.

This allows us to give a definition of electric potential:

> The **electric potential** at a point is equal to the work done in bringing unit positive charge from infinity to that point.

Electric potential is a scalar quantity. To calculate the potential at a point caused by more than one charge, find each potential separately and add them. Remember that positive charges cause positive potentials and negative charges cause negative potentials.

Field strength and potential gradient

We can picture electric potential in the same way that we thought about gravitational potential. A negative charge attracts a positive test charge, so we can regard it as a potential 'well'. A positive charge is the opposite – a 'hill' (Figure 23.12). The strength of the field is shown by the slope of the hill or well:

field strength = –potential gradient

The minus sign is needed because, if we are going up a potential hill, the force on us is pushing us back down the slope, in the opposite direction.

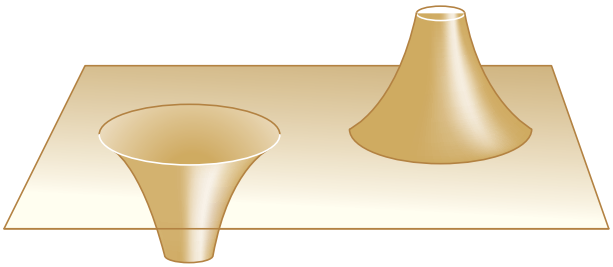


Figure 23.12 A 'potential well' near a negative charge, and a 'potential hill' near a positive charge.

This relationship applies to all electric fields. For the special case of a uniform field, the potential gradient E is constant. Its value is given by

$$E = \frac{\Delta V}{\Delta d}$$

where V is the potential difference between two points separated by a distance d. (This is the same as the relationship $E = \frac{V}{d}$ quoted in Chapter 8, page 120.)

Worked example 2 shows how to determine the field strength from a potential–distance graph.

QUESTIONS

3 **a** What would be the electrical potential energy of a charge of +1 C placed at each of the points A, B, C, D between the charged, parallel plates shown in Figure 23.13?

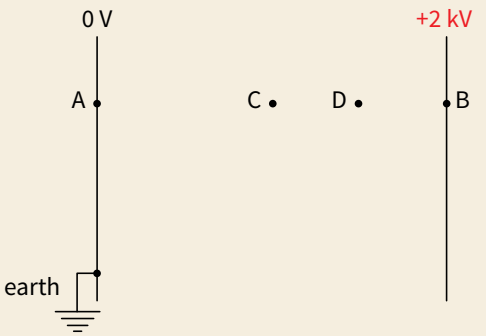


Figure 23.13 A uniform electric field – see Question 3.

b What would be the potential energy of a +2 C charge at each of these points? (C is half-way between A and B; D is half-way between C and B.)

4 A Van de Graaff generator has a spherical dome of radius 10 cm. It is charged up to a potential of 100 000 V (100 kV). How much charge does it store? What is the potential at a distance of 10 cm from the dome?

5 **a** How much work would be done in moving a +1 C charge along the following paths shown in Figure 23.14: from E to H; from E to F; from F to G; from H to E?

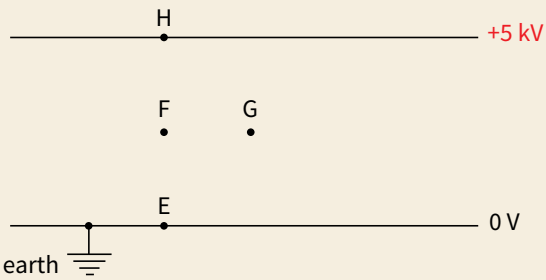


Figure 23.14 A uniform electric field – see Question 5.

b How would your answers differ for:

i a –1 C charge?

ii a +2 C charge?

WORKED EXAMPLE

2 The graph (Figure 23.15) shows how the electric potential varies near a charged object. Calculate the electric field strength at a point 5 cm from the centre of the object.

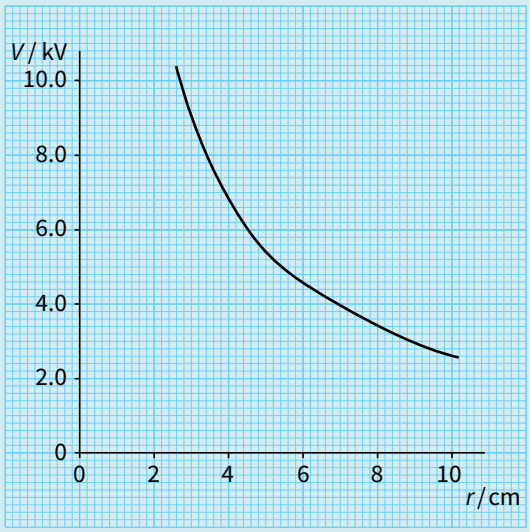


Figure 23.15 Variation of the potential V near a positively charged object.

Step 1 Draw the tangent to the graph at the point 5.0 cm. This is shown in Figure 23.16.

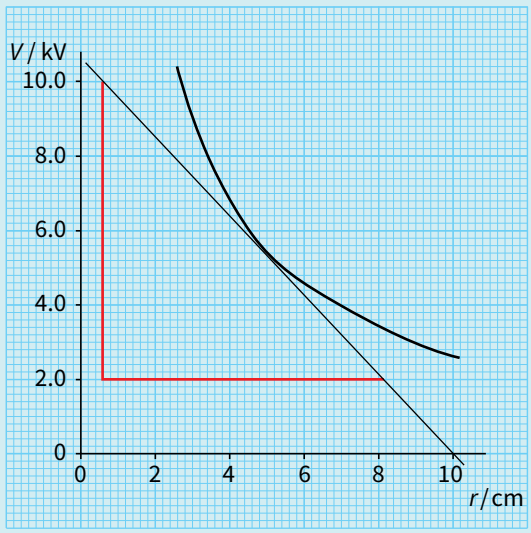


Figure 23.16 Drawing the tangent to the V–r graph to find the electric field strength E.

Step 2 Calculate the gradient of the tangent:

$$\text{gradient} = \frac{\Delta v}{\Delta r}$$

$$= \frac{(10.0 - 2.0)}{(0.6 - 8.2)}$$

$$= -1.05\ \text{kV cm}^{-1} = -1.05 \times 10^5\ \text{V m}^{-1}$$

$$\approx -1.1 \times 10^5\ \text{V m}^{-1}$$

The electric field strength is therefore $+1.1 \times 10^5\ \text{V m}^{-1}$ or $+1.1 \times 10^5\ \text{N C}^{-1}$.

Remember E = −potential gradient.

Comparing gravitational and electric fields

There are obvious similarities between the ideas we have used in this chapter to describe electric fields and those we used in Chapter 18 for gravitational fields. This can be helpful, or it can be confusing! The summary given in Table 23.1 is intended to help you to sort them out.

An important difference is this: electric charges can be positive or negative, so they can attract or repel. There are no negative masses, so there is only attraction in a gravitational field.

QUESTION

You will need the data below to answer the question.

proton mass = 1.67×10^{-27} kg

proton charge = $+1.60 \times 10^{-19}$ C

$\varepsilon_0 = 8.85 \times 10^{-12}\ \text{F m}^{-1}$

$G = 6.67 \times 10^{-11}\ \text{N m}^2\ \text{kg}^{-2}$

6 Two protons in the nucleus of an atom are separated by a distance of 10^{-15} m. Calculate the electrostatic force of repulsion between them, and the force of gravitational attraction between them. (Assume the protons behave as point charges and point masses.) Is the attractive gravitational force enough to balance the repulsive electrical force? What does this suggest to you about the forces between protons within a nucleus?

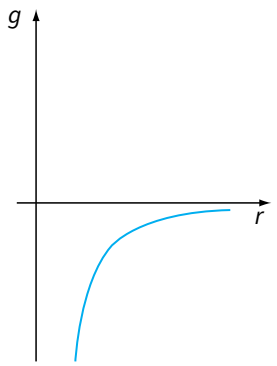
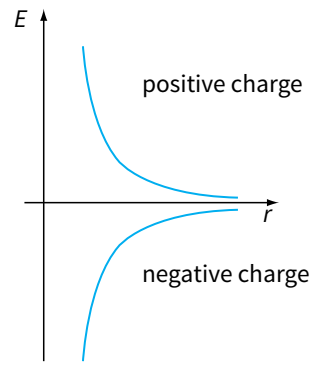
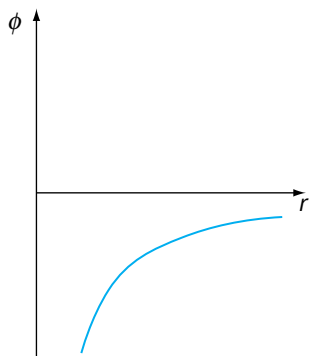
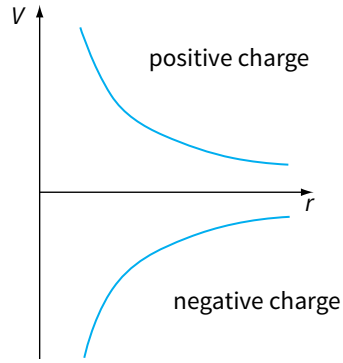
Gravitational fields	Electric fields
Origin arise from masses	**Origin** arise from electric charges
Vector forces only gravitational attraction, no repulsion	**Vector forces** both electrical attraction and repulsion are possible (because of positive and negative charges)
All gravitational fields field strength $g = \frac{F}{m}$ i.e. field strength is force per unit mass	**All electric fields** field strength $E = \frac{F}{Q}$ i.e. field strength is force per unit positive charge
Units F in N, g in N kg^{-1} or m s^{-2}	**Units** F in N, E in N C^{-1} or V m^{-1}
Uniform gravitational fields parallel gravitational field lines g = constant	**Uniform electric fields** parallel electric field lines $E = \frac{V}{d}$ = constant
Spherical gravitational fields radial field lines force given by Newton's law: $F = \frac{GMm}{r^2}$ field strength is therefore: $g = \frac{GM}{r^2}$ (Gravitational forces are always attractive, so we show g on a graph against r as negative.) force and field strength obey an inverse square law with distance 	**Spherical electric fields** radial field lines force given by Coulomb's law: $F = \frac{Q_1 Q_2}{4\pi\varepsilon_0 r^2}$ field strength is therefore: $E = \frac{Q}{4\pi\varepsilon_0 r^2}$ (A negative charge gives an attractive field, a positive charge gives a repulsive field.) force and field strength obey an inverse square law with distance 
Gravitational potential given by: $\phi = -\frac{GM}{r}$ potential obeys an inverse relationship with distance and is zero at infinity potential is a scalar quantity and is always negative 	**Electric potential** given by: $V = \frac{Q}{4\pi\varepsilon_0 r}$ potential obeys an inverse relationship with distance and is zero at infinity potential is a scalar quantity 

Table 23.1 Gravitational and electric fields compared.

Summary

- Coulomb's law states that two point charges exert an electrical force on each other that is proportional to the product of their charges and inversely proportional to the square of the distance between them.
- The equation for Coulomb's law is:

$$F = \frac{Q_1 Q_2}{4\pi\varepsilon_0 r^2}$$

- A point charge Q gives rise to a radial field. The electric field strength is given by the equation:

$$E = \frac{Q}{4\pi\varepsilon_0 r^2}$$

- The electric potential at a point is defined as the work done per unit positive charge in bringing charge from infinity to the point.
- For a point charge, the electric potential is given by:

$$V = \frac{Q}{4\pi\varepsilon_0 r}$$

End-of-chapter questions

1 On a copy of Figure 23.17, draw the electric fields between the charged objects. [5]

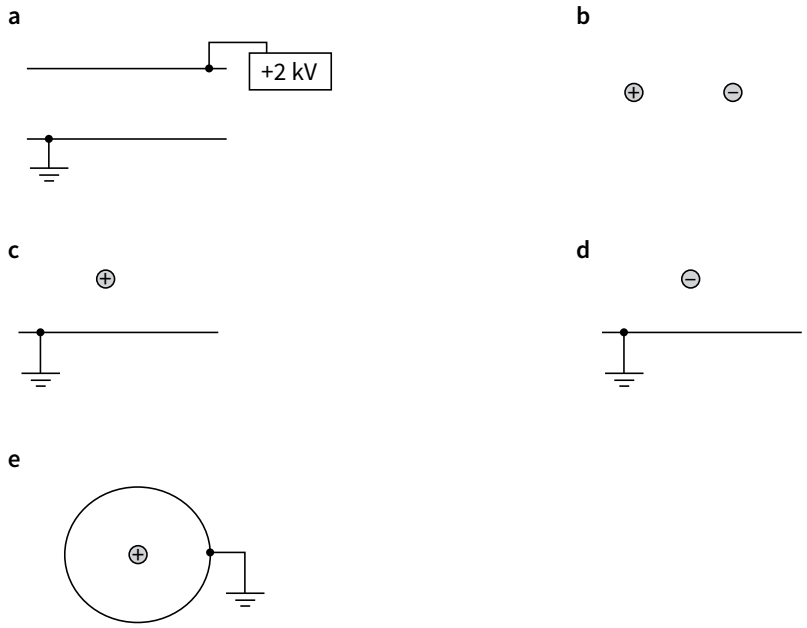


Figure 23.17 For End-of-chapter Question 1.

2 Two parallel plates are 4 cm apart and have a potential difference of 2.5 kV between them.

a Calculate the electric field strength between the plates. [2]

b A small piece of dust carrying a charge of +2.4 nC moves into the space between the plates.

i Calculate the force on the dust particle. [2]

ii The mass of the dust particle is 4.2 μg. Calculate the acceleration of the particle towards the negative plate. [2]

3 A small sphere carries a charge of 2.4×10^{-9} C. Calculate the electric field strength at a distance of:

a 2 cm from the centre of the sphere [2]

b 4 cm from the centre of the sphere. [2]

4 A conducting sphere of diameter 6.0 cm is mounted on an insulating base. The sphere is connected to a power supply which has an output voltage of 20 kV.

a Calculate the charge on the sphere. [3]

b Calculate the electric field strength at the surface of the sphere. [2]

5 The nucleus of a hydrogen atom carries a charge of $+1.60 \times 10^{-19}$ C. Its electron is at a distance of 1.05×10^{-10} m from the nucleus. Calculate the ionisation potential of hydrogen. [3]
(Hint: This is equal to the work per unit charge needed to remove the electron to infinity.)

6 **a** Define **electric field strength**. [2]

b Two charged conducting spheres, each of radius 1.0 cm, are placed with their centres 10 cm apart, as shown in Figure 23.18.

Sphere A carries a charge of $+2 \times 10^{-9}$ C.

Figure 23.19 shows how the electric field strength between the two spheres varies with distance x.

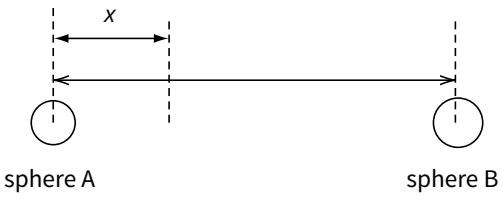


Figure 23.18 For End-of-chapter Question 6.

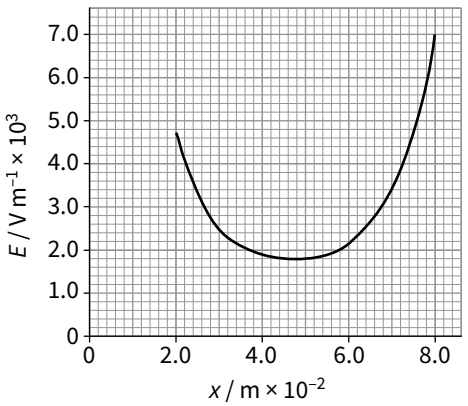


Figure 23.19 Variation of electric field strength with distance x from the centre of sphere A.

i Calculate the field produced by sphere A at the 5.0 cm mark. [2]

ii Use your result to **b i** to calculate the charge on sphere B. [3]

c **i** Sphere B is now removed. Calculate the potential at the surface of sphere A. [2]

ii Suggest and explain how the potential at the surface of sphere A would compare before and after sphere B was removed. [2]

7 An α-particle emitted in the radioactive decay of radium has a kinetic energy of 8.0×10^{-13} J.

a **i** Calculate the potential difference that an α-particle, initially at rest, would have to be accelerated through to gain this energy. [2]

ii Calculate the speed of the α-particle at this kinetic energy. [3]

b Figure 23.20 shows the path of an α-particle of this energy as it approaches a gold nucleus head-on.

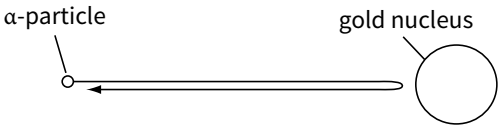


Figure 23.20 For End-of-chapter Question 7.

i State the speed of the α-particle at its point of closest approach to the gold nucleus. [1]

ii Write down the kinetic energy of the α-particle at this point. [1]

iii Write down the potential energy of the α-particle at this point. [1]

c Use your answer to **b iii** to show that the α-particle will reach a distance of 4.5×10^{-14} m from the centre of the gold nucleus. [2]

d Suggest and explain what this information tells us about the gold nucleus. [2]

(Mass of an α-particle = 6.65×10^{-27} kg, charge on an α-particle = +2*e*, charge on a gold nucleus = +79*e*.)

8 **a** Define **electric potential at a point**. [2]

b Figure 23.21 shows the potential well near a hydrogen nucleus.

The first electron orbital can be considered to be a circle of diameter 1.04×10^{-10} m.

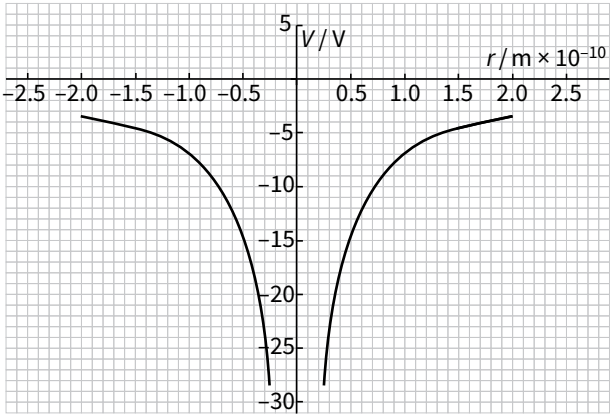


Figure 23.21 For End-of-chapter Question 8.

i Determine the potential at a point on this orbital. [2]

ii Calculate the energy required to ionise the atom. [2]

c Use the graph to estimate the electric field strength at a distance of 1.0×10^{-10} m from the centre of the nucleus. [2]

9 Figure 23.22 shows a conducting sphere of radius 0.80 cm carrying a charge of $+6.0 \times 10^{-8}$ C resting on a balance.

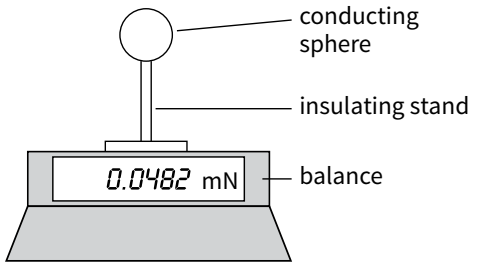


Figure 23.22 For End-of-chapter Question 9.

a Calculate the electric field at the surface of the sphere. [2]

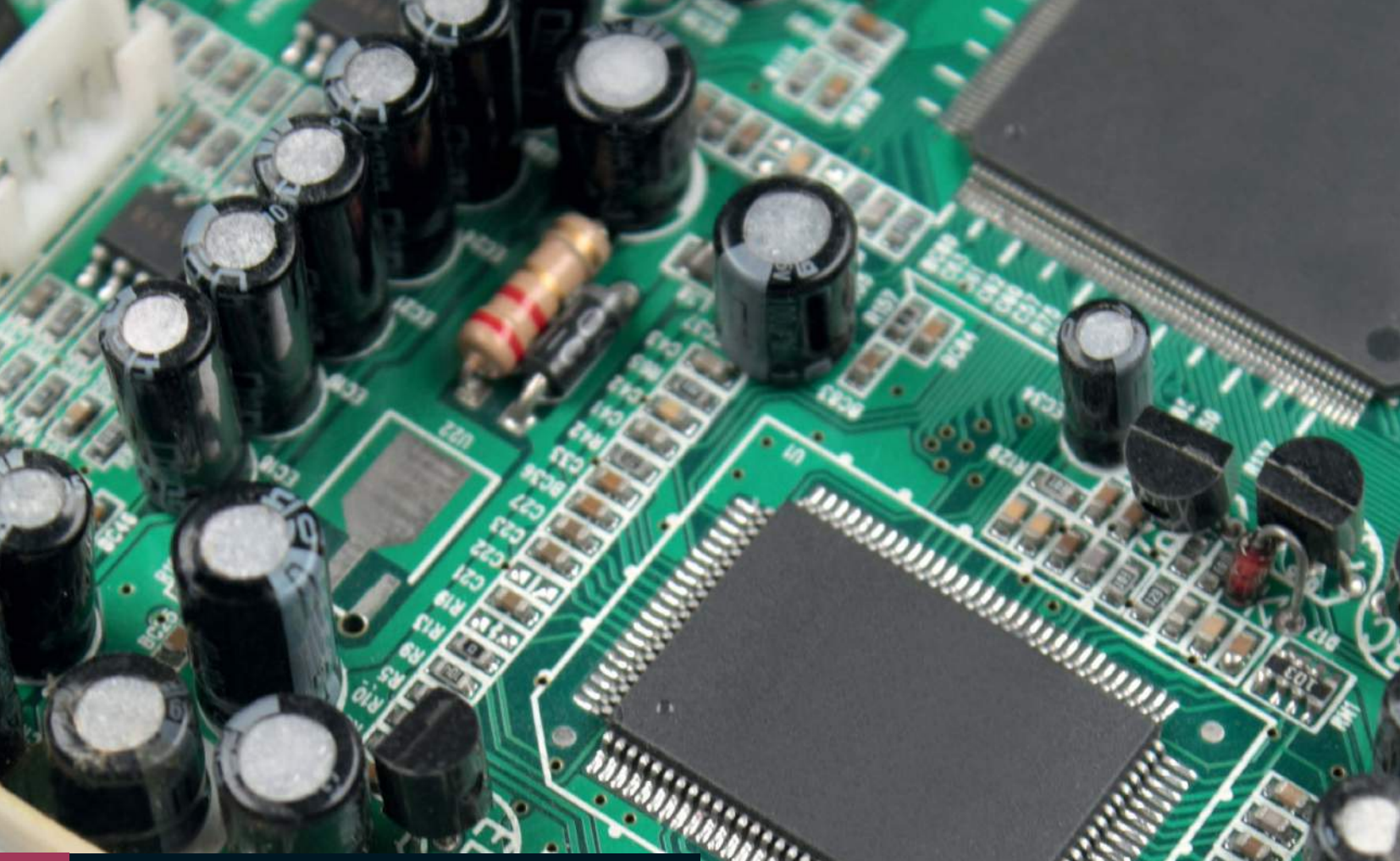
b An identical sphere carrying a charge of -4.5×10^{-8} C is held so that its centre is 5.0 cm vertically above the centre of the first sphere.

i Calculate the electric force between the two spheres. [2]

ii Calculate the new reading on the balance. [1]

c The second sphere is moved vertically downwards through 1.5 cm. Calculate the work done against the electric field in moving the sphere. [3]

Chapter 24: Capacitance



Learning outcomes

You should be able to:

- define capacitance and its unit, the farad
- solve problems involving charge, voltage and capacitance
- deduce and use a formula for the energy stored by a capacitor
- derive and use formulae for capacitances in series and parallel

Capacitors

Most electronic devices, such as radios, computers and MP3 players, make use of components called capacitors. These are usually quite small, but Figure 24.1 shows a giant capacitor, specially constructed to store electrical energy at the Fermilab particle accelerator in the United States.



Figure 24.1 One of the world's largest capacitors, built to store energy at the Fermilab particle accelerator.

Capacitors in use

Capacitors are used to store energy in electrical and electronic circuits. This means that they have many valuable applications. For example, capacitors are used in computers; they store energy in normal use, and then they gradually release this energy if there is a power failure, so that the computer will operate long enough to save valuable data. The photograph (Figure 24.2) shows a variety of shapes and sizes of capacitors.



Figure 24.2 A variety of capacitors.

Every capacitor has two leads, each connected to a metal plate. To store energy, these two plates must be given equal and opposite electric charges. Between the plates is an insulating material called the **dielectric**. Figure 24.3 shows a simplified version of the construction of a capacitor; in practice, many have a spiral 'Swiss-roll' form.

To move charge onto the plates of a capacitor, it must be connected to a voltage supply. The negative terminal of the supply pushes electrons onto one plate, making it negatively charged. Electrons are repelled from the other plate, making it positively charged. Figure 24.4 shows that there is a flow of electrons all the way round the circuit.

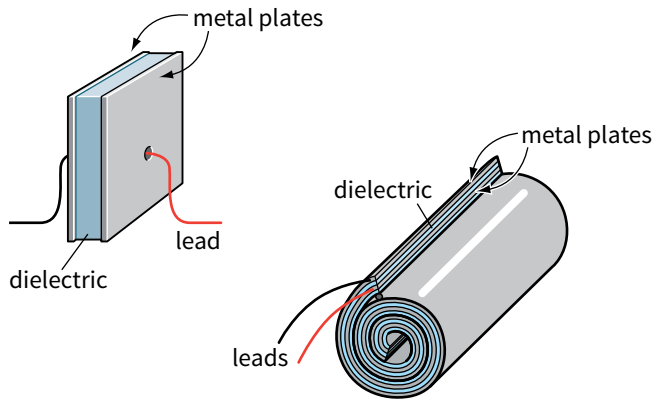


Figure 24.3 The construction of two types of capacitor.

The two ammeters will give identical readings. The current stops when the potential difference (p.d.) across the capacitor is equal to the electromotive force (e.m.f.) of the supply. We then say that the capacitor is 'fully charged'.

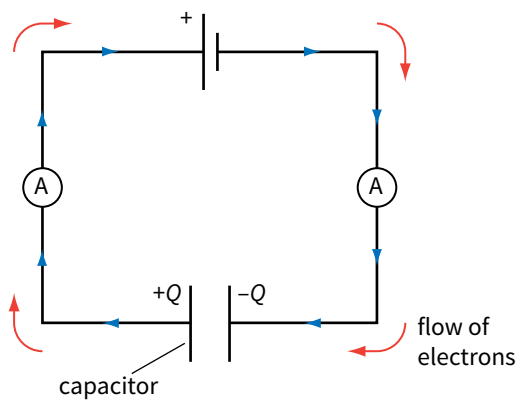


Figure 24.4 The flow of charge when a capacitor is charged up.

Note: The convention is that current is the flow of positive charge. Here, it is free electrons that flow. Electrons are negatively charged; conventional current flows in the opposite direction to the electrons (Figure 24.5).

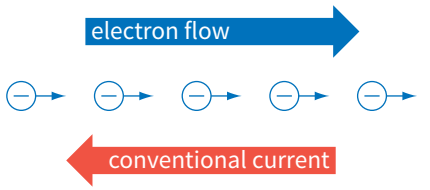


Figure 24.5 A flow of electrons to the right constitutes a conventional current to the left.

Charge on the plates

Think about a capacitor with uncharged plates. Each plate has equal amounts of positive and negative charge. Connecting the capacitor to a supply pulls charge $+Q$ from one plate and transfers it to the other, leaving behind charge $-Q$. The supply does work in separating the charges. Since the two plates now store equal and opposite charges, the total charge on the capacitor is zero. When we talk about the 'charge stored' by a capacitor, we mean the quantity Q, the magnitude of the charge stored on each plate.

To make the capacitor plates store more charge, we would have to use a supply of higher e.m.f. If we connect the leads of the charged capacitor together, electrons flow back around the circuit and the capacitor is discharged.

You can observe a capacitor discharging as follows. Connect the two leads of a capacitor to the terminals of a battery. Disconnect, and then reconnect the leads to a light-emitting diode (LED). It is best to have a protective resistor in series with the LED. The LED will glow briefly as the capacitor discharges.

In any circuit, the charge that flows past a point in a given time is equal to the area under a current–time graph (just as distance is equal to the area under a speed–time graph). So the magnitude of the charge on the plates in a capacitor is given by the area under the current–time graph recorded while the capacitor is being charged up.

The meaning of capacitance

If you look at some capacitors, you will see that they are marked with the value of their **capacitance**. The greater the capacitance, the greater is the charge on the capacitor plates for a given potential difference across it.

The capacitance C of a capacitor is defined by:

$$\text{capacitance} = \frac{\text{charge}}{\text{potential difference}}$$

or $$C = \frac{Q}{V}$$

where Q is the magnitude of the charge on each of the capacitor's plates and V is the potential difference across it the capacitor.

> The capacitance of a capacitor is the charge stored on one plate per unit of potential difference between the plates.

The charge on the capacitor may be calculated using the equation:

$$Q = VC$$

This equation shows that the charge depends on two things: the capacitance C and the voltage V (double the voltage means double the charge). Note that it isn't only capacitors that have capacitance. Any object can become charged by connecting it to a voltage. The object's capacitance is then the ratio of the charge to the voltage.

Units of capacitance

The unit of capacitance is the **farad**, F. From the equation that defines capacitance, you can see that this must be the same as the units of charge (coulombs, C) divided by voltage (V):

$$1\,\text{F} = 1\,\text{C}\,\text{V}^{-1}$$

(It is unfortunate that the letter 'C' is used for both capacitance and coulomb. There is room for confusion here!)

In practice, a farad is a large unit. Few capacitors have a capacitance of 1F. Capacitors usually have their values marked in picofarads (pF), nanofarads (nF) or microfarads (μF):

$$1\,\text{pF} = 10^{-12}\,\text{F} \qquad 1\,\text{nF} = 10^{-9}\,\text{F} \qquad 1\,\mu\text{F} = 10^{-6}\,\text{F}$$

Other markings on capacitors

Many capacitors are marked with their highest safe working voltage. If you exceed this value, charge may leak across between the plates, and the dielectric will cease to be an insulator. Some capacitors (electrolytic ones) must be connected correctly in a circuit. They have an indication to show which end must be connected to the positive of the supply. Failure to connect correctly will damage the capacitor, and can be extremely dangerous.

QUESTIONS

1 Calculate the charge on a 220 μF capacitor charged up to 15 V. Give your answer in microcoulombs (μC) and in coulombs (C).

2 A charge of 1.0×10^{-3} C is measured on a capacitor with a potential difference across it of 500 V. Calculate the capacitance in farads (F), microfarads (μF) and picofarads (pF).

3 Calculate the average current required to charge a 50 μF capacitor to a p.d. of 10 V in a time interval of 0.01 s.

4 A student connects an uncharged capacitor of capacitance C in series with a resistor, a cell and a switch. The student closes the switch and records the current I at intervals of 10 s. The results are shown in Table 24.1. The potential difference across the capacitor after 60 s was 8.5 V. Plot a current–time graph, and use it to estimate the value of C.

t/s	0	10	20	30	40	50	60
I/μA	200	142	102	75	51	37	27

Table 24.1 Data for Question 4.

Energy stored in a capacitor

When you charge a capacitor, you use a power supply to push electrons onto one plate and off the other. The power supply does work on the electrons, so their potential energy increases. You recover this energy when you discharge the capacitor.

If you charge a large capacitor (1000 μF or more) to a potential difference of 6.0 V, disconnect it from the supply, and then connect it across a 6.0 V lamp, you can see the lamp glow as energy is released from the capacitor. The lamp will flash briefly. Clearly, such a capacitor does not store much energy when it is charged.

In order to charge a capacitor, work must be done to push electrons onto one plate and off the other (Figure 24.6). At first, there is only a small amount of negative charge on the left-hand plate. Adding more electrons is relatively easy, because there is not much repulsion. As the charge on the plate increases, the repulsion between the electrons on the plate and the new electrons increases, and a greater amount of work must be done to increase the charge on the plate.

This can be seen qualitatively in Figure 24.7a. This graph shows how the p.d. V increases as the amount of charge Q increases. It is a straight line because Q and V are related by:

$$V = \frac{Q}{C}$$

We can use Figure 24.7a to calculate the work done in charging up the capacitor.

First, consider the work done W in moving charge Q through a constant p.d. V. This is given by:

$$W = QV$$

(You studied this equation in Chapter 9.) From the graph of Q against V (Figure 24.7b), we can see that the quantity $Q \times V$ is given by the area under the graph.

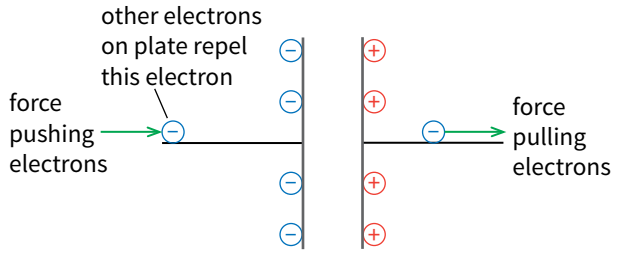


Figure 24.6 When a capacitor is charged, work must be done to push additional electrons against the repulsion of the electrons that are already present.

The area under a graph of p.d. against charge is equal to work done.

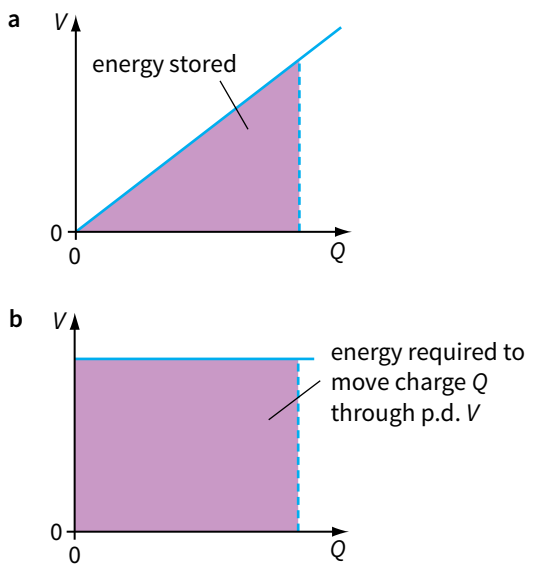


Figure 24.7 The area under a graph of voltage against charge gives a quantity of energy. The area in **a** shows the energy stored in a capacitor; the area in **b** shows the energy required to drive a charge through a resistor.

If we apply the same idea to the capacitor graph (Figure 24.7a), then the area under the graph is the shaded triangle, with an area of $\frac{1}{2}$ base × height. Hence the work done in charging a capacitor to a particular p.d. is given by:

$$W = \tfrac{1}{2}QV$$

Substituting $Q = CV$ into this equation gives two further equations:

$$W = \tfrac{1}{2}CV^2$$

and

$$W = \tfrac{1}{2}\frac{Q^2}{C}$$

These three equations show the work done in charging up the capacitor. This is equal to the energy stored by the capacitor, since this is the amount of energy released when the capacitor is discharged.

We can also see from the second formula ($W = \frac{1}{2}CV^2$) that the energy W that a capacitor stores depends on its capacitance C and the potential difference V to which it is charged.

The energy W stored is proportional to the square of the potential difference V ($W \propto V^2$). It follows that doubling the charging voltage means that four times as much energy is stored.

WORKED EXAMPLE

1 A 2000 μF capacitor is charged to a p.d. of 10 V. Calculate the energy stored by the capacitor.

Step 1 Write down the quantities we know:

$C = 2000\,\mu F$

$V = 10\,V$

Step 2 Write down the equation for energy stored and substitute values:

$W = \frac{1}{2}CV^2$

$W = \frac{1}{2} \times 2000 \times 10^{-6} \times 10^2 = 0.10\,J$

This is a small amount of energy – compare it with the energy stored by a rechargeable battery, typically of the order of 10 000 J. A charged capacitor will not keep an MP3 player running for any length of time.

QUESTIONS

5 State the quantity represented by the gradient of the straight line shown in Figure 24.7a.

6 The graph of Figure 24.8 shows how V depends on Q for a particular capacitor.

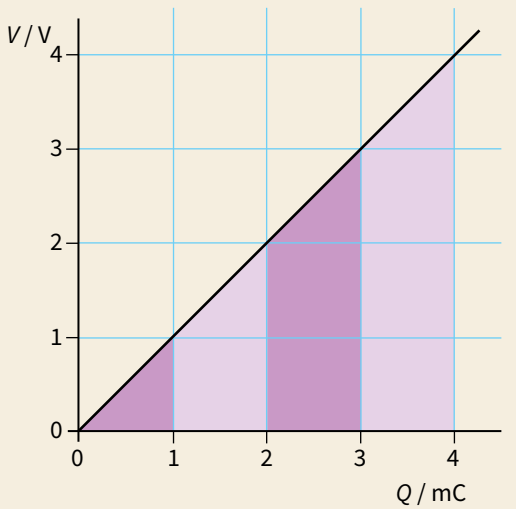


Figure 24.8 The energy stored by a capacitor is equal to the area under the graph of voltage against charge.

The area under the graph has been divided into strips to make it easy to calculate the energy stored. The first strip (which is simply a triangle) shows the energy stored when the capacitor is charged up to 1.0 V. The energy stored is:

$\frac{1}{2}QV = \frac{1}{2} \times 1.0\,mC \times 1.0\,V = 0.5\,mJ$

a Calculate the capacitance C of the capacitor.

b Copy Table 24.2 and complete it by calculating the areas of successive strips, to show how W depends on V.

c Plot a graph of W against V. Describe the shape of this graph.

Q / mC	V / V	Area of strip ΔW / mJ	Sum of areas W / mJ
1.0	1.0	0.5	0.5
2.0	2.0	1.5	2.0
3.0			
4.0			

Table 24.2 Data for Question 6.

BOX 24.1: Investigating energy stored in a capacitor

If you have a sensitive joulemeter (capable of measuring millijoules, mJ), you can investigate the equation for energy stored. A suitable circuit is shown in Figure 24.9.

The capacitor is charged up when the switch connects it to the power supply. When the switch is altered, the capacitor discharges through the joulemeter. (It is important to wait for the capacitor to discharge completely.) The joulemeter will measure the amount of energy released by the capacitor.

By using capacitors with different values of C, and by changing the charging voltage V, you can investigate how the energy W stored depends on C and V.

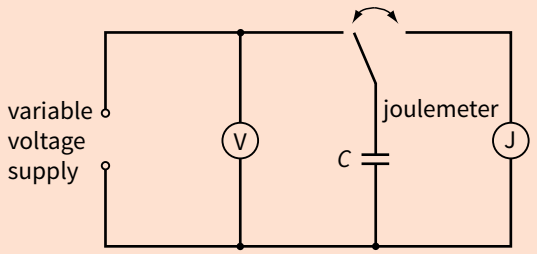


Figure 24.9 With the switch to the left, the capacitor C charges up; to the right, it discharges through the joulemeter.

QUESTIONS

7 Calculate the energy stored in the following capacitors:
- **a** a 5000 μF capacitor charged to 5.0 V
- **b** a 5000 pF capacitor charged to 5.0 V
- **c** a 200 μF capacitor charged to 230 V.

8 Which involves more charge, a 100 μF capacitor charged to 200 V or a 200 μF capacitor charged to 100 V? Which stores more energy?

9 A 10 000 μF capacitor is charged to 12 V, and then connected across a lamp rated at '12 V, 36 W'.
- **a** Calculate the energy stored by the capacitor.
- **b** Estimate the time the lamp stays fully lit. Assume that energy is dissipated in the lamp at a steady rate.

10 In a simple photographic flashgun, a 0.20 F capacitor is charged by a 9.0 V battery. It is then discharged in a flash of duration 0.01 s. Calculate:
- **a** the charge on and energy stored by the capacitor
- **b** the average power dissipated during the flash
- **c** the average current in the flash bulb
- **d** the approximate resistance of the bulb.

Capacitors in parallel

Capacitors are used in electric circuits to store energy. Situations often arise where two or more capacitors are connected together in a circuit. In this section, we will look at capacitors connected in parallel. The next section deals with capacitors in series.

When two capacitors are connected in parallel (Figure 24.10), their combined or total capacitance C_{total} is simply the sum of their individual capacitances C_1 and C_2:

$$C_{total} = C_1 + C_2$$

This is because, when two capacitors are connected together, they are equivalent to a single capacitor with larger plates. The bigger the plates, the more charge that can be stored for a given voltage, and hence the greater the capacitance.

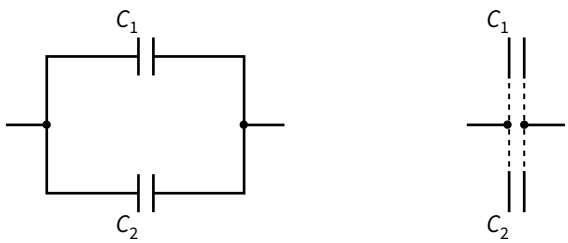


Figure 24.10 Two capacitors connected in parallel are equivalent to a single, larger capacitor.

The total charge Q on two capacitors connected in parallel and charged to a potential difference V is simply given by:

$$Q = C_{total} \times V$$

For three or more capacitors connected in parallel, the equation for their total capacitance becomes:

$$C_{total} = C_1 + C_2 + C_3 + \cdots$$

Capacitors in parallel: deriving the formula

We can derive the equation for capacitors in parallel by thinking about the charge on the two capacitors. As shown in Figure 24.11, C_1 stores charge Q_1 and C_2 stores charge Q_2. Since the p.d. across each capacitor is V, we can write:

$$Q_1 = C_1V \quad \text{and} \quad Q_2 = C_2V$$

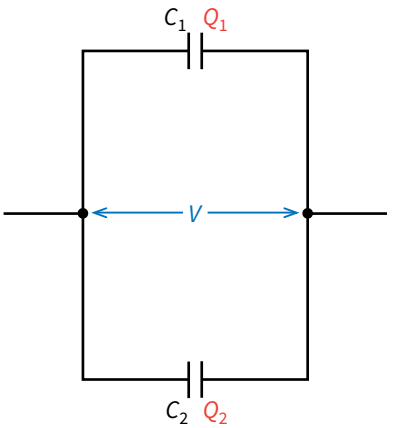


Figure 24.11 Two capacitors connected in parallel have the same p.d. across them, but different amounts of charge.

The total charge is given by the sum of these:

$$Q = Q_1 + Q_2 = C_1V + C_2V$$

Since V is a common factor:

$$Q = (C_1 + C_2)V$$

Comparing this with $Q = C_{total}V$ gives the required $C_{total} = C_1 + C_2$. It follows that for three or more capacitors connected in parallel, we have:

$$C_{total} = C_1 + C_2 + C_3 + \cdots$$

Capacitors in parallel: summary

For capacitors in parallel, the following rules apply:

- The p.d. across each capacitor is the same.
- The total charge on the capacitors is equal to the sum of the charges:
  $$Q_{total} = Q_1 + Q_2 + Q_3 + \cdots$$
- The total capacitance C_{total} is given by:
  $$C_{total} = C_1 + C_2 + C_3 + \cdots$$

QUESTIONS

11 **a** Calculate the total capacitance of two 100 μF capacitors connected in parallel.

b Calculate the total charge they store when charged to a p.d. of 20 V.

12 A capacitor of capacitance 50 μF is required, but the only values available to you are 10 μF, 20 μF and 100 μF (you may use more than one of each value). How would you achieve the required value by connecting capacitors in parallel? Give at least **two** answers.

Capacitors in series

In a similar way to the case of capacitors connected in parallel, we can consider two or more capacitors connected in series (Figure 24.12). The total capacitance C_{total} of two capacitors of capacitances C_1 and C_2 is given by:

$$\frac{1}{C_{total}} = \frac{1}{C_1} + \frac{1}{C_2}$$

Here, it is the reciprocals of the capacitances that must be added to give the reciprocal of the total capacitance. For three or more capacitors connected in series, the equation for their total capacitance is:

$$\frac{1}{C_{total}} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \cdots$$

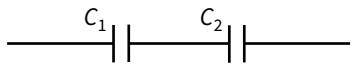


Figure 24.12 Two capacitors connected in series.

Capacitors in series: deriving the formula

The same principles apply here as for the case of capacitors in parallel. Figure 24.13 shows the situation. C_1 and C_2 are connected in series, and there is a p.d. V across them. This p.d. is divided (it is shared between the two capacitors), so that the p.d. across C_1 is V_1 and the p.d. across C_2 is V_2. It follows that:

$$V = V_1 + V_2$$

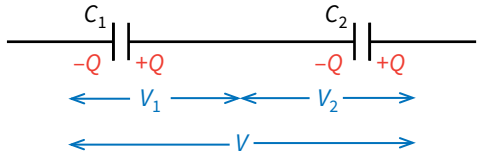


Figure 24.13 Capacitors connected in series store the same charge, but they have different p.d.s across them.

Now we must think about the charge stored by the combination of capacitors. In Figure 24.13, you will see that both capacitors are shown as storing the same charge Q. How does this come about? When the voltage is first applied, charge $-Q$ arrives on the left-hand plate of C_1. This repels charge $-Q$ off the right-hand plate, leaving it with charge $+Q$. Charge $-Q$ now arrives on the left-hand plate of C_2, and this in turn results in charge $+Q$ on the right-hand plate.

Note that charge is not arbitrarily created or destroyed in this process – the total amount of charge in the system is constant. This is an example of the conservation of charge.

Notice also that there is a central isolated section of the circuit between the two capacitors. Since this is initially uncharged, it must remain so at the end. This requirement is satisfied, because there is charge $-Q$ at one end and $+Q$ at the other. Hence we conclude that capacitors connected in series store the same charge. This allows us to write equations for V_1 and V_2:

$$V_2 = \frac{Q}{C_1} \quad \text{and} \quad V_2 = \frac{Q}{C_2}$$

The combination of capacitors stores charge Q when charged to p.d. V, and so we can write:

$$V = \frac{Q}{C_{\text{total}}}$$

Substituting these in $V = V_1 + V_2$ gives:

$$\frac{Q}{C_{\text{total}}} = \frac{Q}{C_1} + \frac{Q}{C_2}$$

Cancelling the common factor of Q gives the required equation:

$$\frac{1}{C_{\text{total}}} = \frac{1}{C_1} + \frac{1}{C_2}$$

Worked example 2 shows how to use this relationship.

QUESTIONS

13 Calculate the total capacitance of three capacitors of capacitances 200 μF, 300 μF and 600 μF, connected in series.

14 You have a number of identical capacitors, each of capacitance C. Determine the total capacitance when:

- **a** two capacitors are connected in series
- **b** n capacitors are connected in series
- **c** two capacitors are connected in parallel
- **d** n capacitors are connected in parallel.

WORKED EXAMPLE

2 Calculate the total capacitance of a 300 μF capacitor and a 600 μF capacitor connected in series.

Step 1 The calculation should be done in two steps; this is relatively simple using a calculator with a $\frac{1}{x}$ or x^{-1} key.

Substitute the values into the equation:

$$\frac{1}{C_{\text{total}}} = \frac{1}{C_1} + \frac{1}{C_2}$$

This gives:

$$\frac{1}{C_{\text{total}}} = \frac{1}{300} + \frac{1}{600}$$

$$\frac{1}{C_{\text{total}}} = 0.005\,\mu\text{F}^{-1}$$

Step 2 Now take the reciprocal of this value to determine the capacitance in μF:

$$C_{\text{total}} = \frac{1}{0.005} = 200\,\mu\text{F}$$

Notice that the total capacitance of two capacitors in series is less than either of the individual capacitances.

Using the x^{-1} key on your calculator, you can also do this calculation in one step:

$C_{\text{total}} = (300^{-1} + 600^{-1})^{-1} = 200\,\mu\text{F}$

Comparing capacitors and resistors

It is helpful to compare the formulae for capacitors in series and parallel with the corresponding formulae for resistors (Table 24.3).

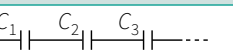
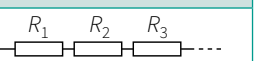
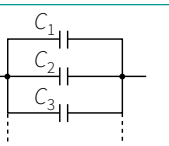
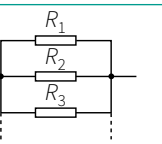
	Capacitors	Resistors
In series		
	store same charge	have same current
	$\frac{1}{C_{\text{total}}} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} + \cdots$	$R_{\text{total}} = R_1 + R_2 + R_3 + \cdots$
In parallel		
	have same p.d.	have same p.d.
	$C_{\text{total}} = C_1 + C_2 + C_3 + \cdots$	$\frac{1}{R_{\text{total}}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \cdots$

Table 24.3 Capacitors and resistors compared.

Notice that the reciprocal formula applies to capacitors in series but to resistors in parallel. This comes from the definitions of capacitance and resistance. Capacitance indicates how good a capacitor is at storing charge for a given voltage, and resistance indicates how **bad** a resistor is at letting current through for a given voltage.

QUESTION

15 The conductance G of a resistor indicates how **good** a resistor is at letting current through for a given voltage.

It is the reciprocal of the resistance: $G = \frac{1}{R}$.

Write down equations for the combined conductance G_{total} of two resistors whose conductances are G_1 and G_2, connected:

a in series

b in parallel.

Capacitor networks

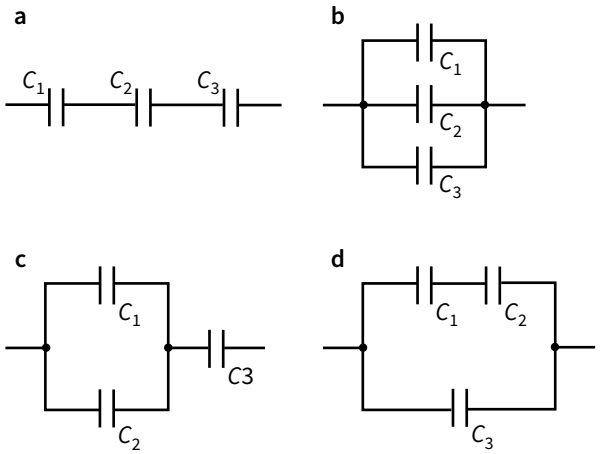


Figure 24.14 Four ways to connect three capacitors.

There are four ways in which three capacitors may be connected together. These are shown in Figure 24.14. The combined capacitance of the first two arrangements (three capacitors in series, three in parallel) can be calculated using the formulae above. The other combinations must be dealt with in a different way:

- Figure 24.14a – All in series. Calculate C_{total} as in Table 24.3.
- Figure 24.14b – All in parallel. Calculate C_{total} as in Table 24.3.
- Figure 24.14c – Calculate C_{total} for the two capacitors of capacitances C_1 and C_2, which are connected in parallel, and then take account of the third capacitor of capacitance C_3, which is connected in series.
- Figure 24.14d – Calculate C_{total} for the two capacitors of capacitances C_1 and C_2, which are connected in series, and then take account of the third capacitor of capacitance C_3, which is connected in parallel.

These are the same approaches as would be used for networks of resistors.

QUESTIONS

16 For each of the four circuits shown in Figure 24.14, calculate the total capacitance in μF if each capacitor has capacitance 100 μF.

17 Given a number of 100 μF capacitors, how might you connect networks to give the following values of capacitance:

a 400 μF **b** 25 μF **c** 250 μF?

(Note that, in each case, there is more than one correct answer; try to find the answer that requires the minimum number of capacitors.)

18 You have three capacitors of capacitances 100 pF, 200 pF and 600 pF. Determine the maximum and minimum values of capacitance that you can make by connecting them together to form a network. State how they should be connected in each case.

19 Calculate the capacitance in μF of the network of capacitors shown in Figure 24.15.

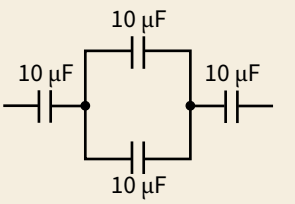


Figure 24.15 A capacitor network – see Question 19.

Sharing charge, sharing energy

If a capacitor is charged and then connected to a second capacitor (Figure 24.16), what happens to the charge and the energy that it stores? Note that, when the capacitors are connected together, they are in parallel, because they have the same p.d. across them. Their combined capacitance C_{total} is equal to the sum of their individual capacitances.

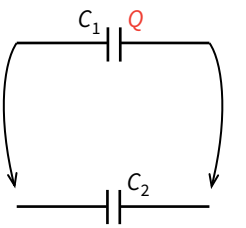


Figure 24.16 Capacitor of capacitance C_1 is charged and then connected across C_2.

Now we can think about the charge stored, Q. This is shared between the two capacitors; the total amount of charge stored must remain the same, since charge is conserved. The charge is shared between the two capacitors in proportion to their capacitances. Now the p.d. can be calculated from $V = \frac{Q}{C}$ and the energy from $W = \frac{1}{2}CV^2$.

If we look at a numerical example, we find an interesting result (Worked example 3).

Figure 24.17 shows an analogy to the situation described in Worked example 3. Capacitors are represented by containers of water. A wide (high capacitance) container is filled to a certain level (p.d.). It is then connected to a container with a smaller capacitance, and the levels equalise. (The p.d. is the same for each.) Notice that the potential energy of the water has decreased, because the height of its centre of gravity above the base level has decreased. Energy is dissipated as heat, as there is friction both within the moving water and between the water and the container.

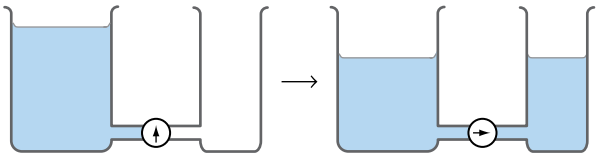


Figure 24.17 An analogy for the sharing of charge between capacitors.

WORKED EXAMPLE

3 Consider two 100 mF capacitors. One is charged to 10 V, disconnected from the power supply, and then connected across the other. Calculate the energy stored by the combination.

Step 1 Calculate the charge and energy stored for the single capacitor.

initial charge $Q = VC = 10 \times 0.10 = 1.0\,\text{C}$

initial stored energy $\frac{1}{2}CV^2 = \frac{1}{2} \times 0.10 \times 10^2$

$= 5.0\,\text{J}$

Step 2 Calculate the final p.d. across the capacitors. The capacitors are in parallel and have a total stored charge of 1.0 C.

$C_{\text{total}} = C_1 + C_2 = 100 + 100 = 200\,\text{mF}$

The p.d. V can be determined using $Q = VC$.

$$V = \frac{Q}{C} = \frac{1.0}{200} \times 10^{-3} = 5.0\,\text{V}$$

This is because the charge is shared equally, with the original capacitor losing half of its charge.

Step 3 Now calculate the total energy stored by the capacitors.

total energy $\frac{1}{2}CV^2 = \frac{1}{2} \times 200 \times 10^{-3} \times 5.0^2$

$= 2.5\,\text{J}$

The charge stored remains the same, but half of the stored energy is lost. This energy is lost in the connecting wires as heat as electrons migrate between the capacitors.

QUESTIONS

20 Three capacitors, each of capacitance 120 μF, are connected together in series. This network is then connected to a 10 kV supply. Calculate:

- **a** their combined capacitance in μF
- **b** the charge stored
- **c** the total energy stored.

21 A 20 μF capacitor is charged up to 200 V and then disconnected from the supply. It is then connected across a 5.0 μF capacitor. Calculate:

- **a** the combined capacitance of the two capacitors in μF
- **b** the charge they store
- **c** the p.d. across the combination
- **d** the energy dissipated when they are connected together.

Capacitance of isolated bodies

It is not just capacitors that have capacitance – all bodies have capacitance. Yes, even you have capacitance! You may have noticed that, particularly in dry conditions, you may become charged up, perhaps by rubbing against a synthetic fabric. You are at a high voltage and store a significant amount of charge. Discharging yourself by touching an earthed metal object would produce a spark.

If we consider a conducting sphere of radius r insulated from its surroundings and carrying a charge Q it will have a potential at its surface of V, where

$$V = \frac{1}{4\pi\varepsilon_0}\frac{Q}{r}$$

Since $C = \frac{Q}{V}$ it follows that the capacitance of a sphere is $C = 4\pi\varepsilon_0 r$.

QUESTION

22 Estimate the capacitance of the Earth given that it has a radius of 6.4×10^6 m.

Summary

- Capacitors are constructed from two metal sheets ('plates'), separated by an insulating material. A capacitor stores equal and opposite amounts of charge on its plates.
- For a capacitor, the charge stored is directly proportional to the p.d. between the plates: $Q = VC$.
- Capacitance is the charge stored per unit of p.d.
- A farad is a coulomb per volt: $1\,\text{F} = 1\,\text{C}\,\text{V}^{-1}$.
- Capacitors store energy. The energy W stored at p.d. V is:

$$W = \tfrac{1}{2}QV = \tfrac{1}{2}CV^2 = \tfrac{1}{2}\frac{Q^2}{C}$$

- The formula $W = \frac{1}{2}QV$ is deduced from the area under a graph of potential difference against charge.
- For capacitors connected in parallel and in series, the combined capacitances are as follows:

parallel: $C_{\text{total}} = C_1 + C_2 + C_3 + \cdots$

series: $\dfrac{1}{C_{\text{total}}} = \dfrac{1}{C_1} + \dfrac{1}{C_2} + \dfrac{1}{C_3} + \cdots$

- These formulae are derived from conservation of charge and addition of p.d.s.

End-of-chapter questions

1 A 470 μF capacitor is connected across the terminals of a battery of e.m.f. 9 V. Calculate the charge on the plates of the capacitor. [1]

2 Calculate the p.d. across the terminals of a 2200 μF capacitor when it has a charge of 0.033 C on its plates. [1]

3 Calculate the capacitance of a capacitor if it stores a charge of 2.0 C when there is a potential difference of 5000 V across its plates. [1]

4 Calculate the energy stored when a 470 μF capacitor has a potential difference of 12 V across its plates. [1]

5 Calculate the energy stored on a capacitor if it stores 1.5 mC of charge when there is a potential difference of 50 V across it. [1]

6 A 5000 μF capacitor has a p.d. of 24 V across its plates.

a Calculate the energy stored on the capacitor. [1]

b The capacitor is briefly connected across a bulb and half the charge flows off the capacitor. Calculate the energy dissipated in the lamp. [3]

7 A 4700 μF capacitor has a p.d. of 12 V across its terminals. It is connected to a resistor and the charge leaks away through the resistor in 2.5 s.

a Calculate the energy stored on the capacitor. [1]

b Calculate the charge stored on the capacitor. [1]

c Estimate the average current through the resistor. [1]

d Estimate the resistance of the resistor. [1]

e Suggest why the last two quantities can only be estimates. [1]

8 An electronic engineer is designing a circuit in which a capacitor of capacitance of 4700 μF is to be connected across a potential difference of 9.0 V. He has four 4700 μF, 6 V capacitors available. Draw a diagram to show how the four capacitors could be used for this purpose. [1]

9 Calculate the different capacitances that can be made from three 100 μF capacitors. For each value, draw the network that is used. [4]

10 Figure 24.18 shows three capacitors connected in series with a cell of e.m.f. 1.5 V.

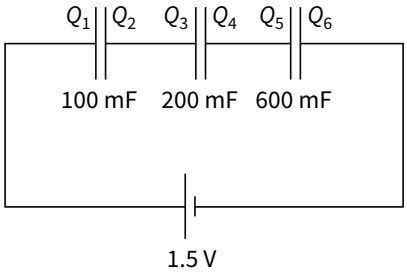


Figure 24.18 For End-of-chapter Question 10.

a Calculate the charges Q_1 to Q_6 on each of the plates. [5]

b Calculate the p.d. across each capacitor. [3]

11 **a** State **one** use of a capacitor in a simple electric circuit. [1]

b Figure 24.19 shows a circuit used to investigate the discharge of a capacitor, and Figure 24.20 is a graph showing the change in current with time when the capacitor is discharged.

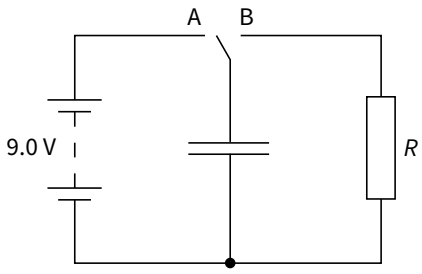


Figure 24.19 Circuit for End-of-chapter Question 11.

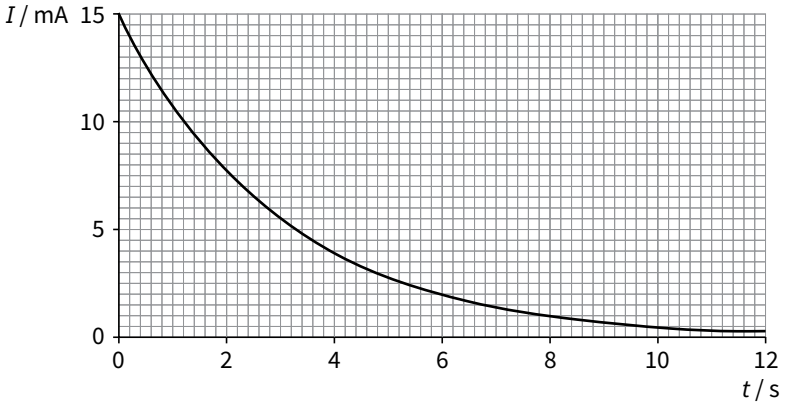


Figure 24.20 Graph for End-of-chapter Question 11.

i Deduce the resistance R of the resistor. [2]

ii Explain why the current decreases as the capacitor discharges. [2]

iii The charge on the capacitor is equal to the area under the graph. Estimate the charge on the capacitor when the potential difference across it is 9.0 V. [2]

iv Calculate the capacitance of the capacitor. [2]

12 The spherical dome on a Van de Graaff generator is placed near an earthed metal plate. The dome has a diameter of 40 cm and the potential at its surface is 54 kV.

a **i** Calculate the charge on the dome. [2]

ii Calculate the capacitance of the dome. [2]

The metal plate is moved slowly towards the sphere but does not touch it. The sphere discharges through the air to the plate. The graph in Figure 24.21 shows how the potential at the surface of the sphere changes during the discharge.

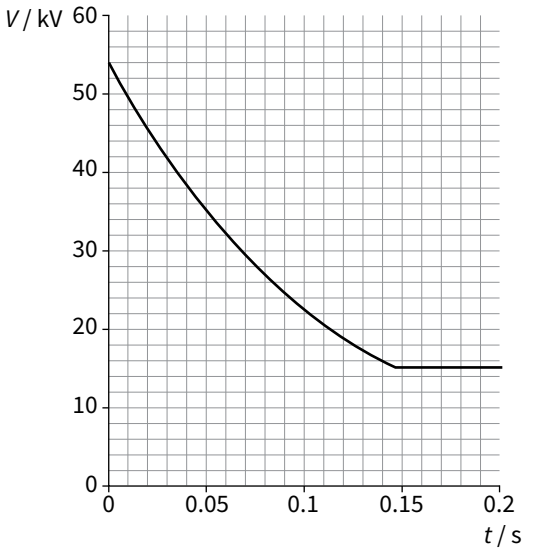


Figure 24.21 For End-of-chapter Question 12.

b Calculate the energy that is dissipated during the discharge. [4]

c Suggest why the discharge ceases while there is still some charge on the dome. [2]

13 **a** Show that the capacitance C of an isolated conducting sphere of radius r is given by the formula:

$C = 4\pi\varepsilon_0 r$ [2]

Figure 24.22 shows two identical conducting brass spheres of radius 10 cm mounted on insulating stands. Sphere A has a charge of $+5.0 \times 10^{-8}$ C and sphere B is uncharged.

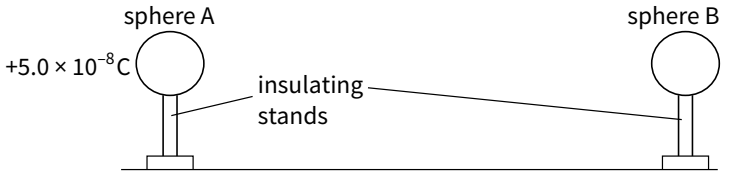


Figure 24.22 For End-of-chapter Question 13.

b **i** Calculate the potential at the surface of sphere A. [2]

ii Calculate the energy stored on sphere A. [2]

Sphere B is brought up to sphere A and is touched to it so that the charge is shared between the two spheres, before being taken back to its original position.

c **i** Calculate the energy stored on each sphere. [3]

ii Suggest why there is a change in the total energy of the system. [1]

14 **a** Define the term **capacitance** of a capacitor. [2]

b Figure 24.23 shows a circuit which can be used to measure the capacitance of a capacitor.

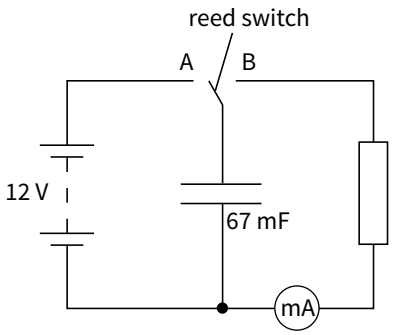


Figure 24.23 For End-of-chapter Question 14.

The reed switch vibrates back and forth at a frequency of 50 Hz. Each time it makes contact with A the capacitor is charged by the battery so that there is a p.d. of 12 V across it. Each time it makes contact with B it is fully discharged through the resistor.

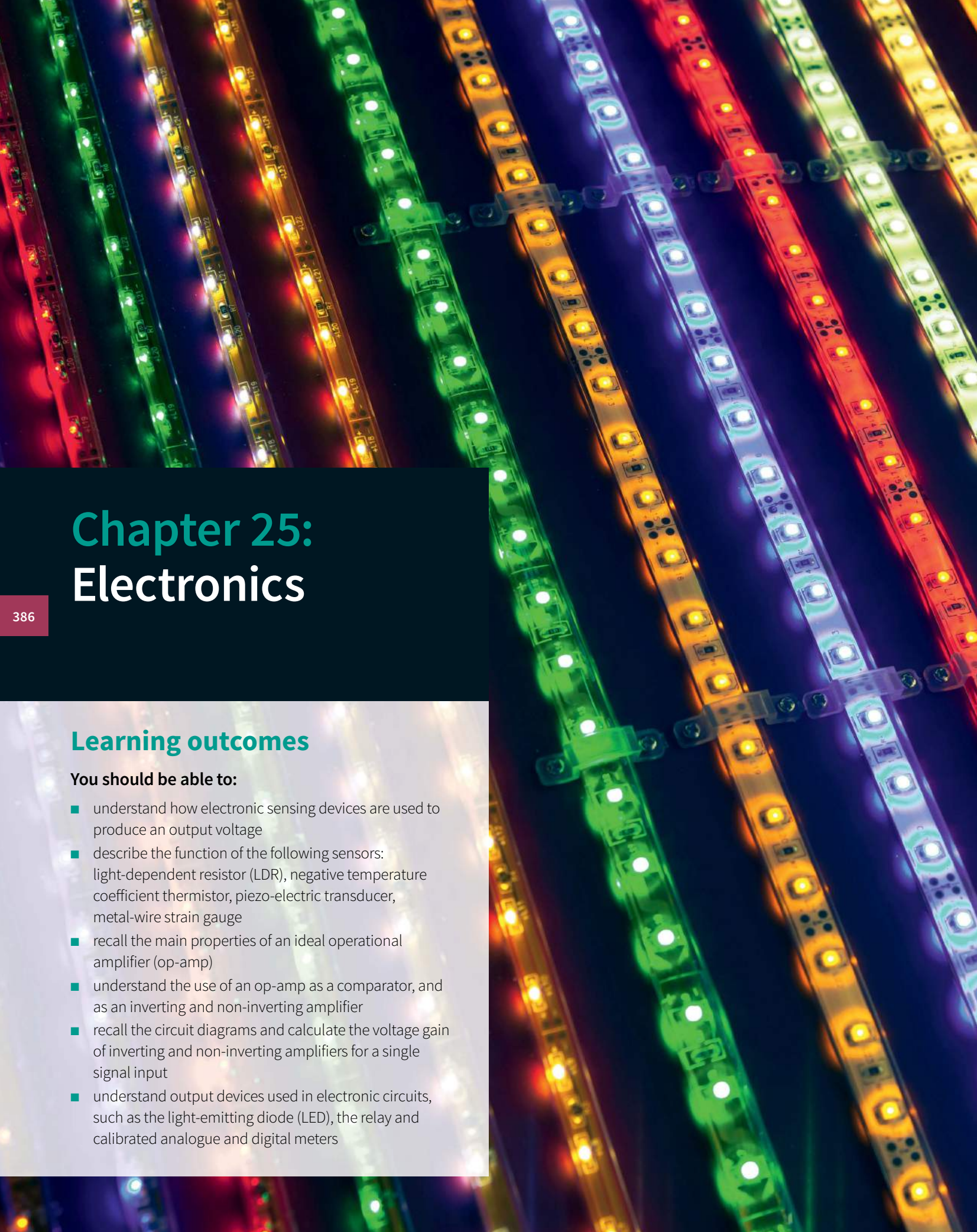
i Calculate the charge that is stored on the capacitor when there is a p.d. of 12 V across it. [2]

ii Calculate the average current in the resistor. [2]

iii Calculate the average power dissipated in the resistor. [3]

c A second capacitor of the same value is connected in series with the first capacitor.

Discuss the effect on both the current recorded and the power dissipated in the resistor. [4]



Chapter 25: Electronics

Learning outcomes

You should be able to:

- understand how electronic sensing devices are used to produce an output voltage
- describe the function of the following sensors: light-dependent resistor (LDR), negative temperature coefficient thermistor, piezo-electric transducer, metal-wire strain gauge
- recall the main properties of an ideal operational amplifier (op-amp)
- understand the use of an op-amp as a comparator, and as an inverting and non-inverting amplifier
- recall the circuit diagrams and calculate the voltage gain of inverting and non-inverting amplifiers for a single signal input
- understand output devices used in electronic circuits, such as the light-emitting diode (LED), the relay and calibrated analogue and digital meters

Electronic devices

Electronic devices are widespread in the modern world. From birth to death, electronic sensors watch over and control many of our environments. The chances of survival of the newborn baby in Figure 25.1 are improved as a sensor inside the incubator measures the temperature of the air. The output of the sensor operates heating and cooling devices which keep the temperature as constant as possible. Other sensors sound alarms if the baby's breathing or heart rate falls to dangerous levels.

In this chapter, we will explore how some common electronic sensors work. You will use what you have learned about electricity and electric circuits in Chapters 8–12.



Figure 25.1 A newborn baby in an incubator. Sensors inside the incubator monitor and control the temperature, oxygen and humidity levels.

Components of an electronic sensing system

In its basic form, an electronic sensor may be represented as a sensing device, a processor that provides an output voltage, and an output device (Figure 25.2).



Figure 25.2 Block diagram of an electronic sensor.

The sensing device is an electronic component with a property that changes when a physical quantity such as temperature or light intensity alters. We have already looked at one such component, the thermistor (Chapter 11), with a resistance that changes when temperature changes. A similar component, called a light-dependent resistor (LDR), has a resistance that changes when the light level alters. This change in resistance causes the processor to produce an output voltage that drives the output device. In this way the change in a physical quantity such as resistance can be used to trigger the output device. When the air in the incubator in Figure 25.1 becomes cold, the change in resistance of the sensor causes the processor to switch on the output device, a heater, to keep the baby warm. When it becomes too hot, the same sensing device and processor will switch the heater off.

It is important to note that we are not concerned with all the details of how a particular sensing device works or what there is inside the processor. We are only interested in knowing the output for various inputs and how the sensor can be used in a range of useful applications. Sensors make use of many basic ideas about electric circuits: resistance and current, for example. A processor may contain many transistors, but it is not necessary to understand how a transistor works to appreciate the function of a processor. Some detail will be given, but if you wish to delve further into how these devices work, there are many useful websites that will help you to understand how physicists have used their knowledge and inspiration to good effect.

The sensing device is sometimes called a **transducer**. A transducer changes energy from one form into another. A microphone is an obvious example, as it changes sound into electrical energy. However, even the thermistor used as the sensing device in the incubator in Figure 25.1 can be thought of as a transducer; a change in the internal energy of the air alters the electrical energy in the thermistor circuit.

Piezo-electric transducers

Some crystals such as quartz produce an electric field when a force is applied and the shape of the crystal changes. This is known as the piezo-electric effect. A piezo-electric crystal consists of positive and negative ions in a regular arrangement. When the crystal is stressed, a small voltage is produced between the faces of the crystal. The crystal acts as a **transducer** since energy is changed

from one form to another. There is more about piezo-electric transducers in Chapter 32, where their use in ultrasound scanning is described.

For use in a microphone, the piezo-electric crystal is made into a thin sheet with metal connections on opposite sides. When a sound wave hits one side of the sheet, the compressions and rarefactions cause the pressure to increase and decrease. The crystal changes shape in response to these pressure changes and a small voltage is created across the connections. Figure 25.3 shows the symbol for a microphone.

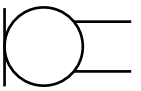


Figure 25.3 The symbol for a microphone.

Acoustic guitars and other instruments often use a piezo-electric transducer to produce an electrical output. The microphone is stuck to the body of the guitar and the electrical output can be amplified and played back through loudspeakers.

The light-dependent resistor (LDR)

A **light-dependent resistor (LDR)** is made of a high-resistance semiconductor. If light falling on the LDR is of high enough frequency, photons are absorbed by the semiconductor. As some photons are absorbed, electrons are released from atoms in the semiconductor. The resulting free electrons conduct electricity and the resistance of the semiconductor is reduced.

The graph in Figure 25.4 shows the variation of the resistance of a typical LDR with light intensity. Only a narrow range of light intensity, measured in lux, is shown. A typical LDR will have a resistance of a few hundred ohms in sunlight, but in the dark its resistance will be millions of ohms.

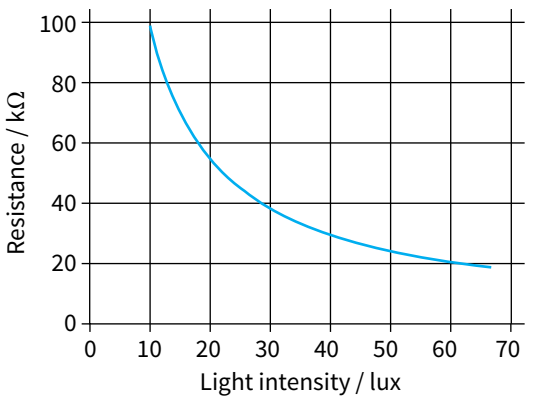


Figure 25.4 Resistance plotted against light intensity for an LDR.

How is an LDR used as a sensing device? A voltage is needed to drive the output device, yet the LDR only produces a change in resistance. The sensor must use this change in resistance to generate the change in voltage. The solution is to place the LDR in series with a fixed resistor, as shown in Figure 25.5.

The voltage of the supply is shared between the two resistors in proportion to their resistance, so as the light level changes and the LDR's resistance changes, so does the voltage across each of the resistors. The two resistors form a potential divider, one of the practical circuits discussed in Chapter 12.

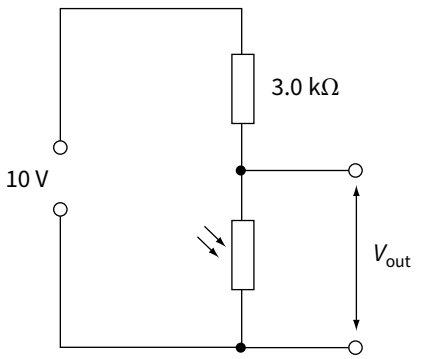


Figure 25.5 An LDR used in a sensor.

WORKED EXAMPLE

1 Using the graph in Figure 25.4, calculate V_{out} in Figure 25.5 when the light intensity is 60 lux.

Step 1 Find the resistance of the LDR at 60 lux.

$R_{LDR} = 20\,k\Omega$

Step 2 Divide the total voltage of 10 V in the ratio 3:20. The total number of parts is 23 so:

$$V_{out} = \frac{20}{23} \times 10 = 8.70\,V$$

Hint: *The answer on your calculator might be 8.69565. When you give your answer to 3 significant figures, do not write 8.69 – you must round correctly.*

QUESTIONS

1 What is the voltage across the 3.0 kΩ resistor in Figure 25.5 when the light intensity is 10 lux?

2 The circuit shown in Figure 25.5 produces a decreasing output voltage when the light intensity increases. How can the circuit be altered to produce an increasing output voltage as the light intensity increases?

The thermistor

The thermistors that we deal with are known as **negative temperature coefficient** thermistors. This means that when the temperature rises, the resistance of the thermistor falls. This happens because the thermistor is made from a semiconductor material. One property of a semiconductor is that when the temperature rises the number of free electrons increases, and thus the resistance falls.

Figure 25.6 shows a graph of the resistance of a thermistor and the resistance of a metal wire plotted against temperature. You can see that the resistance of a metal wire increases with increase in temperature. A metal wire is not a negative temperature device, but it could be used as a sensing device. A thermistor is more useful than a metal wire because there is a much larger change in resistance with change in temperature. However, the change in resistance of a thermistor is not linear with temperature; indeed, it is likely to be an exponential decrease. This means that any device used to measure temperature electronically must be calibrated to take into account the resistance–temperature graph. The scale on an ordinary laboratory thermometer between 0 °C and 100 °C is divided up into 100 equal parts, each of which represents 1 °C. If the resistance of a thermistor were divided like this, the scale would be incorrect.

The thermistor can be used as a sensing device in the same way as an LDR. Instead of sensing a change in light level, it senses a change in temperature.

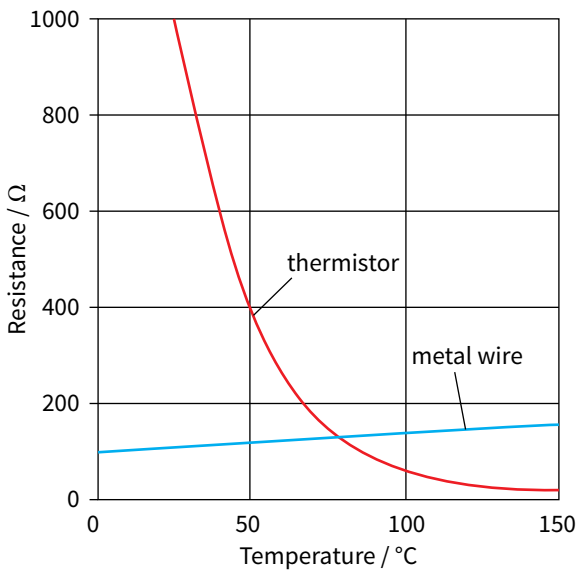


Figure 25.6 Variation of resistance with temperature.

QUESTIONS

3 Explain how a thermistor can be used as a transducer.

4 State **two** similarities between an LDR and a thermistor.

5 Design a circuit using the thermistor in Figure 25.6 that uses a cell of 10 V and produces an output voltage of 5 V at 50 °C. Explain whether the voltage output of your circuit increases or decreases as the temperature rises.

The metal-wire strain gauge

A **strain gauge** takes advantage of the change in resistance of a metal wire as its length and cross-sectional area change. When stretched, a metal wire becomes narrower and longer; both these changes increase the electrical resistance. When compressed, a metal wire becomes shorter and wider; as long as it does not buckle, these changes decrease its electrical resistance.

Figure 25.7 shows the structure of a strain gauge. A thin metal wire is placed between thin sheets of plastic. The metal wire zigzags up and down its plastic base so that the length of wire used is longer than the actual strain gauge.

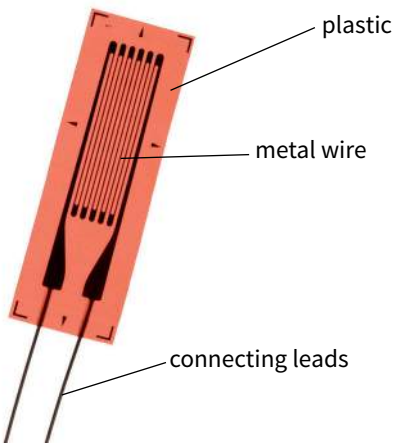


Figure 25.7 A metal-wire strain gauge.

Strain gauges are used in many situations. For example, an engineer may find a crack in a wall and want to know whether the crack is growing bigger. By sticking a strain gauge over the crack the engineer can measure the resistance many days or even years later and see if there has been any movement.

A gauge used in a dynamic experiment may measure the stress changes in an aircraft wing. A processor is used to produce a voltage output and show changes in length as

they happen. This provides important information when an aircraft is tested.

In Chapter 11, we saw that a wire of length L, cross-sectional area A and resistivity ρ has a resistance R given by:

$$R = \frac{\rho L}{A}$$

What happens to the resistance when the wire is stretched depends on the construction of the strain gauge. The simplest approximation is to assume that the cross-sectional area of the wire remains unchanged.

If the wire increases in length by a small amount δL and the cross-sectional area A is unchanged, then the resistance of the wire increases by δR, where:

$$R + \delta R = \frac{\rho(L + \delta L)}{A}$$

Subtracting the two equations gives:

$$\delta R = \frac{\rho \delta L}{A}$$

If the expression is divided by $R = \frac{\rho L}{A}$, we get:

$$\frac{\delta R}{R} = \frac{\delta L}{L}$$

When a wire is stretched not only does its length increase; at the same time, its area decreases (because its volume remains roughly constant). This also increases the wire's resistance. The increase in resistance can be shown to be the same as that caused by the increase in length so that, for small changes:

$$\delta R = \frac{2\rho \delta L}{A} \quad \text{or} \quad \frac{\delta R}{R} = \frac{2\delta L}{L}$$

In either case, the change in resistance is directly proportional to the increase in length, $\delta R \propto \delta L$.

The change in resistance is likely to be very small but with a suitable series resistor and a cell used as a sensor, the change in length can produce a measurable change in output voltage.

QUESTION

6 Using the data given in Worked example 2:

- **a** calculate the increase in length of the wire when the output voltage is 5.1 V
- **b** calculate the strain in the wire when the output voltage is 5.1 V
- **c** calculate the output voltage when the wire is stretched by 0.05 cm.

WORKED EXAMPLE

2 The wire in a strain gauge when unstretched has length 10.00 cm and resistance 120.0 Ω. When the wire is stretched by 0.10 cm, the resistance becomes 122.4 Ω. The strain gauge is connected in the circuit shown in Figure 25.8. What is the change in length of the wire in the strain gauge when the output voltage is 5.06 V?

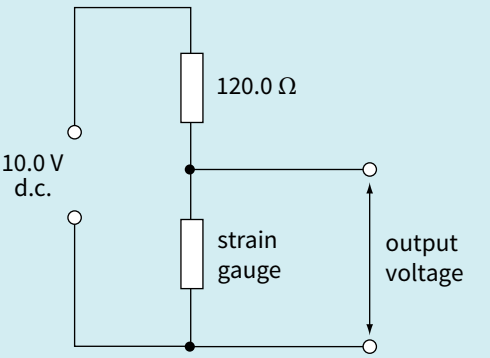


Figure 25.8 Circuit for Worked example 2.

Step 1 Determine the resistance of the strain gauge R_s by considering the two resistors as a potential divider. The 10 V supply is split in the ratio 5.06 V : 4.94 V and so:

$$R_s = \frac{5.06}{4.94} \times 120 = 122.9\ \Omega$$

Alternatively, apply the potential divider formula

$$V_{out} = \left(\frac{R_2}{R_1 \times R_2}\right) \times V_{in}$$

$$4.94 = \left(\frac{120}{120 + R_s}\right) \times 10.0$$

It is easier to use 4.94 V, the voltage across the 120 Ω resistor. This gives a value for R_s of 122.9 Ω.

Step 2 The resistance of the wire has increased by 2.9 Ω. Since $\delta R \propto \delta L$, and an increase in length of 0.10 cm increases the resistance by 2.4 Ω, the length has increased by:

$$\frac{2.9}{2.4} \times 0.10 = 0.12\ \text{cm}$$

If the strain gauge is glued to a metal support, the strain in the metal support is the same as the strain in the gauge, in this case $\frac{0.12}{10} = 1.2\%$. Thus the strain gauge can easily measure strain.

BOX 25.1: Potential dividers in use

Potential divider circuits are often used in electronic circuits. They are useful when a sensor is connected to a processing circuit. Suitable sensors include thermistors and light-dependent resistors (Figure 25.9). These can be used as sensors because:

- the resistance of a negative temperature coefficient (NTC) thermistor **decreases** as its temperature **increases**
- the resistance of a light-dependent resistor (LDR) **decreases** as the incident intensity of light **increases**.

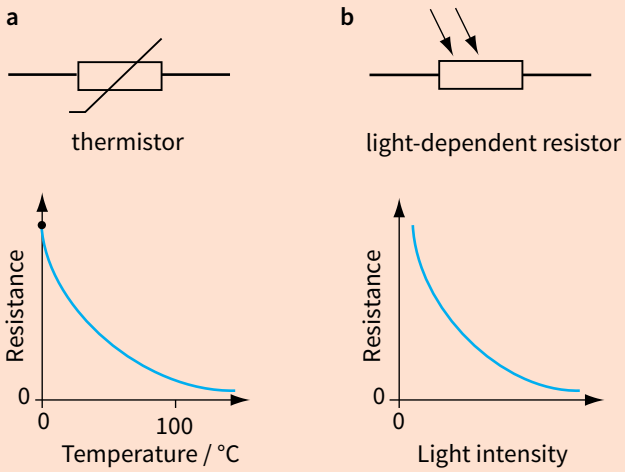


Figure 25.9 Two components with variable resistances: **a** the thermistor's resistance changes with temperature; **b** the light-dependent resistor's resistance depends on the intensity of light.

This means that a thermistor can be used in a potential divider circuit to provide an output voltage V_{out} which depends on the temperature. A light-dependent resistor can be used in a potential divider circuit to provide an output voltage V_{out} which depends on the intensity of light.

Figure 25.10 shows how a sensor can be used in a potential divider circuit. Here, a thermistor is being used to detect temperature, perhaps the temperature of a fish tank. If the temperature rises, the resistance of the thermistor decreases and the output voltage V_{out} increases. If the output voltage V_{out} is across the thermistor, as shown in Figure 25.10b, it will decrease as the temperature rises. By changing the setting of the variable resistor R_2, you can control the range over which V_{out} varies. This would allow you to set the temperature at which a heater operates, for example.

When designing a practical circuit like this, it is necessary to know how the voltage output depends on the temperature. You can investigate the voltage against

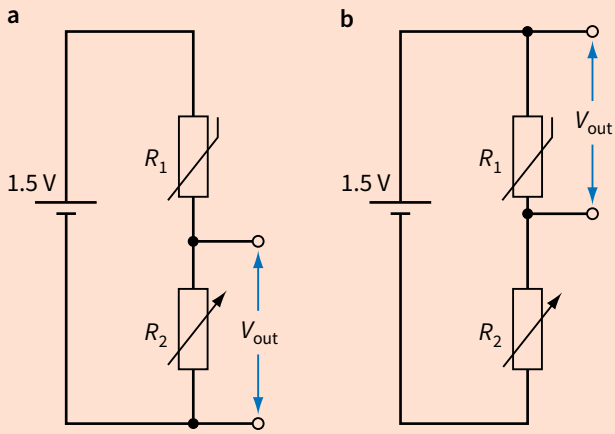


Figure 25.10 Using a thermistor in a potential divider circuit. The output voltage V_{out} may be **a** across the variable resistor, or **b** across the thermistor.

temperature characteristics of such a circuit using a datalogger (Figure 25.11). The temperature probe of the datalogger records the temperature of the water bath and the second input to the datalogger records the voltage output of the potential divider circuit.



Figure 25.11 Using a datalogger to investigate the characteristics of a thermistor in a potential divider circuit. During the experiment, the screen shows how temperature and output voltage change with time. After the experiment, the same data can be displayed as a graph of p.d. against temperature.

BOX 25.1: Potential dividers in use (continued)

The temperature can be raised rapidly by pouring amounts of water into the water bath. The datalogger then records both temperature and voltage and the computer gives a display of the voltage against temperature. Dataloggers are very good at processing the collected data.

Potential divider circuits are especially useful in circuits with very small currents but where voltages are important. Electronic devices such as transistors and integrated circuits draw only very small currents, so potential dividers are very useful where these devices are used. Where large currents are involved, because there will be some current through both R_1 and R_2, there will be wasted power in the resistors of the potential divider circuit.

QUESTIONS

7 An NTC thermistor is used in the circuit shown in Figure 25.12. The supply has an e.m.f. of 10 V and negligible internal resistance. The resistance of the thermistor changes from 20 kΩ at 20 °C to 100 Ω at 60 °C. Calculate the output voltage V_{out} at these two temperatures.

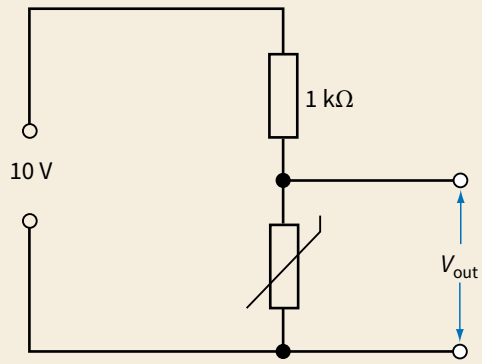


Figure 25.12 Thermistor used in a potential divider circuit. For Questions 7 and 8.

8 The thermistor in Figure 25.12 is replaced with a light-dependent resistor (LDR). Explain whether the output voltage V_{out} will increase or decrease when a bright light is shone on to the LDR.

9 The light-dependent resistor (LDR) in Figure 25.13 has a resistance of 300 Ω in full sunlight and 1 MΩ in darkness. What values will the output voltage V_{out} have in these two conditions?

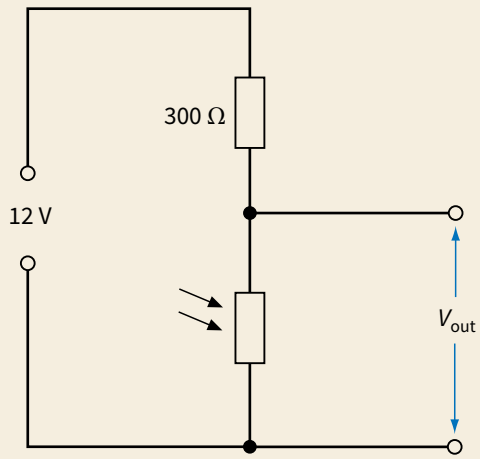


Figure 25.13 For Question 9.

10 A potential divider circuit is required which will give an output voltage that increases as the temperature increases. A thermistor is to be used whose resistance decreases as the temperature increases. Draw a suitable circuit for the potential divider, showing the connections for the output voltage.

The operational amplifier (op-amp)

Operational amplifiers (op-amps) are among the most widely used electronic devices today, being used in a vast array of consumer, industrial and scientific devices. The amplifier is the basic building block of many electronic systems. The electrical output from the musicians in the concert shown in Figure 25.14 must be amplified before it can be passed to the loudspeakers and turned into sound.

An amplifier produces an output with more power and usually more voltage than the input. A perfect amplifier should produce an exact copy of the input. In particular, the different frequencies produced by the musicians and their instruments must be amplified by the same amount. If, for example, high-frequency notes are made louder than low-frequency notes then the whole performance will be altered. You may have noticed that people sound different when talking on the telephone. This is often because some frequencies are not amplified as much as others.



Figure 25.14 A concert – the loud music has been greatly amplified.

The goal of producing an amplifier with a constant amplification or **gain** might seem simple, but it is hard to achieve. Unfortunately, electronic components, such as capacitors and transistors, amplify signals of different frequencies by different amounts.

One approach is to use an amplifier with a very high gain and then provide an external circuit which reduces the gain but ensures that the overall gain is the same for signals of a greater range of frequencies. Such a device is the **operational amplifier (op-amp)**.

Figure 25.15 shows an operational amplifier. Inside the plastic casing there are many transistors, resistors and other components. The op-amp has two inputs; one is marked (–) and is known as the **inverting input** and the other is marked (+) and is known as the **non-inverting input**. The one function of this device is to use the potential difference between the two inputs to produce as large an output voltage V_{out} as possible. V^- is the potential at the inverting input (–) and V^+ is the potential at the non-inverting input (+).

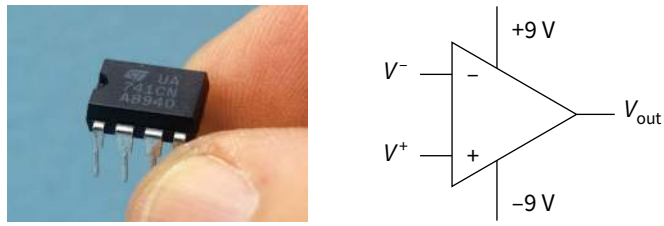


Figure 25.15 An operational amplifier and its symbol.

The open-loop voltage gain G_0 is given by:

$$G_0 = \frac{\text{output voltage}}{\text{input voltage}}$$

For the op-amp in Figure 25.15 the open-loop voltage gain is given by:

$$G_0 = \frac{V_{out}}{(V^+ - V^-)}$$

It is called the open-loop gain because there is no loop of resistors or other components linking the output back to the input – it is just the operational amplifier alone. (As we shall see on page 396, an external circuit or loop between the output and the input can be used to provide **negative feedback**, which reduces the gain.)

The op-amp is not like a transformer. A transformer's output power comes directly from from the input power. By contrast, an op-amp's output power is much greater than its input power. To achieve this, it needs two power supplies; these are shown in Figure 25.15 as the +9 V and –9 V connections to the op-amp. There is also a zero volt line, or earth. One power supply will be between the +9 V and the 0 V line and the other between the –9 V connection and the 0 V line.

The zero volt line, or earth, is very important as all voltages are measured relative to this potential. If the potential of the inverting input (–) is V^-, then the potential difference or voltage between that input and the zero volt line is also V^-.

The actual voltage used for the power supplies can vary in different circuits. The positive and negative supply voltages are of equal magnitude and may be written as $+V_s$ and $-V_s$. The connections of the power supplies to the op-amp are often left out of a circuit diagram for clarity but they are always there. They provide the power for the op-amp.
The largest voltage that an op-amp can produce is a value close to the supply voltage. The op-amp in Figure 25.15 can

only produce outputs between +9 V and −9 V. When the output voltage reaches either supply voltage, the highest or lowest value that it can achieve, the amplifier is said to be **saturated**.

The properties of an ideal op-amp

The ideal op-amp has the following properties.

Infinite open-loop voltage gain

This means that when an op-amp is used on its own, with no feedback loop, then a small input signal will be amplified to an 'infinite' output signal. Clearly this is not physically possible (the output cannot exceed the supply voltage) and at its maximum output the amplifier is saturated with output value $+V_s$ or $-V_s$. However, when a feedback loop is applied, the overall gain of the circuit is reduced to a realistic value. Infinite open-loop voltage gain means that signals of a wide range of frequencies have equal gain before the feedback is applied. An actual op-amp may have an open-loop gain of 10^5 but this can be much higher.

Infinite input resistance (or impedance)

The input to an op-amp is a voltage. If, for example, a piezo-electric microphone is connected to the op-amp, then the microphone is acting as the voltage supply. It acts just like an electrical battery but the voltage it produces changes with time. Any voltage supply has an internal resistance. You may remember that one of the effects of this is to reduce the terminal p.d. when a current is supplied. The infinite input resistance of an ideal op-amp means that no current is drawn from the supply, there are no 'lost volts' and the input voltage to the op-amp is as large as possible. The resistance for an alternating voltage is known as impedance, so the ideal op-amp has infinite impedance and no current passes into the input terminals. The input impedance of an actual op-amp may be as high as $10^{12}\,\Omega$, but $10^6\,\Omega$ is typical.

Zero output resistance (or impedance)

The output from an op-amp is a voltage. The op-amp is itself acting as a voltage supply to the next part of a circuit. An ideal op-amp has zero output resistance and so it acts just like an electrical battery with zero internal resistance. This means that there will be no 'lost volts' when current is supplied by the op-amp. An actual op-amp typically has an output resistance of around 75 Ω.

Infinite bandwidth

The **bandwidth** of an op-amp is the range of frequencies that are amplified by the same amount. An ideal op-amp will amplify signals of all frequencies and should therefore have an infinite bandwidth. However, in actual op-amps the bandwidth can be as low as a few kilohertz.

Infinite slew rate

An ideal op-amp should change the output instantaneously as the input is changed. The slew rate of the op-amp is the factor that affects this time delay. An infinite slew rate means there is no time delay.

Zero noise contribution

Any signal includes a small amount of noise. The ideal op-amp does not produce any noise itself, although it will amplify any noise that is present in its input.

The op-amp as a comparator

The op-amp shown in Figure 25.16 is connected to two power supplies. One battery of 9 V is connected between the zero volt line and the +9 V positive supply terminal of the op-amp and the other between the zero volt line and the negative power supply terminal. These batteries are not shown.

The output voltage is given by $V_{out} = G_0 \times (V^+ - V^-)$, where G_0 is the open-loop voltage gain. Notice that all of these voltages are measured with reference to the zero volt line, which is often connected to earth.

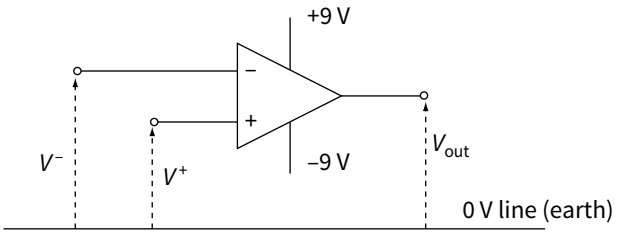


Figure 25.16 An op-amp used as a comparator.

Suppose that $G_0 = 10^5$ and that $V^+ = 0.15$ V and $V^- = 0.10$ V. The equation then gives:

$$V_{out} = 10^5 \times (0.15 - 0.10) = 5000\text{ V}$$

Clearly this is impossible. The op-amp is therefore saturated and V_{out} will be close to one of the power supply voltages, in this case +9 V because V^+ is bigger than V^-. In fact:

- If V^+ is slightly greater in magnitude than V^-, then V_{out} will have a magnitude equal to the positive power supply voltage.
- If V^+ is slightly smaller in magnitude than V^-, then V_{out} will have a magnitude equal to the negative power supply voltage.

The op-amp is serving as a comparator, comparing V^+ and V^-. If these two voltages are slightly different, then the output voltage tells us which one is larger. It is unlikely that the voltages will be exactly the same.

A comparator circuit can be used to compare two temperatures or two light levels. The circuit shown in Figure 25.17 is used to give a warning when the temperature sensed by thermistor T becomes smaller than a certain value.

To understand the action of this circuit, you should notice that the positive power supply to the op-amp is also used to supply voltage and current to the thermistor T and a 10 kΩ resistor connected as a potential divider. There is also another potential divider circuit connected to the inverting input (–) of the op-amp. (Note that there is no connection where the two connecting wires are shown crossing.)

Worked example 3 explains how this circuit operates.

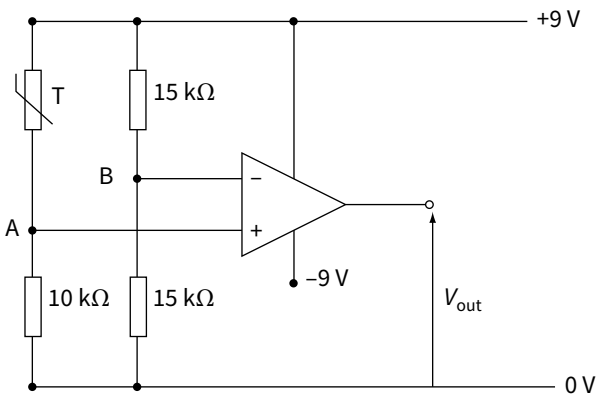


Figure 25.17 An op-amp used as a comparator to monitor temperature.

QUESTIONS

11 With an open-loop voltage gain $G_0 = 10^5$ and power supply voltages of +9 V and −9 V, what is the smallest difference between V^+ and V^- for which the op-amp is not saturated?

12 What happens in the circuit shown in Figure 25.17 if the supply voltages are changed to +15 V and −15 V? Is the temperature at which it switches over the same? Are the output voltages the same?

13 How can the circuit of Figure 25.17 be altered so that the output switches from −9 V when the temperature is hot to +9 V when the temperature is cold?

14 Calculate values of V^+, V^- and V_{out} in Figure 25.17 when the resistance of T is 3 kΩ.

WORKED EXAMPLE

3 For the circuit shown in Figure 25.17, the resistance of the thermistor T is 8 kΩ at a temperature of 15 °C. What are V^- and V^+, the potentials at the inverting and non-inverting inputs? And what happens when the temperature falls so that the resistance of T rises above 10 kΩ?

Step 1 V^- and V^+ can be found by using the potential divider formula to find the potentials at points A and B. The potential at A is the p.d. across the 10 kΩ resistor. So:

$$\text{potential at A} = 9 \times \frac{10}{18} = 5.0\,\text{V}$$

The potential at B is easier to find, as the two 15 kΩ resistors share the 9 V equally.

$$\text{potential at B} = \frac{9}{2} = 4.5\,\text{V}$$

The op-amp acts as a comparator and, since V^+ is larger than V^-, the output will be the highest voltage that the op-amp can produce, in this case +9 V.

Step 2 The thermistor T is a negative temperature coefficient thermistor and so its resistance rises sharply and eventually becomes larger than 10 kΩ. Suppose it becomes 12 kΩ. Then:

$$\text{potential at A} = 9 \times \frac{10}{22} = 4.1\,\text{V}$$

Now V^+ is smaller than V^- and the op-amp output voltage is the lowest it can provide, near the negative supply voltage, in this case −9 V.

This switch from +9 V to −9 V is quite sudden because of the large open-loop voltage gain. The value of the temperature when the output voltage switches from +9 V to −9 V can be altered by adjusting the resistance of the resistor in series with the thermistor.

Negative feedback

What happens when the op-amp in Figure 25.17 is connected to a heater which warms the air around the thermistor? The op-amp senses when the room is cold and switches on the heater. The heater then warms the room and this information is fed back to the thermistor, which then senses that the room is now warm enough and switches off the heater. This process is known as **feedback** and keeps the room at a reasonably constant temperature.

The effect is shown in Figure 25.18.

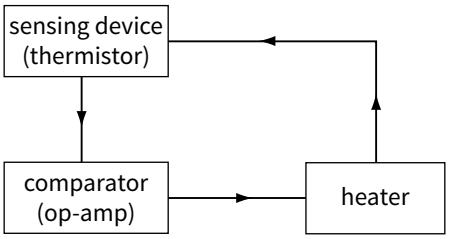


Figure 25.18 Feedback keeping temperature constant.

There are many examples of feedback. Your body temperature is remarkably constant because of a number of feedback mechanisms. When your skin senses that you are cold, a signal is sent via the nerves to your brain which closes down some of the blood vessels in your skin to reduce energy loss. If you are very cold the brain makes muscles contract and expand uncontrollably as you shiver. This will make the core of your body warm again and the effect is fed back to the brain, which switches off the shivering.

Electrical feedback is also very important. Consider the circuit shown In Figure 25.19. In this circuit V_{in}, the input to the op-amp, is connected to the non-inverting input (+). The whole of the voltage output V_{out} is fed back to the inverting input (–) of the op-amp. How does the op-amp behave with this feedback loop?

The potential V^- at the inverting input (–) is always the same as V_{out} because they are connected by the feedback loop. We will assume that the open-loop voltage gain is

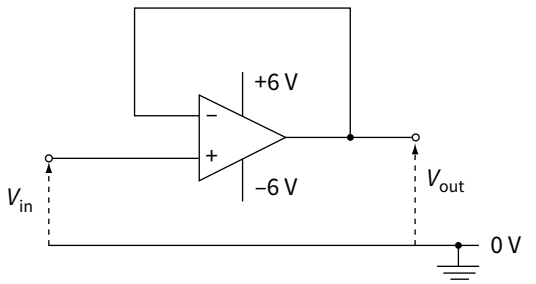


Figure 25.19 An op-amp with the output connected to the inverting input.

infinite and that the op-amp is not saturated. In this case V^- must be equal in value to V^+, and this is an important principle.

Suppose the system starts with both V_{in} and V_{out} as 0 V and then the input voltage V_{in} changes to +0.1 V. The op-amp multiplies the potential difference between V^+ and V^- by the open-loop voltage gain to produce V_{out}. The gain is very high so V_{out} increases quickly. What does V_{out} become? As V_{out} is 'fed back' to V^-, the value of V^- increases and this reduces the difference between V^- and V^+. Very quickly this difference becomes zero again and $V_{out} = V_{in}$.

We know that $V_{out} = G_0 \times (V^+ - V^-)$, where G_0 is the open-loop voltage gain. Since $V_{out} = V^-$ and $V_{in} = V^+$ we have:

$$V_{out} = G_0(V_{in} - V_{out})$$

$$V_{out}(1 + G_0) = G_0 V_{in}$$

The closed-loop gain G is given by:

$$G = \frac{V_{out}}{V_{in}} = \frac{G_0}{(1 + G_0)}$$

Because G_0 is very high, about 10^5, there is little difference between G_0 and $(G_0 + 1)$, so the closed-loop gain is very nearly 1. Since the input voltage was +0.1 V the output voltage is also +0.1 V. This analysis is true as long as the output voltage is smaller than the supply voltage, in this case as long as V_{out} is between −6 V and +6 V.

This may seem strange. We have taken an op-amp with a gain of 10^5 and turned it into a device with a gain of 1, so the input voltage and output voltage are equal. This is hardly an amplifier, but it is useful. The op-amp draws very little current from the input, yet it can supply a reasonable current from its output. This circuit is often used as a buffer between electronic circuits. If something happens to one circuit it does not affect the other circuit.

A piezo-electric microphone has a high internal resistance and cannot supply much current. If it is connected to the input of the op-amp in Figure 25.19, the same voltage is output but the current can be larger.

Another advantage is that it does not matter whether the frequency of the input signal is high or low; the gain is the same. So the output signal is exactly the same as the input signal and there is no distortion. This is only true when the open-loop voltage gain is high. At very high frequencies the open-loop voltage gain falls and eventually the closed-loop gain falls. The **bandwidth**, the range of frequencies for which the gain is constant, is increased by using negative feedback.

To summarise, the benefits of using negative feedback to reduce the gain of an op-amp are:

- less distortion
- increased bandwidth
- the gain is more stable and not affected by changes in temperature, etc.
- the output resistance (impedance) can be low and the input resistance (impedance) high.

The inverting amplifier

The **inverting amplifier** shown in Figure 25.20 uses negative feedback, but not all of the output voltage is fed back to the inverting input (–). When an op-amp is connected as an inverting amplifier:

- the non-inverting input (+) is connected to the 0 V line
- part of the output voltage (or signal) is connected to the inverting input (–)
- the input voltage (or signal) is connected to the inverting input.

To understand how the inverting amplifier works, you need to understand the concept of the **virtual earth approximation**. In this approximation the potential at the inverting input (–) is very close to 0 V. Why is this true? There are two steps in the argument:

- The op-amp multiplies the difference in potential between the inverting and non-inverting inputs, V^- and V^+, to produce the output voltage V_{out}. Because the open-loop voltage gain is very high, the difference between V^- and V^+ must be almost zero.
- The non-inverting input (+) is connected to the zero volt line so $V^+ = 0$. Thus V^- must be close to zero and the inverting input (–) is almost at earth potential.

Point P is known as a **virtual earth**. It cannot actually be 0 V but it is very close to 0 V. This approximation is true as long as the op-amp is not saturated, and for frequencies where the open-loop voltage gain is high.

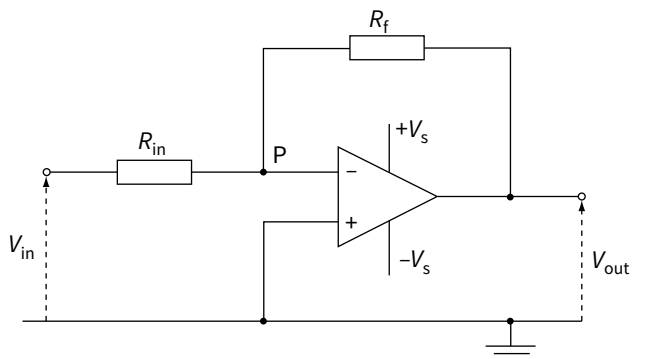


Figure 25.20 An inverting amplifier.

The virtual earth approximation can be used to find an expression for the gain of an inverting amplifier, as follows. If the current in the input resistor R_{in} is I_{in} and the current in the feedback resistor R_f is I_f, then, because point P is at 0 V:

$$I_{in} = \frac{V_{in}}{R_{in}}$$

and

$$I_f = \frac{V_{out}}{R_f}$$

The input resistance of the op-amp is very high and so virtually no current enters or leaves the inverting input (–) of the op-amp. This means that I_{in} and I_f must be equal in size.

If V_{in} is a positive potential, then the current in the two resistors flows from left to right. V_{out} will be negative because the current flows from P, which is at 0 V, to the output connection, which must have a lower voltage than 0 V. Thus:

$$I_f = -I_{in}$$

and

$$\frac{V_{out}}{R_f} = -\frac{V_{in}}{R_{in}}$$

The gain of the inverting amplifier is thus given by:

$$G = \frac{V_{out}}{V_{in}} = -\frac{R_f}{R_{in}}$$

The negative sign shows that when the input voltage is positive the output voltage is negative and when the input is positive the output is negative. If the input voltage is alternating then there will be a phase difference of 180° or π rad between the input and the output voltages.

QUESTIONS

15 Draw the circuit diagram of an inverting amplifier with a gain of –100 and an input resistor R_{in} of 10 kΩ. Include the value of the feedback resistor R_f.

16 The supply voltage to an op-amp is ±15 V. The op-amp is connected as an inverting amplifier of gain –20. Calculate the output voltage for the following input voltages:

a +20 mV

b –400 mV

c +1.0 V

WORKED EXAMPLE

4 In the circuit shown in Figure 25.20, $R_{in} = 4\,k\Omega$ and $R_f = 20\,k\Omega$. Calculate:

a the output voltage when the input voltage is 0.5 V

b the maximum input voltage if the supply voltage is ±12 V

c the input resistance of the amplifier.

The solutions can be found using the formula for gain G.

Step 1 Calculate the gain $G = -\frac{R_f}{R_{in}} = -\frac{20}{4} = -5$

Step 2 Calculate V_{out} = gain × V_{in} = −5 × 0.5 = −2.5 V

Step 3 The op-amp becomes saturated when $V_{out} = -12$ V, so:

$$V_{in} = \frac{V_{out}}{G} = \frac{12}{5} = 2.4\,V$$

Step 4 Point P is at 0 V, so the input resistance is just caused by $R_{in} = 4\,k\Omega$.

The non-inverting amplifier

Figure 25.21 shows the circuit for a **non-inverting amplifier**. The input voltage is applied to the non-inverting input; part of the output voltage is fed back to the inverting input.

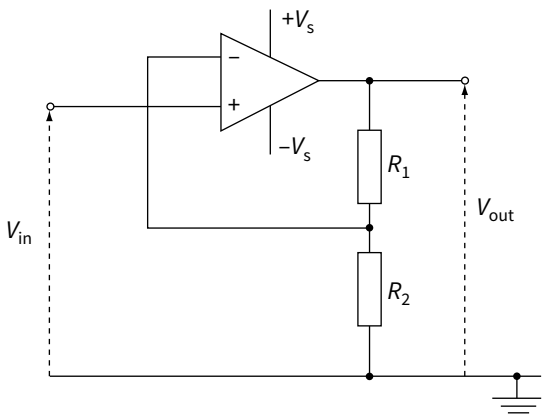


Figure 25.21 A non-inverting amplifier.

As long as the op-amp is not saturated, the potential difference between the inverting (–) and non-inverting inputs (+) is almost zero. So $V^- = V^+$.

Since the non-inverting input (+) is connected to the input voltage, $V^+ = V_{in}$. Thus $V^- = V_{in}$.

The two resistors R_1 and R_2 form a potential divider. The total voltage across R_1 and R_2 is V_{out} and the voltage across R_2 alone is V_{in}.

The current in the two resistors can be written as:

$$\frac{V_{out}}{(R_1 + R_2)} = \frac{V_{in}}{R_2}$$

The gain is calculated from:

$$G = \frac{V_{out}}{V_{in}} = \frac{(R_1 + R_2)}{R_2} = 1 + \left(\frac{R_1}{R_2}\right)$$

Thus for a non-inverting amplifier the gain is given by:

$$G = 1 + \left(\frac{R_1}{R_2}\right)$$

For a non-inverting amplifier, the output is in phase with the input. When the input voltage is positive, so is the output voltage.

QUESTIONS

17 The circuit shown in Figure 25.19 is a non-inverting amplifier. What are the values of R_1 and R_2? Show that the gain of the amplifier is 1.

18 Draw the circuit of a non-inverting amplifier with a gain of 10. Explain whether the input resistance is very high or very low.

19 Suppose that the non-inverting amplifier in Figure 25.21 has $R_1 = 50\,k\Omega$, $R_2 = 5\,k\Omega$ and $V_s = 10$ V.

a Calculate the gain.

b Calculate the output voltage when the input voltage is:

i −0.10 V **ii** +1.0 V.

Output devices

The output voltage of an op-amp may be used to operate a device according to the changing input voltage. We will now look at three types of output device.

The relay

Although the output resistance of an op-amp is low, a typical op-amp can only provide a maximum output current of 25 mA. The maximum voltage output from an op-amp is also limited to the supply voltage, typically 15 V. To switch on larger currents and larger voltages the op-amp is connected to a relay.

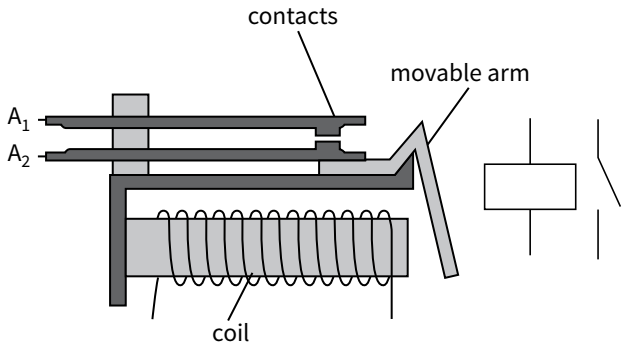


Figure 25.22 A relay and its circuit symbol.

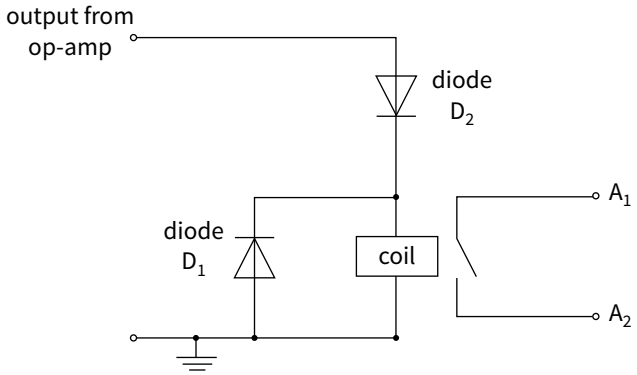


Figure 25.23 The output of an op-amp connected to a relay.

The **relay** is just an electromagnetic switch operated by a small current in the coil. Notice that there are two circuits, one to the coil and one involving the switch contacts A_1 and A_2. When a small current passes through the coil of the relay in Figure 25.22, the iron core attracts a movable arm and the contacts connected to A_1 and A_2 close, completing the second circuit.

The coil of the relay is the part connected to the output of an op-amp. The op-amp can easily provide the small current required for the coil. When the contacts A_1 and A_2 close they can switch large voltages or currents in another circuit.

There is, however, a problem using a relay connected to an op-amp. The current from the op-amp causes the coil to act as an electromagnet and creates a magnetic field. When this current is turned off there is a very rapid fall in the magnetic flux within the coil and a large e.m.f. is induced across the terminals of the coil, large enough to damage the op-amp. Switching off a relay can damage an op-amp.

To avoid this damage, a reverse-biased diode is placed across the relay coil. This is shown in Figure 25.23, where D_1 is the reverse-biased diode. When the op-amp switches off, the induced voltage in the coil causes the bottom of the coil to be more positive than the top of the coil. Diode D_1 is able to pass current round the coil without any damage to the op-amp.

The output of the op-amp can be negative as well as positive. Without diode D_2, the relay contacts close whether there is a negative output or a positive output from the op-amp because it does not matter which direction the current flows in the coil in Figure 25.22.

Diode D_2 ensures that current can only flow from the op-amp when the output of the op-amp is positive. Thus the relay contacts are closed only when the output of the op-amp is positive.

Since diode D_1 is reverse-biased, no current from the op-amp flows through D_1.

The light-emitting diode (LED)

The light-emitting diode is a very convenient device to attach to the output of an op-amp. LEDs come in several different colours and only require a current of about 20 mA to produce a reasonable light output. When placed on the output of the op-amp they readily show the state of that output, whether it is positive, negative or zero.

In practice, an LED cannot be placed directly between the output of an op-amp and the zero-volt line. The current–voltage characteristic for an LED is shown in Chapter 11 (page 160). The LED starts to conduct when the voltage across it is greater than about 2 V, although this value depends on the type of LED used. Once the output of an op-amp is much bigger than 2 V, the current in the LED and the op-amp will be very high and will damage both of them. A series resistor is required to limit the current, as shown in Figure 25.24.

The value of the resistance of the series resistor R can be calculated. If the current is to be 20 mA and the maximum voltage output from the op-amp is 12 V, then there will be just over 2 V across the LED and $12 - 2 = 10$ V across the series resistor R. The series resistor required is:

$$R = \frac{10}{0.02} = 500\,\Omega$$

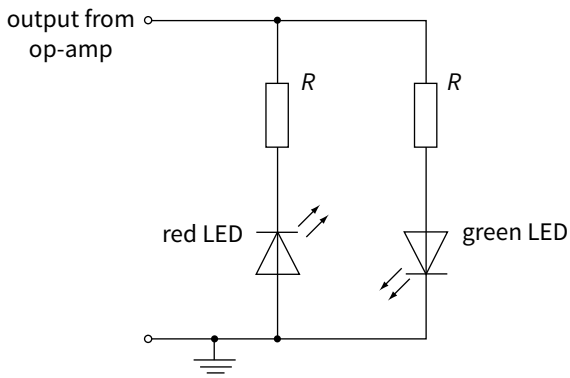


Figure 25.24 LEDs connected to the output of an op-amp.

In Figure 25.24, when the output of the op-amp is sufficiently positive relative to earth the green LED will light. When the output is negative relative to earth, the red LED will light.

QUESTIONS

20 Draw a diagram of a relay connected to the output of an op-amp. The relay should switch on a separate circuit containing mains voltage and a heater.

21 Why is a reverse-biased diode required for a relay but not for an LED?

The calibrated meter

An op-amp may be monitoring a physical quantity such as temperature or light intensity. It is helpful to display the value of that physical quantity directly on a meter and not to have to measure an output voltage and then calculate the physical quantity each time.

The output voltage of the op-amp is unlikely to be proportional to the physical quantity being measured, for example temperature. The numbers on a voltmeter connected to the output of an op-amp cannot simply be changed to read values of the physical quantity, but the voltmeter can still be calibrated in terms of the physical quantity.

Suppose that temperature is being measured and a digital voltmeter is connected to the output. Calibration can be achieved by placing the temperature sensor and a thermometer in a water bath and recording the voltmeter reading and the temperature of the water bath at a number of different temperatures. A calibration graph is then drawn, as shown in Figure 25.25.

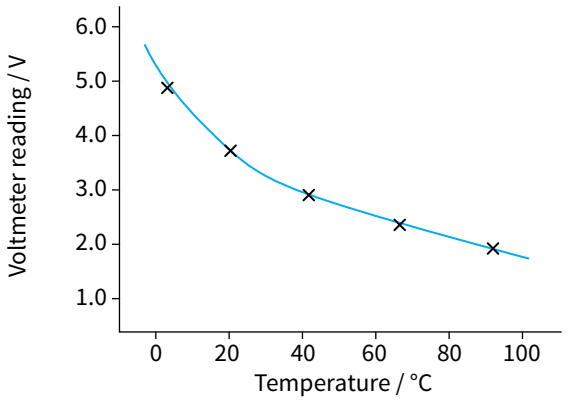


Figure 25.25 A calibration curve relates the output voltage of an op-amp to the variable it is being used to measure.

The calibration curve is used to change any other voltmeter reading into a value for the physical quantity, in this case temperature.

An analogue meter can be calibrated in the same way, but if the same meter is used all the time as the output, then the scale on the meter can be marked directly with values of the physical quantity. This means that the 'voltmeter' will measure temperature directly. Since the voltmeter reading is not proportional to the change in the physical quantity, the scale will not be linear. Figure 25.26 shows a meter with three scales. The bottom two scales are linear and might be the voltage recorded by a voltmeter. The top scale is non-linear and might be the value of the physical quantity being measured. Care is needed when reading a non-linear scale when the pointer is between the markings.

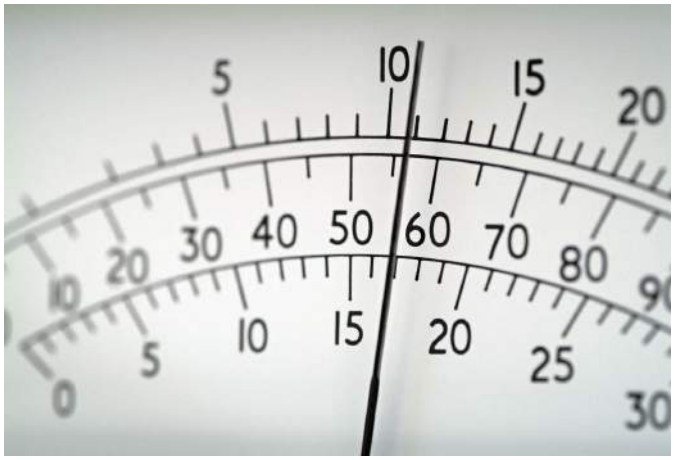


Figure 25.26 Linear and non-linear scales.

Summary

- A sensor contains a sensing device and a processor to produce an output voltage.
- The resistance of an LDR decreases as the level of the light intensity increases.
- The resistance of a negative temperature coefficient thermistor falls as temperature rises.
- Thermistors and light-dependent resistors can be used in potential divider circuits to provide output voltages that are dependent on temperature and light intensity, respectively.
- A piezo-electric transducer produces a variation in output voltage when a crystal is stressed, for example by pressure changes in sound.
- The change in resistance of a metal-wire strain gauge is proportional to extension.
- An operational amplifier used as a comparator produces either a high positive or a high negative voltage output, depending on which of two input voltage levels is larger.
- An operational amplifier can be used as an inverting or a non-inverting amplifier to amplify voltage signals.
- The gain of an inverting amplifier is $-\frac{R_f}{R_{in}}$
- The gain of a non-inverting amplifier is $1+\left(\frac{R_1}{R_2}\right)$.
- Feedback reduces gain but increases bandwidth and stability and reduces distortion.
- Output devices which may be connected to an op-amp include relays, LDRs and calibrated meters.

End-of-chapter questions

1 How can the circuit of Figure 25.17 be altered so that the op-amp switches from +9 V to −9 V when the temperature rises, instead of when it falls? Draw the circuit diagram. [1]

2 Draw the circuit diagram of a comparator that switches the output voltage from positive to negative when the light intensity falling on an LDR decreases. [3]

3 Explain what is meant by:

a **feedback** [1]

b the **virtual earth approximation**. [2]

4 In the inverting amplifier shown in Figure 25.20, $R_{in} = 10\,k\Omega$, $R_f = 200\,k\Omega$ and $V_s = 12\,V$. Calculate:

a the gain of the amplifier [2]

b the input voltage when the output voltage is 8.0 V [2]

c the maximum input voltage. [1]

5 **a** Explain how an inverting amplifier can have a gain of −0.5. [2]

b Explain why the gain of a non-inverting amplifier cannot be less than 1. [2]

6 The non-inverting amplifier in Figure 25.21 on page 398 has $R_1 = 15\,k\Omega$, $R_2 = 5\,k\Omega$ and $V_s = 10\,V$.
An a.c. signal of amplitude 0.20 V is applied to the input. Draw a single sketch graph which shows both the input and the output voltage over **two** cycles of the input on the same axes. [3]

7 Is the circuit shown in Figure 25.27 an inverting or non-inverting amplifier? Calculate its gain. [3]

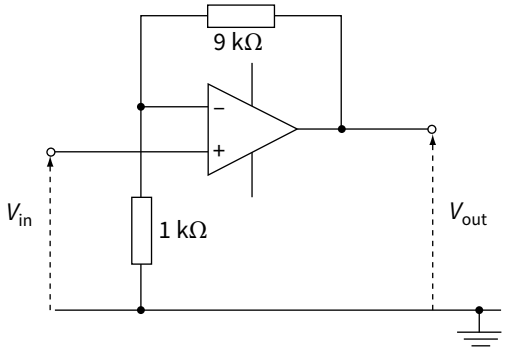


Figure 25.27 For End-of-chapter Question 7.

8 A light-dependent resistor is used with additional components to make a sensor. The voltage output increases as light intensity increases.

a Sketch the diagram of a suitable circuit. [2]

b Explain how your circuit works. [2]

c State which parts of your circuit are the sensor, the processor and the output device. [2]

9 A strain gauge is used to measure the extension of a metal bar. The gauge is stuck to the bar and the change in resistance of the wire in the gauge enables the extension to be determined. The strain gauge is connected in the circuit shown in Figure 25.28.

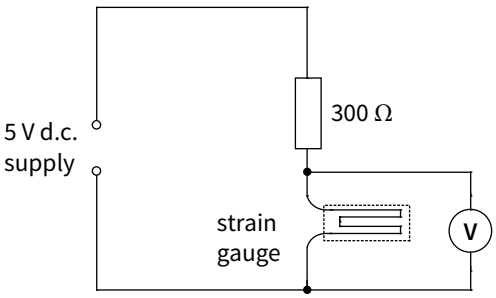


Figure 25.28 For End-of-chapter Question 9.

a Explain why the resistance of the wire increases when it is stretched and show that the change in resistance is proportional to the extension, provided that the cross-sectional area is constant. [2]

b The resistance of the strain gauge when unstretched is 150 Ω. Calculate the reading on the voltmeter. [2]

c State and explain the effect on the voltmeter reading when the strain gauge is stretched. [2]

10 In a fibre optic system, sound into a microphone is used to produce different intensities of light that are passed along the glass fibre.

The microphone is connected to a non-inverting amplifier, the output of which is connected to an LED. An operational amplifier with dual 15 V power supplies is used as a non-inverting amplifier and the microphone produces a maximum output of +20 mV.

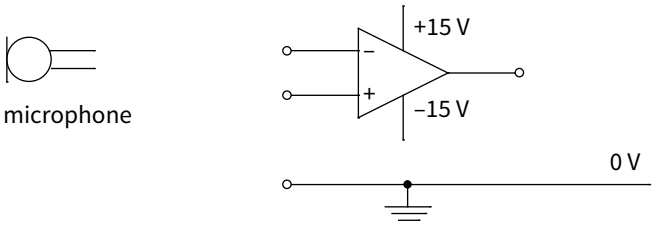


Figure 25.29 For End-of-chapter Question 10.

a **i** Copy and complete the circuit in Figure 25.29 to show the resistors necessary for a non-inverting amplifier. Make the necessary connection between the microphone and the amplifier. [2]

ii Add to your diagram the necessary output components for the amplifier to work as described. [2]

b State **two** advantages of using negative feedback in the amplifier circuit. [2]

c **i** Calculate the maximum possible gain for the amplifier. [2]

ii The amplifier is used with the maximum possible gain. Suggest suitable values for the resistance of all the resistors drawn in **a i**. [2]

d Explain what happens in the system when the voltage output from the microphone rises. [2]

11 A student builds an electronic circuit to control a motor using two identical LDRs, P and Q (Figure 25.30). In a room with the same amount of light on each LDR, the resistance of each LDR is 500 Ω. When the light from a torch shines on either LDR its resistance becomes 200 Ω.

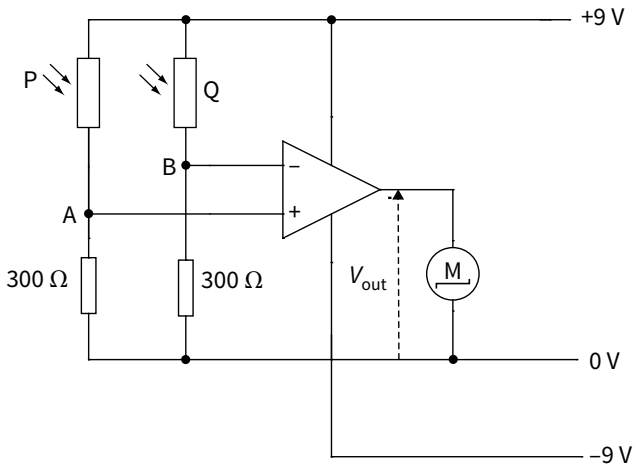


Figure 25.30 For End-of-chapter Question 11.

a State the value of the output voltage V_{out} of the op-amp when:

i the potential at point A is greater than the potential at point B [1]

ii the potential at point B is greater than the potential at point A. [1]

b Calculate the potential at point A when a torch shines on P. [2]

c **i** Explain what happens in the circuit when a torch shines on P but not on Q. [1]

ii Explain what happens in the circuit when a torch shines on Q but not on P. [1]

d Explain why it is difficult to know what will happen if no torch shines on P or Q. [1]

12 **a** Explain what is meant by a **virtual earth** and use your explanation to show that the gain of the inverting amplifier in Figure 25.31 is $-\frac{R_1}{R_2}$. [5]

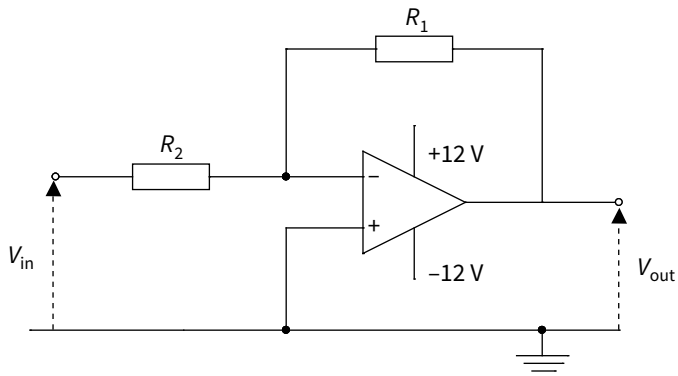


Figure 25.31 For End-of-chapter Question 12.

b The value of R_1 is 20 kΩ and the value of R_2 is 4 kΩ. The value of V_{in} is 0.50 V.

i Calculate the value of V_{out}. [2]

ii A resistor of 4 kΩ is placed in parallel with R_2. Calculate the new value of V_{out}. [2]

13 In order to switch on a heater when the temperature falls below a set value, a potential divider is connected to a switching circuit. When the input voltage to the switching circuit falls below 0.5 V, it switches the heater on.

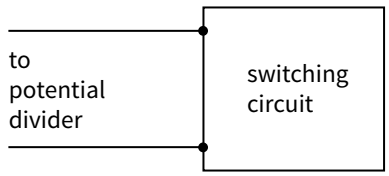


Figure 25.32 For End-of-chapter Question 13.

a Copy Figure 25.32 and add a suitable potential divider circuit to trigger the switching circuit. [2]

b Explain how the operator could lower the temperature at which the heater is switched on. [1]

14 Figure 25.33 shows a circuit used to monitor the variation of light intensity in a room.

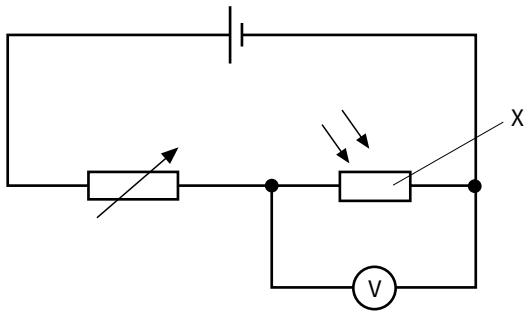


Figure 25.33 For End-of-chapter Question 14.

a Identify the component X and describe how the circuit works. [3]

b Suggest the reason for including the variable resistor in the circuit. [1]

Chapter 26: Magnetic fields and electromagnetism

Learning outcomes

You should be able to:

- describe a magnetic field as a field of force and use field lines to represent a field
- determine the magnitude and direction of the force on a current-carrying conductor in a magnetic field
- define magnetic flux density and describe how it can be measured using a current balance
- explain the forces between current-carrying conductors

Magnets and currents

The train shown in Figure 26.1 is supported at a precise distance above the track by computer-controlled electromagnets. In this chapter, we will look at magnetic forces and fields, how they arise and how they interact.



Figure 26.1 This high-speed train is magnetically levitated so that it avoids friction with the track.

Producing and representing magnetic fields

A magnetic field exists wherever there is force on a magnetic pole. As we saw with electric and gravitational fields, a magnetic field is a field of force.

You can make a magnetic field in two ways: using a permanent magnet, or using an electric current. You should be familiar with the magnetic field patterns of bar magnets (Figure 26.2). These can be shown using iron filings or plotting compasses.

We represent magnetic field patterns by drawing magnetic field lines:

- The magnetic field lines come out of north poles and go into south poles.
- The direction of a field line at any point in the field shows the direction of the force that a 'free' magnetic north pole would experience at that point.
- The field is strongest where the field lines are closest together.

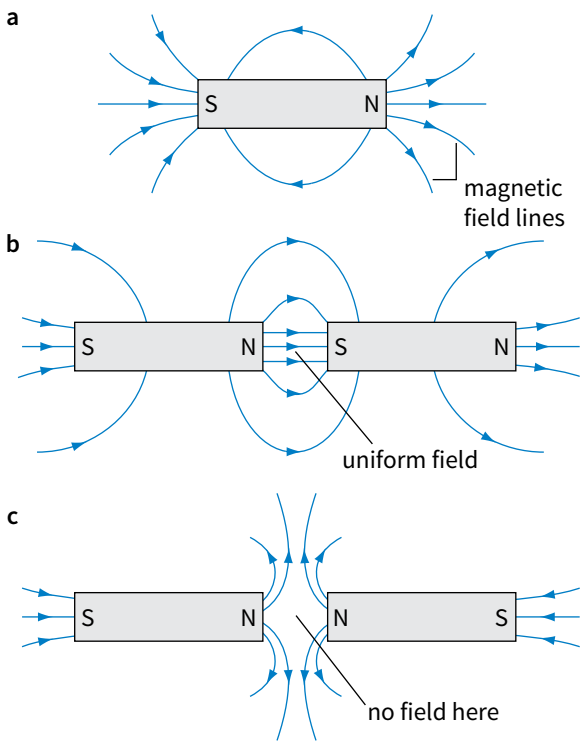


Figure 26.2 Magnetic field patterns: **a** for a bar magnet; **b** for two attracting bar magnets; **c** for two repelling bar magnets.

An electromagnet makes use of the magnetic field created by an electric current (Figure 26.3a). A coil is used because this concentrates the magnetic field. One end becomes a north pole (field lines emerging), while the other end is the south pole. Another name for a coil like this is a **solenoid**. The field pattern for the solenoid looks very similar to that of a bar magnet (see Figure 26.2a), with field lines emerging from a north pole at one end and returning to a south pole at the other. The strength of the magnetic field of a solenoid can be greatly increased by adding a core made of a ferrous (iron-rich) material. For example, an iron rod placed inside the solenoid can act as a core; when the current flows through the solenoid, the iron core itself becomes magnetised and this produces a much stronger field. A flat coil (Figure 26.3b) has a similar field to that of a solenoid.

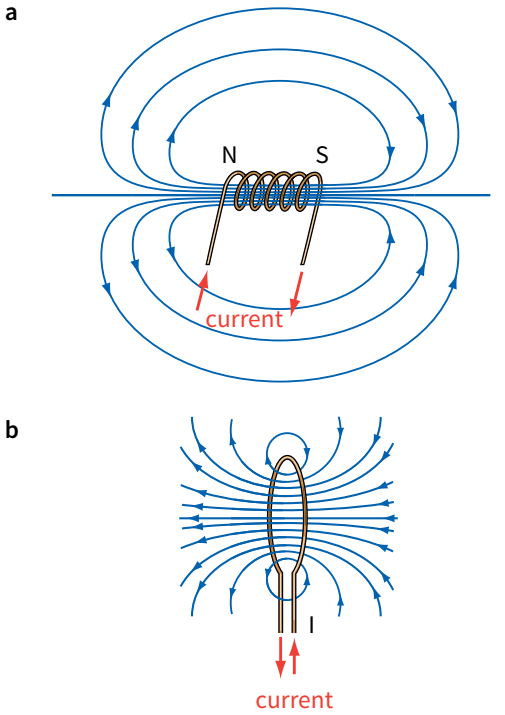


Figure 26.3 Magnetic field patterns for **a** a solenoid, and **b** a flat circular coil.

If we unravel an electromagnet, we get a weaker field. This, too, can be investigated using iron filings or compasses. The magnetic field pattern for a long current-carrying wire is very different from that of a solenoid. The magnetic field lines shown in Figure 26.4 are circular, centred on the long current-carrying wire. Further away from the wire, the field lines are drawn further apart, representing the weaker field at this distance. Reversing the current reverses the direction of the field.

All magnetic fields are created by **moving** charges. (In the case of a wire, the moving charges are free electrons.) This is even true for a permanent bar magnet. In a permanent magnet, the magnetic field is produced by the movement of electrons within the atoms of the magnet. Each electron represents a tiny current as it circulates around within its atom, and this current sets up a magnetic field. In a ferrous material such as iron, the weak fields due to all the electrons combine together to make a strong field, which spreads out into the space beyond the magnet. In non-magnetic materials, the fields produced by the electrons cancel each other out.

Field direction

The idea that magnetic field lines emerge from north poles and go into south poles is simply a convention. Figure 26.5 shows some useful rules for remembering the direction of the magnetic field produced by a current.

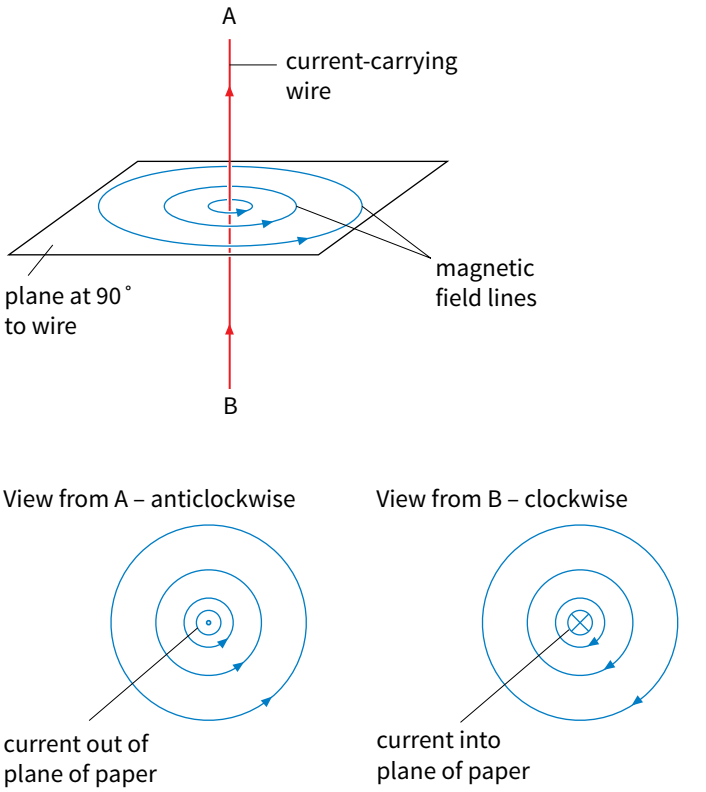


Figure 26.4 The magnetic field pattern around a current-carrying wire. The diagram also shows the convention used to indicate the direction of current.

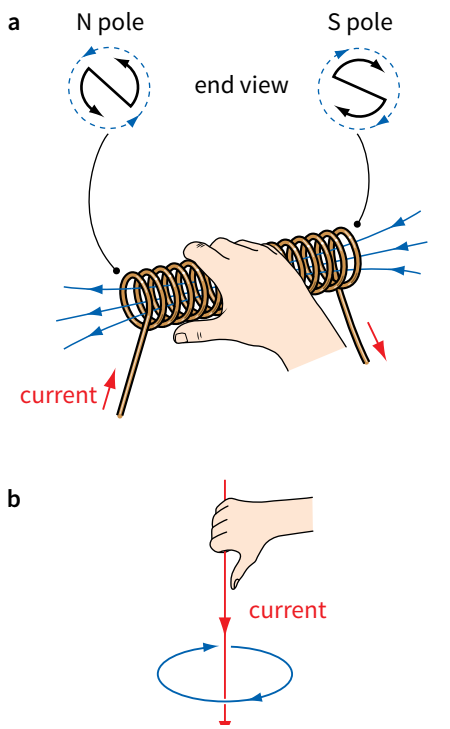


Figure 26.5 Two rules for determining the direction of a magnetic field, **a** inside a solenoid and **b** around a current-carrying wire.

The **right-hand grip rule** gives the direction of magnetic field lines in an electromagnet. Grip the coil so that your fingers go around it following the direction of the current. Your thumb now points in the direction of the field lines inside the coil, i.e. it points towards the electromagnet's north pole.

Another way to identify the poles of an electromagnet is to look at it end on, and decide which way round the current is flowing. The top diagrams in Figure 26.5 show how you can remember that clockwise is a south pole, anticlockwise is a north pole.

The circular field around a wire carrying a current does not have magnetic poles. To find the direction of the magnetic field you need to use another rule, the **right-hand rule**. Grip the wire with your right hand, pointing your thumb in the direction of the current. Your fingers curl around in the direction of the magnetic field.

Note that these two rules are slightly different. The right-hand grip rule applies to a solenoid; the fingers are curled in the direction of the current and the thumb then gives the direction of the field. The right-hand rule applies to a current in a straight wire; the thumb is pointed in the direction of the current and the fingers then give the direction of the field lines.

QUESTIONS

1 Sketch the magnetic field pattern around a long straight wire carrying an electric current. Now, alongside this first sketch, draw a second sketch to show the field pattern if the current flowing is doubled and its direction reversed.

2 Sketch the diagram in Figure 26.6, and label the north and south poles of the electromagnet. Show on your sketch the direction of the magnetic field (as shown by the needle of a plotting compass) at each of the positions A, B, C and D.

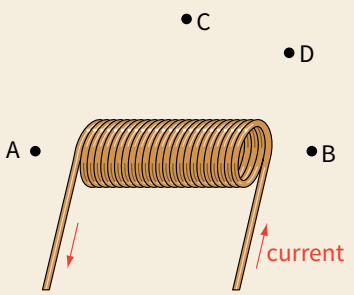


Figure 26.6 A current-carrying solenoid. For Question 2.

QUESTION

3 State which of the pairs of electromagnets shown in Figure 26.7 attract one another, and which repel.

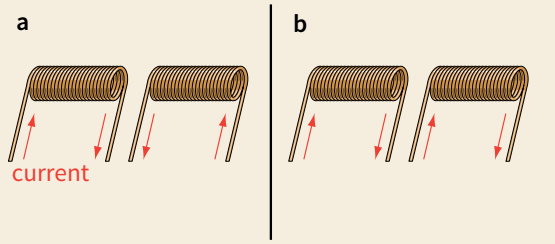


Figure 26.7 Two pairs of solenoids. For Question 3.

Magnetic force

A current-carrying wire is surrounded by a magnetic field. This magnetic field wil interact with an external magnetic field, giving rise to a force on the conductor, just like the fields of two interacting magnets. A simple situation is shown in Figure 26.8.

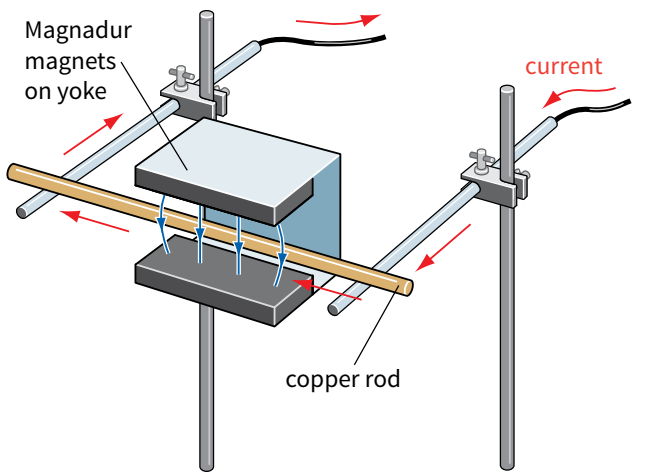


Figure 26.8 The copper rod is free to roll along the two horizontal aluminium 'rails'.

The magnets create a fairly uniform magnetic field. As soon as the current in the copper rod is switched on, the rod starts to roll, showing that a force is acting on it. We use **Fleming's left-hand (motor) rule** to predict the direction of the force on the current-carrying conductor, as explained in Box 26.1.

BOX 26.1: Using Fleming's left-hand rule

Look at Figure 26.9. There are three things here, all of which are mutually at right angles to each other – the magnetic field, the current in the rod, and the force on the rod. These can be represented by holding the thumb and the first two fingers of your left hand so that they are mutually at right angles (Figure 26.9). Your thumb and fingers then represent:

- thuMb – direction of Motion
- First finger – direction of external magnetic Field
- seCond finger – direction of conventional Current.

If the thumb and first two fingers of the left hand are held at right angles to one another, with the First finger pointing in the direction of the Field and seCond finger in the direction of the Current, then the thuMb points in the direction of the Motion or force.

You should practise using your left hand to check that the rule correctly predicts these directions.

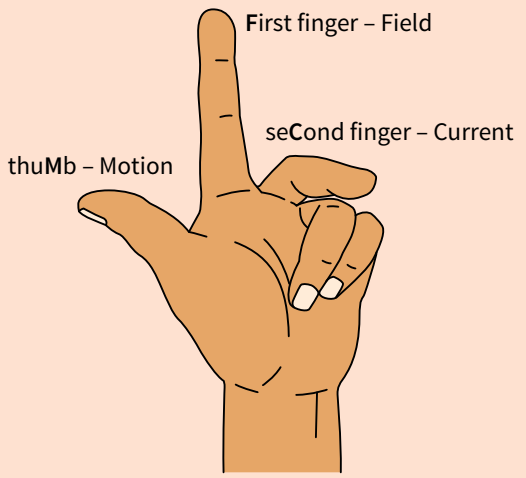


Figure 26.9 Fleming's left-hand (motor) rule.

Explaining the magnetic force

We can explain this force by thinking about the magnetic fields of the magnets and the current-carrying conductor. These fields combine or interact to produce the force on the rod.

Figure 26.10 shows:

- the external magnetic field of the magnets
- the magnetic field of the current-carrying conductor
- the combined fields of the current-carrying conductor and the magnets.

If you think of the magnetic field lines as elastic bands then you can see why the wire is pushed out in the direction shown.

The production of this force is known as the **motor effect**, because this force is used in electric motors. In a simple motor, a current in a coil produces a magnetic field; this field interacts with a second field produced by a permanent magnet.

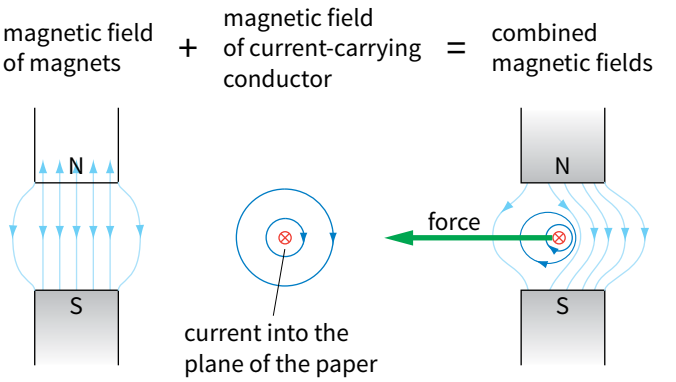


Figure 26.10 In the field of a permanent magnet, a current-carrying conductor experiences a force in accordance with Fleming's left-hand rule. The fields due to the permanent magnet and the current (left and centre) combine as shown on the right.

QUESTION

4 Figure 26.11 shows three examples of current-carrying conductors in magnetic fields. For each example, decide whether there will be a magnetic force on the conductor. If there is a force, in what direction will it act?

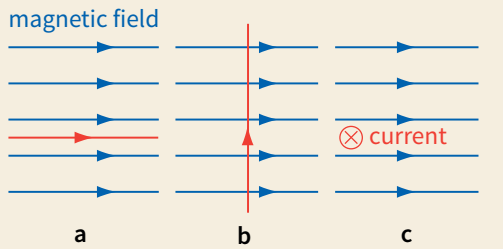


Figure 26.11 Three conductors in a magnetic field.

Magnetic flux density

In electric or gravitational field diagrams, the strength of the field is indicated by the separation between the field lines. The field is strongest where the field lines are closest together. The same is true for magnetic fields. The **strength** of a magnetic field is known as its **magnetic flux density**, with symbol B. (You can imagine this quantity to represent the number of magnetic field lines passing through a region per unit area.) The magnetic flux density is greater close to the pole of a bar magnet, and gets smaller as you move away from it.

We define gravitational field strength g at a point as the force per unit mass:

$$g = \frac{F}{m}$$

Electric field strength E is defined as the force per unit positive charge:

$$E = \frac{F}{Q}$$

In a similar way, magnetic flux density is defined in terms of the magnetic force experienced by a current-carrying conductor placed at **right angles** to a magnetic field. For a uniform magnetic field, the flux density B is defined by the equation:

$$B = \frac{F}{IL}$$

where F is the force experienced by a current-carrying conductor, I is the current in the conductor and L is the length of the conductor in the uniform magnetic field of flux density B. The direction of the force F is given by Fleming's left-hand rule.

Magnetic flux density is defined as follows:

> The magnetic flux density at a point in space is the force experienced per unit length by a long straight conductor carrying unit current and placed at right angles to the field at that point.

The unit for magnetic flux density is the tesla, T. It follows from the equation for B that $1\,\text{T} = 1\,\text{N}\,\text{A}^{-1}\,\text{m}^{-1}$.

The tesla is defined as follows:

> The magnetic flux density is 1 T when a wire carrying a current of 1 A placed at right angles to the magnetic field experiences a force of 1 N per metre of its length.

The force on the conductor is given by the equation:

$$F = BIL$$

Note that you can only use this equation when the field is at right angles to the current.

QUESTIONS

5 A current of 0.20 A flows in a wire of length 2.50 m placed at right angles to a magnetic field of flux density 0.06 T. Calculate the force on the wire.

6 A 20 cm length of wire is placed at right angles to a magnetic field. When a current of 1.5 A flows in the wire, a force of 0.015 N acts on it. Determine the flux density of the field.

7 A wire of length 50 cm carrying a current lies at right angles to a magnetic field of flux density 5 mT.

a If 10^{18} electrons pass a point in the wire each second, what current is flowing? (Electron charge $e = 1.60 \times 10^{-19}$ C.)

b What force acts on the wire?

Measuring magnetic flux density

Box 26.2 looks at two practical methods for measuring magnetic flux density.

BOX 26.2: Measuring magnetic flux density

Measuring *B* with a Hall probe

The simplest device for measuring magnetic flux density B is a Hall probe (Figure 26.12). When the probe is held so that the field lines are passing at right angles through the flat face of the probe, the meter gives a reading of the value of B. Some instruments are calibrated so that they give readings

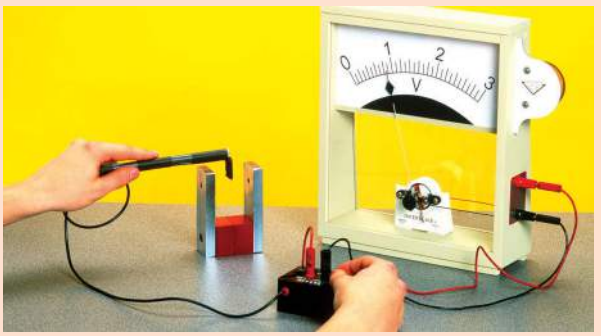


Figure 26.12 Using a Hall probe to measure the flux density between two magnets.

BOX 26.2: Measuring magnetic flux density (continued)

in microteslas (μT) or milliteslas (mT). Others are not calibrated, so you must either calibrate them or use them to obtain relative measurements of *B*.

A Hall probe must be held so that the field lines are passing directly through it, at right angles to the flat surface of the probe (Figure 26.13). If the probe is not held in the correct orientation, the reading on the meter will be reduced.

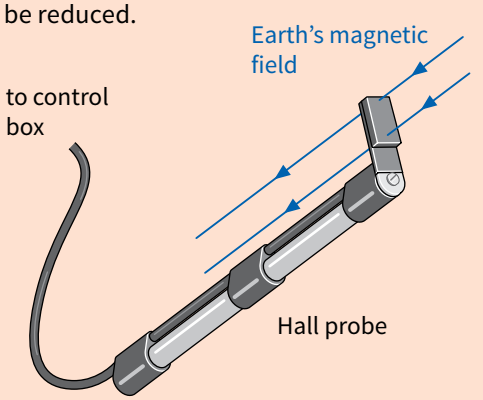


Figure 26.13 Magnetic flux lines must pass through the probe at 90° to the surface.

A Hall probe is sensitive enough to measure the Earth's magnetic flux density. The probe is first held so that the Earth's field lines are passing directly through it, as shown in Figure 26.13. In this orientation, the reading on the voltmeter will be a maximum. The probe is then rotated through 180° so that the magnetic field lines are passing through it in the opposite direction. The change in the reading of the meter is twice the Earth's magnetic flux density.

There is more about how the Hall probe works in Chapter 27.

Measuring *B* with a current balance

Figure 26.14 shows a simple arrangement that can be used to determine the flux density between two magnets. The magnetic field between these magnets is (roughly) uniform. The length *L* of the current-carrying wire in the uniform magnetic field can be measured using a ruler.

When there is no current in the wire, the magnet arrangement is placed on the top pan and the balance is zeroed. Now, when a current *I* flows in the wire, its value is shown by the ammeter. The wire experiences an upward force and, according to Newton's third law of motion, there is an equal and opposite force on the magnets. The magnets are pushed downwards and a reading appears on the scale of the balance. The force *F* is given by *mg*,

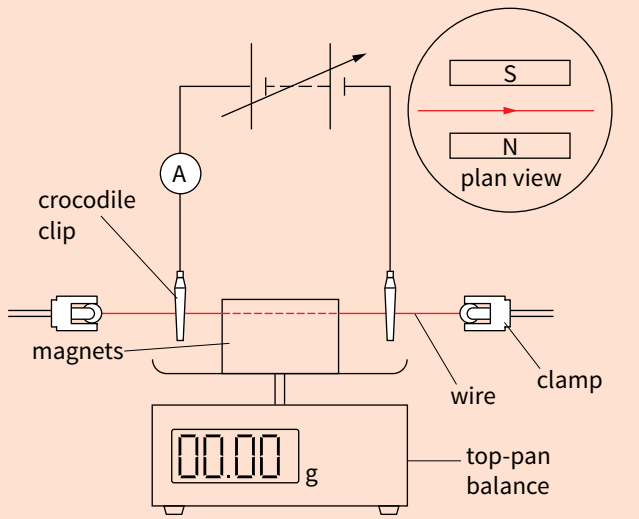


Figure 26.14 An arrangement to determine magnetic flux density in the laboratory.

where *m* is the mass indicated on the balance in kilograms and *g* is the acceleration of free fall (9.81 m s^{-2}).

Knowing *F*, *I* and *L*, the magnetic flux density *B* between the magnets can be determined using the equation:

$$B = \frac{F}{IL}$$

You can also use the arrangement in Figure 26.14 to show that the force is directly proportional to the current.

A system like this in effect 'weighs' the force on the current-carrying conductor, and is an example of a current balance. Another version of a current balance is shown in Figure 26.15. This consists of a wire frame which is balanced on two pivots. When a current flows through the frame, the magnetic field pushes the frame downwards. By adding small weights to the other side of the frame, you can restore it to a balanced position.

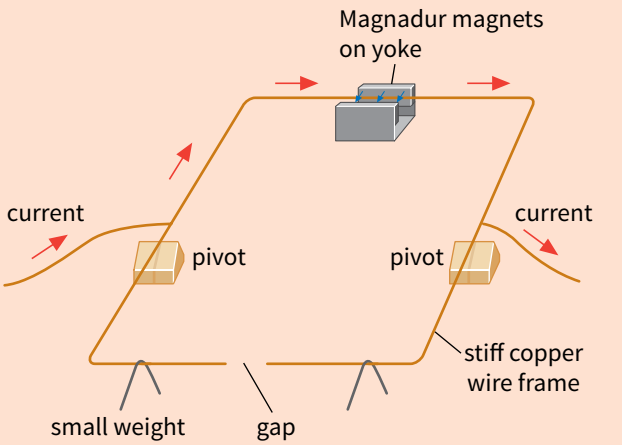


Figure 26.15 A simple laboratory current balance.

QUESTIONS

8 In the examples shown in the diagrams in Figure 26.16, which current balances will tilt? Will the side carrying the current tilt upwards or downwards?

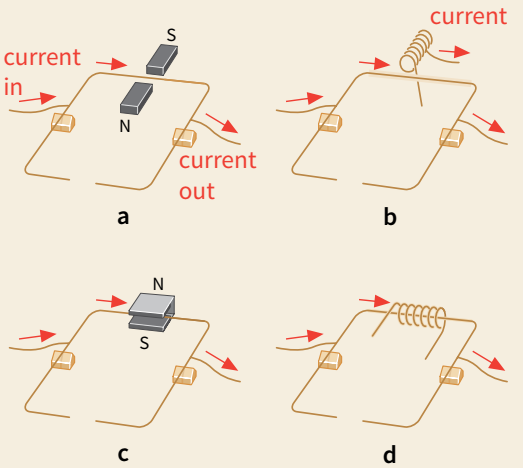


Figure 26.16 Four current balances – will they tip? For Question 8.

9 In the current balance shown in Figure 26.17, a current of 0.50 A is flowing. A student finds that a counterweight of mass 0.02 g is needed to restore balance. The section of the conductor in the field is 5.0 cm long. What is the flux density of the field?

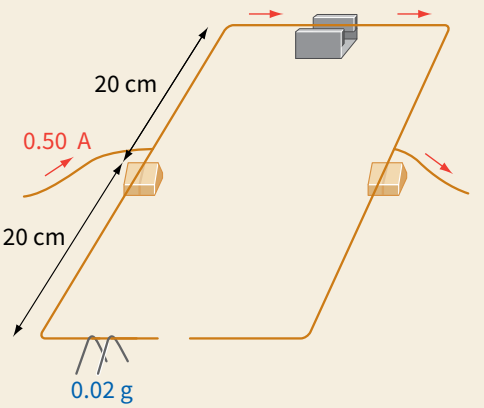


Figure 26.17 A current balance – see Question 9.

Currents crossing fields

At right angles

We explained the force on a current-carrying conductor in a field in terms of the interaction of the two magnetic fields: the field due to the current and the external field. Here is another, more abstract, way of thinking of this.

Whenever an electric current cuts across magnetic field lines (Figure 26.18), a force is exerted on the current-carrying conductor. This helps us to remember that a conductor experiences no force when the current is parallel to the field.

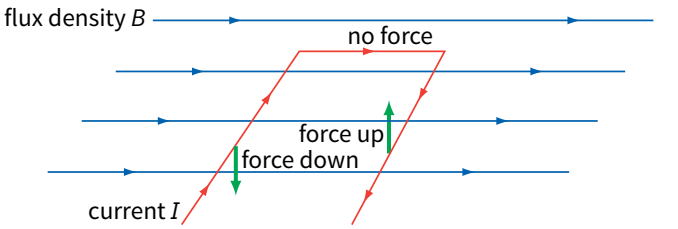


Figure 26.18 The force on a current-carrying conductor crossing a magnetic field.

This is a useful idea, because it saves us thinking about the field due to the current. In Figure 26.18, we can see that there is only a force when the current cuts across the magnetic field lines.

This force is very important – it is the basis of electric motors. Worked example 1 shows why a current-carrying coil placed in a magnetic field rotates.

QUESTIONS

10 A wire of length 50 cm carrying a current of 2.4 A lies at right angles to a magnetic field of flux density 5.0 mT. Calculate the force on the wire.

11 The coil of an electric motor is made up of 200 turns of wire carrying a current of 1.0 A. The coil is square, with sides of length 20 cm, and it is placed in a magnetic field of flux density 0.05 T.

a Determine the maximum force exerted on the side of the coil.

b In what position must the coil be for this force to have its greatest turning effect?

c List four ways in which the motor could be made more 'powerful' – that is, have greater torque.

WORKED EXAMPLE

1 An electric motor has a rectangular loop of wire with the dimensions shown in Figure 26.19. The loop is in a magnetic field of flux density 0.10 T. The current in the loop is 2.0 A. Calculate the torque that acts on the loop in the position shown.

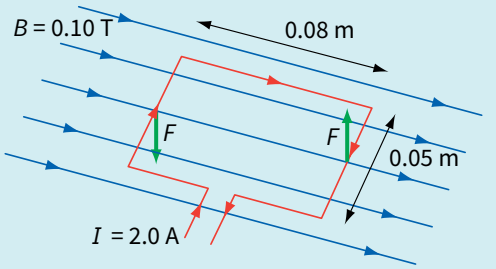


Figure 26.19 A simple electric motor – a current-carrying loop in a magnetic field.

Step 1 The quantities we know are:

$B = 0.10\,\text{T}$, $I = 2.0\,\text{A}$ and $L = 0.05\,\text{m}$

Step 2 Now we can calculate the force on one side of the loop using the equation $F = BIL$:

$F = 0.10 \times 2.0 \times 0.05$

$= 0.01\,\text{N}$

Step 3 The two forces on opposite sides of the loop are equal and anti-parallel. In other words, they form a couple. From Chapter 4, you should recall that the torque (moment) of a couple is equal to the magnitude of one of the forces times the perpendicular distance between them. The two forces are separated by 0.08 m, so:

torque = force × separation

$= 0.01 \times 0.08 = 8.0 \times 10^{-4}\,\text{N m}$

At an angle other than 90°

Now we must consider the situation where the current-carrying conductor cuts across a magnetic field at an angle other than a right angle. In Figure 26.20, the force gets weaker as the conductor is moved round from OA to OB, to OC and finally to OD. In the position OD, there is no force on the conductor. To calculate the force, we need to find the component of the magnetic flux density B at right angles to the current. This component is $B\sin\theta$, where θ is the angle between the magnetic field and the current or the conductor. Substituting this into the equation $F = BIL$ gives:

$$F = (B\sin\theta)IL$$

or simply:

$$F = BIL\sin\theta$$

Now look at Worked example 2.

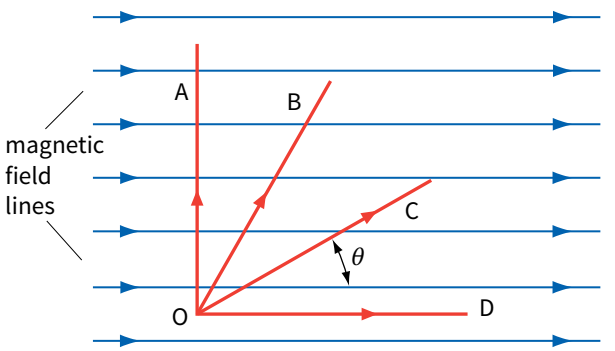


Figure 26.20 The force on a current-carrying conductor depends on the angle it makes with the magnetic field lines.

WORKED EXAMPLE

2 A conductor OC (see Figure 26.20) of length 0.20 m lies at an angle θ of 25° to a magnetic field of flux density 0.050 T. Calculate the force on the conductor when it carries a current of 400 mA.

Step 1 Write down what you know, and what you want to know:

$B = 0.050\,\text{T}$ $L = 0.20\,\text{m}$

$I = 400\,\text{mA} = 0.40\,\text{A}$ $\theta = 25°$

$F = ?$

Step 2 Write down the equation, substitute values and solve:

$F = BIL\sin\theta$

$F = 0.050 \times 0.40 \times 0.20 \times \sin 25° \approx 1.7 \times 10^{-3}\,\text{N}$

Step 3 Give the direction of the force. The force acts at 90° to the field and the current, i.e. perpendicular to the page. The left-hand rule shows that it acts downwards into the plane of the paper.

Note that the component of B parallel to the current is $B\cos\theta$, but this does not contribute to the force; there is no force when the field and current are parallel. The force F is at right angles to both the current and the field.

QUESTION

12 What force will be exerted on each of the currents shown in Figure 26.21, and in what direction will each force act?

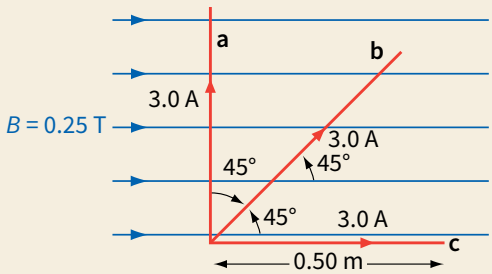


Figure 26.21 Three currents in a magnetic field.

Forces between currents

Any electric current has a magnetic field around it. If we have two currents, each will have its own magnetic field, and we might expect these to interact.

Explaining the forces

There are two ways to understand the origin of the forces between current-carrying conductors. In the first, we draw the magnetic fields around two current-carrying conductors (Figure 26.22a). Figure 26.22a shows two unlike (anti-parallel) currents, one flowing into the page, the other flowing out of the page. Their magnetic fields circle round, and in the space between the wires there is an extra-strong field. We imagine the field lines squashed together, and the result is that they push the wires apart. The diagram shows the resultant field, and the repulsive forces on the two wires.

Figure 26.22b shows the same idea, but for two like (parallel) currents. In the space between the two wires, the magnetic fields cancel out. The wires are pushed together.

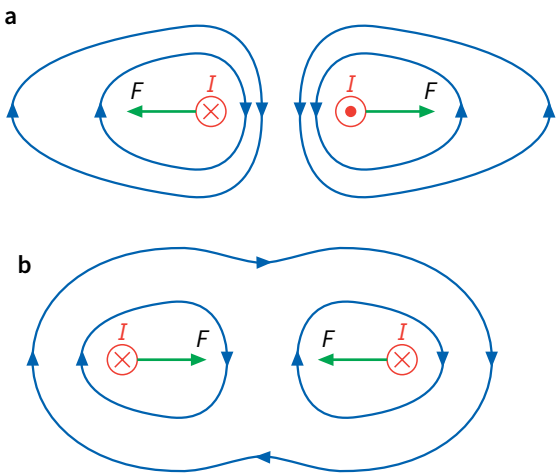


Figure 26.22 The forces between current-carrying wires.

BOX 26.3: Observing the forces between currents

You can observe the attraction and repulsion between two parallel currents using the equipment shown in Figure 26.23.

Two long thin strips of aluminium foil are mounted so that they are parallel and a small distance apart. By connecting them in series with a power supply, you can make a current occur in both of them. By changing the connections, you can make the current first in the same direction through both strips (parallel currents) and then in opposite directions (anti-parallel currents).

If you try this out, you will observe the strips of foil either bending towards each other or away from each other. (Foil is used because it is much more flexible than wire.)

You should find that parallel currents attract one another, while anti-parallel currents repel. This may seem surprising, since we are used to opposite charges attracting, and opposite magnetic poles attracting. Now we have found that opposite currents repel one another.

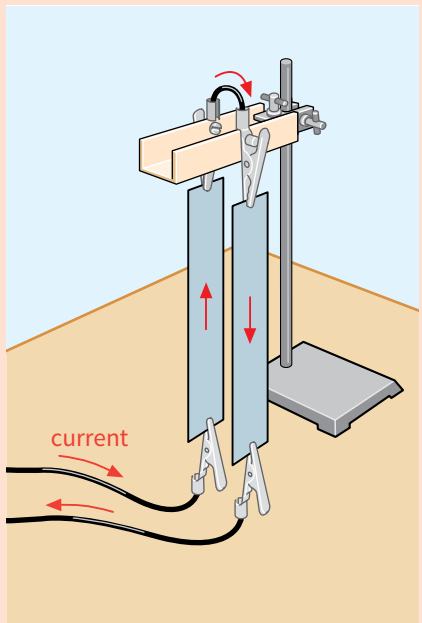


Figure 26.23 Current in two aluminium strips – their magnetic fields interact.

The other way to explain the forces between the currents is to use the idea of the motor effect. Figure 26.24 again shows two like currents, I_1 and I_2, but this time we only consider the magnetic field of one of them, I_1. The second current I_2 is flowing across the magnetic field of I_1; from the diagram, you can see that B is at right angles to I_2. Hence there will be a force on I_2 (the *BIL* force), and we can find its direction using Fleming's left-hand rule. The arrow shows the direction of the force, which is towards I_1. Similarly, there will be a *BIL* force on I_1, directed towards I_2.

These two forces are equal and opposite to one another. They are an example of an action and reaction pair, as described by Newton's third law of motion.

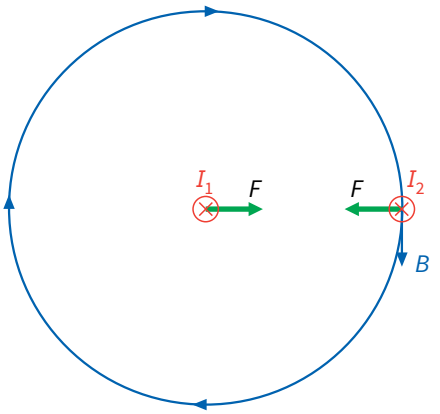


Figure 26.24 Explaining the force between two currents.

QUESTION

13 Two flat circular coils of wire are set up side by side, as shown in Figure 26.25. They are connected in series so that the same current flows around each, and in the same direction. Will the coils attract or repel one another? Explain your answer, first by describing the coils as electromagnets, and secondly by considering the forces between parallel currents. What will happen if the current is reversed in both coils?

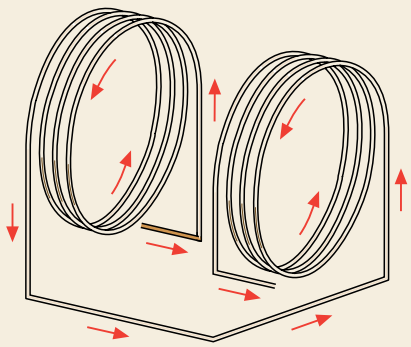


Figure 26.25 Two coils carrying the same current – see Question 13.

Relating SI units

In this chapter, we have seen how one SI unit, the tesla, is defined in terms of three others, the amp, the metre and the newton. It is an essential feature of the SI system that all units are carefully defined; in particular, derived units such as the newton and tesla must be defined in terms of a set of more fundamental units called **base units**.

We met the idea of base units in Chapter 3. The SI system of units has seven base units, of which you have met six. These are:

m kg s A K mol

(The seventh is the candela, cd, the unit of luminous intensity.) Each base unit is carefully defined; for example, the ampere can be defined in terms of the magnetic force between two parallel wires carrying a current. The exact definition is not required, but you should know that the ampere is itself a base unit. Other units are known as **derived units**, and can be deduced from the base units. For example, as shown in Chapter 3, the newton is given by:

$1\,\text{N} = 1\,\text{kg m s}^{-2}$

Similarly, in this chapter, you have learned about the tesla, the unit of magnetic flux density, given by:

$1\,\text{T} = 1\,\text{N A}^{-1}\,\text{m}^{-1}$ or $1\,\text{T} = 1\,\text{kg A}^{-1}\,\text{s}^{-2}$

If you learn formulae relating physical quantities, you can replace the quantities by their units to see how the units are defined. For example:

force = mass × acceleration $F = ma$ $\text{N} = \text{kg m s}^{-2}$

You should be able to picture how the different derived units form a logical sequence, as shown in Table 26.1.

Base units	Derived units	because
m, kg, s	newton N = kg m s^{-2}	$F = ma$
	joule J = kg m^2 s^{-2}	$W = Fd$
	watt W = kg m^2 s^{-3}	$P = \frac{W}{t}$
m, kg, s, A	coulomb C = A s	$Q = It$
	volt V = kg m^2 A^{-1} s^{-3}	$V = \frac{W}{Q}$
	tesla T = kg A^{-1} s^{-2}	$B = \frac{F}{IL}$

Table 26.1 How derived units relate to base units in the SI system.

Comparing forces in magnetic, electric and gravitational fields

We have now considered three types of field: electric (Chapters 8 and 23), gravitational (Chapter 18) and magnetic (this chapter). What are the similarities and differences between these three types of field?

Modern physics sees magnetic fields and electric fields as two parts of a combined whole, an electromagnetic field. Gravitational fields, however, are different in nature to electromagnetic fields.

Gravitational and electric fields are defined in terms of placing a test mass or a test charge at a point to measure the field strength. Similarly, a test wire carrying a current can be placed at a point to measure the magnetic field strength. Therefore all fields are defined in terms of the force on a unit mass, charge or current.

Other features that all fields share include:

- Action at a distance, between masses, between charges or between wires carrying currents.
- Decreasing strength with distance from the source of the field.
- Representation by field lines, the direction of which show the direction of the force at points along the line; the density of field lines indicates the relative strength of the field.

How do the forces arising from these fields compare? The answer depends on the exact situation. Using ideas that you have studied earlier, you should be able to confirm each of the following values:

- The force between two 1 kg masses 1 m apart = 6.7×10^{-11} N
- The force between two charges of 1 C placed 1 m apart = 9.0×10^{9} N
- The force per metre on two wires carrying a current of 1 A placed 1 m apart = 2.0×10^{-7} N

This might suggest that the electric force is strongest and gravity is the weakest. Certainly if you consider an electron in a hydrogen atom moving in a circular orbit around a proton, the electrical force is 10^{39} times the gravitational force. So for an electron, or any other small charged object, electric forces are the most significant. However, over larger distances and with objects of large mass, the gravitational field becomes the most significant. For example, the motions of planets in the Solar System are affected by the gravitational field but the electromagnetic field is comparatively insignificant.

Summary

- Moving charges produce a magnetic field; this is electromagnetism.
- A current-carrying conductor has concentric magnetic field lines. The magnetic field pattern for a solenoid or flat coil resembles that of a bar magnet.
- The separation between magnetic field lines is an indication of the field's strength.
- Magnetic flux density B is defined by the following equation:

  $$B = \frac{F}{IL}$$

  where F is the force experienced by a current-carrying conductor, I is the current in the conductor and L is the length of the conductor in the uniform magnetic field.
- The unit of magnetic flux density is the tesla (T). $1\ \text{T} = 1\ \text{N A}^{-1}\ \text{m}^{-1}$.
- The magnetic flux density is 1 T when a wire carrying a current of 1 A placed at right angles to the magnetic field experiences a force of 1 N per metre of its length.
- The magnetic force on a current-carrying conductor is given by $F = BIL \sin\theta$.
- The force on a current-carrying conductor can be used to measure the flux density of a magnetic field using a current balance.
- A force acts between current-carrying conductors due to the interaction of their magnetic fields.

End-of-chapter questions

1 A current-carrying wire is placed in a uniform magnetic field.

a When does the wire experience the maximum force due to the magnetic field? [1]

b When does the wire experience no force due to the magnetic field? [1]

2 A current-carrying conductor placed at right angles to a uniform magnetic field experiences a force of 4.70×10^{-3} N. Determine the force on the wire when, separately:

a the current in the wire is increased by a factor of 3.0 [2]

b the magnetic flux density is halved [2]

c the length of the wire in the magnetic field is reduced to 40% of its original length. [2]

3 A copper wire carrying a current of 1.2 A has 3.0 cm of its length placed in a uniform magnetic field, as shown in Figure 26.26. The force experienced by the wire is 3.8×10^{-3} N when the angle between the wire and the magnetic field is 50°.

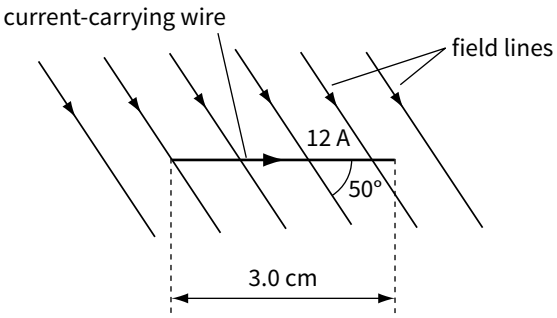


Figure 26.26 For End-of-chapter Question 3.

a Calculate the magnetic field strength. [3]

b What is the direction of the force experienced by the wire? [1]

4 Figure 26.27 shows a view from above of two long, parallel strips of aluminium foil A and B, carrying a current downwards into the paper.



Figure 26.27 For End-of-chapter Question 4.

a On a copy of the diagram, draw the magnetic field around and between the two strips. [2]

b State and explain the direction of the forces caused by the current in the strips. [4]

5 Figure 26.28 shows a wire XY which carries a constant direct current. Plotting compass R, placed alongside the wire, points due north. Compass P is placed below the wire and compass Q is placed above the wire.

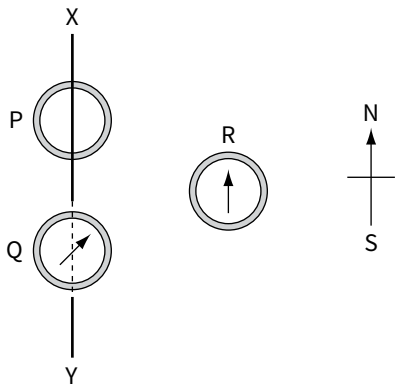


Figure 26.28 For End-of-chapter Question 5.

a State the direction of the current in the wire. [2]

b State in which direction compass P points. [2]

c State in which direction compass Q points if the current in the wire is reversed. [2]

6 Figure 26.29 shows a rectangular metal frame PQRS placed in a uniform magnetic field.

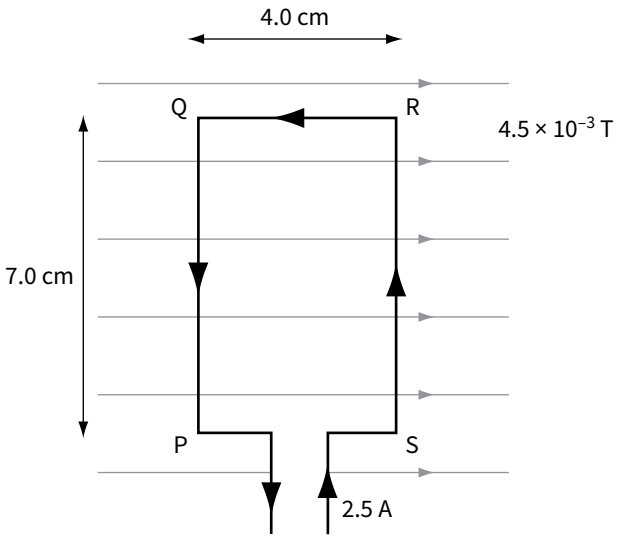


Figure 26.29 For End-of-chapter Question 6.

The magnetic flux density is 4.5×10^{-3} T and the current in the metal frame is 2.5 A.

a Calculate the force experienced by side PQ of the frame. [3]

b Suggest why side QR does not experience a force. [1]

c Describe the motion of the frame immediately after the current in the frame is switched on. [2]

d Calculate the maximum torque (moment) exerted about an axis parallel to side PQ. [2]

7 Figure 26.30 shows a current-carrying wire frame placed between a pair of Magnadur magnets on a yoke. A pointer is attached to the wire.

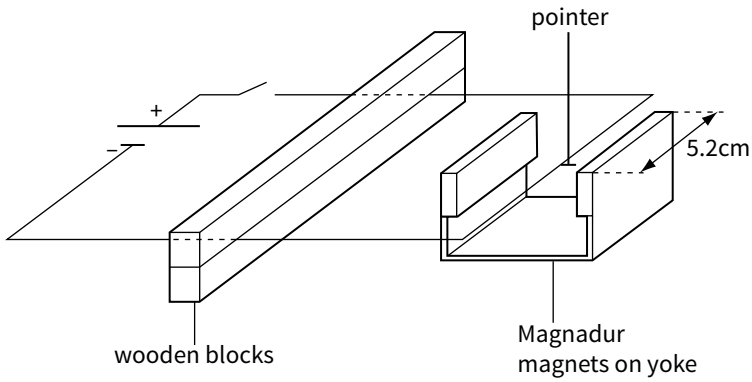


Figure 26.30 For End-of-chapter Question 7.

A current of 8.5 A in the wire causes the pointer to move vertically upwards. A small paper tape is attached to the pointer. The weight of the paper tape causes the pointer to return to its initial position (with no current and no paper tape). The paper's mass is subsequently found to be 60 mg. The section of the wire between the poles of the magnetic has a length of 5.2 cm.

a State the direction of the magnetic field. [1]

b Calculate the force on the wire due to the magnetic field when it carries a current of 8.5 A. [2]

c Calculate the magnetic flux density of the magnetic field between the poles of the magnet. [3]

d Describe what happens to the frame if low-frequency alternating current passes through the wire. [1]

8 **a** The size of the force acting on a wire carrying a current in a magnetic field is proportional to the size of the current in the wire. With the aid of a diagram, describe how this can be demonstrated in a school laboratory. [5]

b At a certain point on the Earth's surface the horizontal component of the Earth's magnetic field is 1.6×10^{-5} T. A piece of wire 3.0 m long and of weight 0.020 N lies in an east–west direction on a laboratory bench. When a large current flows in the wire, the wire just lifts off the surface of the bench.

i State the direction of the current in the wire. [1]

ii Calculate the minimum current needed to lift the wire from the bench. [3]

9 Figure 26.31 shows a fixed horizontal wire passing centrally between the poles of a permanent magnet that is placed on a top-pan balance. With no current flowing, the balance records a mass of 102.45 g. When a current of 4.0 A flows in the wire, the balance records a mass of 101.06 g.

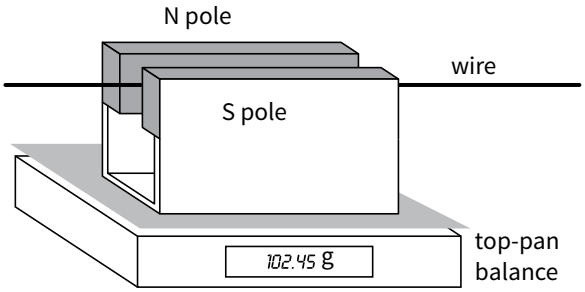


Figure 26.31 For End-of-chapter Question 9.

a Explain why the reading on the top-pan balance decreases when the current is switched on. [2]

b State and explain the direction of the current flow in the wire. [2]

c The length of the wire in the magnetic field is 5.0 cm. Calculate the average magnetic flux density between the poles of the magnet. [2]

d Sketch a graph, with balance reading on the vertical axis and current on the horizontal axis, to show how the balance reading changes when the current is altered. [2]

10 **a** Define **magnetic flux density** and explain the similarity with the definition of electric field strength. [3]

b Two thin horizontal wires are placed in a north–south direction. One wire is placed on a bench and the other wire is held 3.0 cm directly above the first wire.

i When equal currents flow in the two wires, the force exerted on the bench by the lower wire decreases. Explain why this is so. What can you say about the directions of the currents in the two wires? [4]

ii The magnetic flux density B at a distance x from a long straight wire carrying a current I is given by the expression $B = 2.0 \times 10^{-7}\frac{I}{x}$, where x is in metres and I is in amps. When the current in each wire is 4.0 A, calculate the force per unit length on one wire due to the current in the other. [3]

Chapter 27: Charged particles

Learning outcomes

You should be able to:

- determine the magnitude and direction of the force on a charge moving in a magnetic field
- analyse the deflection of beams of charged particles in electric and magnetic fields
- explain the principles of methods for measuring speed and charge-to-mass ratio for electrons
- derive and use an expression for the Hall voltage

Moving particles

The world of atomic physics is populated by a great variety of particles – electrons, protons, neutrons, positrons and many more. Many of these particles are electrically charged, and so their motion is influenced by electric and magnetic fields. Indeed, we use this fact to help us to distinguish one particle from another. Figure 27.1 shows the tracks of particles in a detector called a bubble chamber. A photon (no track) has entered from the top and collided with a proton; the resulting spray of nine particles shows up as the gently curving tracks moving downwards. The tracks curve because the particles are charged and are moving in a magnetic field. The tightly wound spiral tracks are produced by electrons which, because their mass is small, are more dramatically affected by the field.

In this chapter, we will look at how charged particles behave in electric and magnetic fields and how this knowledge can be used to control beams of charged particles. At the end of the chapter, we will look at how this knowledge was used to discover the electron and to measure its charge and mass.



Figure 27.1 A bubble chamber image of the tracks of sub-atomic particles. The tracks curve because the charged particles are affected by the presence of a magnetic field.

Observing the force

You can use your knowledge of how charged particles and electric currents are affected by fields to interpret diagrams of moving particles. You must bear in mind that, by convention, the direction of conventional electric current is the direction of flow of positive charge. When electrons are moving, the conventional current is regarded as flowing in the opposite direction.

An electron beam tube (Figure 27.2) can be used to demonstrate the magnetic force on a moving charge. A beam of electrons is produced by an 'electron gun', and magnets or electromagnets are used to apply a magnetic field.

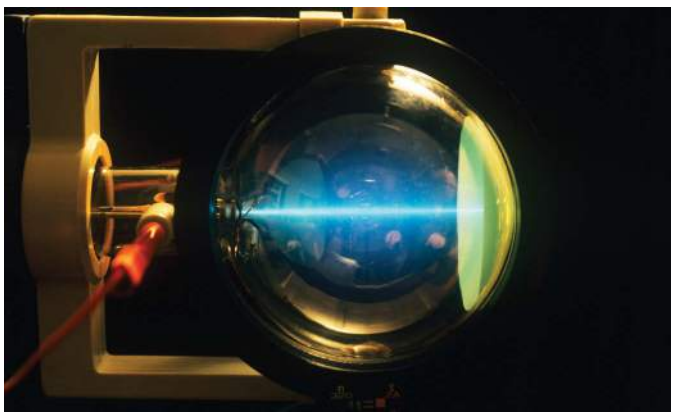


Figure 27.2 An electron beam tube.

You can use such an arrangement to observe the effect of changing the strength and direction of the magnetic field, and the effect of reversing the field.

If you are able to observe a beam of electrons like this, you should find that the force on the electrons moving through the magnetic field can be predicted using Fleming's left-hand rule. In Figure 27.3, a beam of electrons is moving from right to left, into a region where

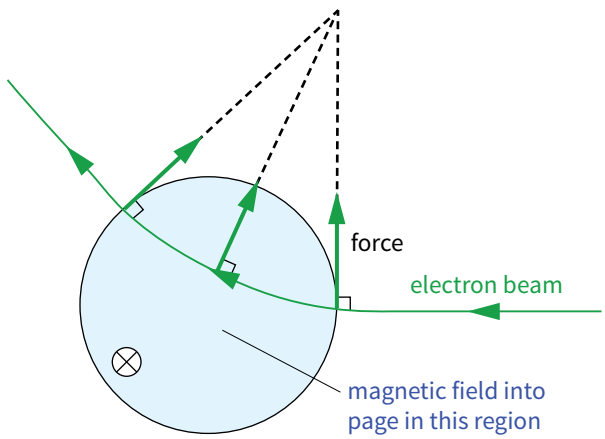


Figure 27.3 A beam of electrons is deflected as it crosses a magnetic field. The magnetic field into the plane of the paper is represented by the cross in the circle.

a magnetic field is directed into the plane of the paper. Since electrons are negatively charged, they represent a conventional current from left to right. Fleming's left-hand rule predicts that, as the electrons enter the field, the force on them will be upwards and so the beam will be deflected up the page. As the direction of the beam changes, so does the direction of the force. The force due to the magnetic field is always at 90° to the velocity of the electrons. It is this force that gives rise to the motor effect. The electrons in a wire experience a force when they flow across a magnetic field, and they transfer the force to the wire itself. In the past, most oscilloscopes, monitors and television sets made use of beams of electrons. The beams were moved about using magnetic and electric fields, and the result was a rapidly changing image on the screen.

BOX 27.1: Electron beam tubes

Figure 27.4 shows the construction of a typical tube. The electron gun has a heated cathode. The positively charged anode attracts electrons from the negative cathode, and they pass through the anode to form a narrow beam in the space beyond. The direction of the beam can be changed using an electric field between two plates (as in Figure 27.4), or a magnetic field created by electromagnetic coils.

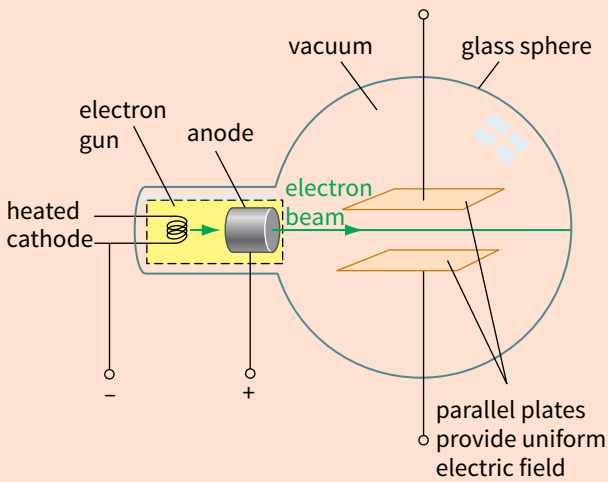


Figure 27.4 The construction of an electron beam tube.

QUESTION

1 In the diagram in Figure 27.5, radiation from a radioactive material passes through a region of uniform magnetic field. State whether each type of radiation has positive or negative charge, or is uncharged.

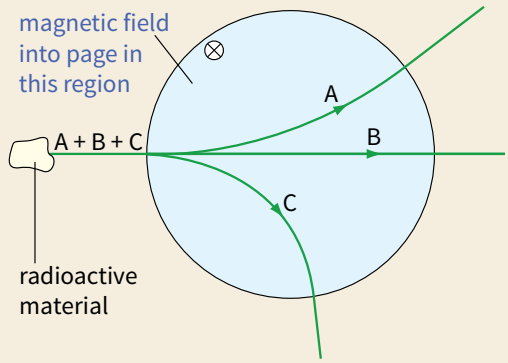


Figure 27.5 Three types of radiation passing through a magnetic field.

The magnetic force on a moving charge

We can make an intelligent guess about the factors that determine the size of the force on a moving charge in a uniform magnetic field (Figure 27.6). It will depend on:

- the magnetic flux density B (strength of the magnetic field)
- the charge Q on the particle
- the speed v of the particle.

The magnetic force F on a moving particle at right angles to a magnetic field is given by the equation:

$$F = BQv$$

The direction of the force can be determined from Fleming's left-hand rule. The force F is always at 90° to the velocity of the particle. Consequently, the path described by the particle will be an arc of a circle.

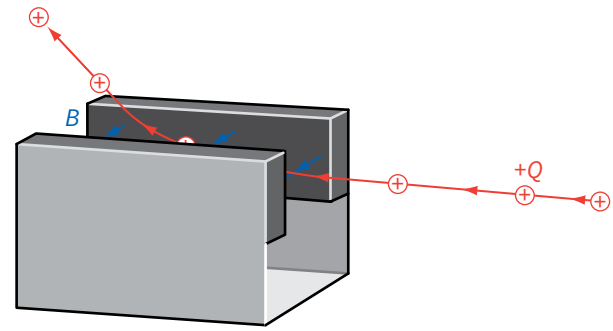


Figure 27.6 The path of a charged particle is curved in a magnetic field.

If the charged particle is moving at an angle θ to the magnetic field, the component of its velocity at right angles to B is $v \sin\theta$. Hence the equation becomes:

$$F = BQv\sin\theta$$

We can show that the two equations $F = BIL$ and $F = BQv$ are consistent with one another, as follows.

Since current I is the rate of flow of charge, we can write:

$$I = \frac{Q}{t}$$

Substituting in $F = BIL$ gives:

$$F = \frac{BQL}{t}$$

Now, $\frac{L}{t}$ is the speed v of the moving particle, so we can write:

$$F = BQv$$

For an electron, with a charge of $-e$, the magnitude of the force on it is:

$$F = Bev \quad (e = 1.60 \times 10^{-19}\,\text{C})$$

The force on a moving charge is sometimes called 'the *Bev* force'; it is this force acting on all the electrons in a wire which gives rise to 'the *BIL* force'.

Here is an important reminder: The force F is always at right angles to the particle's velocity v, and its direction can be found using the left-hand rule (Figure 27.7).

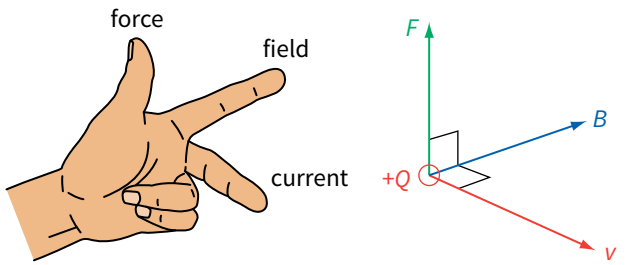


Figure 27.7 Fleming's left-hand rule, applied to a moving positive charge.

QUESTIONS

2 A beam of electrons, moving at $1.0 \times 10^6\,\text{m s}^{-1}$, is directed through a magnetic field of flux density 0.50 T. Calculate the force on each electron when **a** the beam is at right angles to the magnetic field, and **b** the beam is at an angle of 45° to the field.

3 Positrons are particles identical to electrons, except that their charge is positive (+*e*). Use a diagram to explain how a magnetic field could be used to separate a mixed beam of positrons and electrons.

Orbiting charges

Consider a charged particle moving at right angles to a uniform magnetic field. It will describe a circular path because the magnetic force F is always perpendicular to its velocity. We can describe F as a **centripetal force**, because it is always directed towards the centre of the circle.

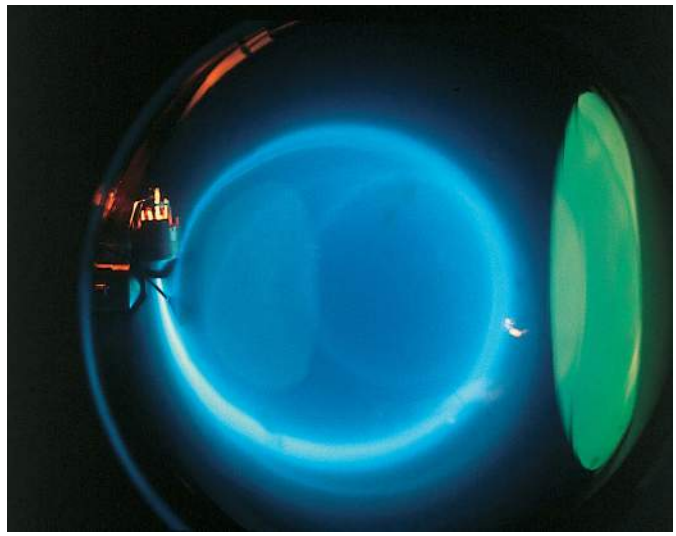


Figure 27.8 In this fine-beam tube, a beam of electrons is bent around into a circular orbit by an external magnetic field. The beam is shown up by the presence of a small amount of gas in the tube. (The electrons travel in an anticlockwise direction.)

Figure 27.8 shows a fine-beam tube. In this tube, a beam of fast-moving electrons is produced by an electron gun. This is similar to the cathode and anode shown in Figure 27.4, but in this case the beam is directed vertically downwards as it emerges from the gun. It enters the spherical tube, which has a uniform horizontal magnetic field. The beam is at right angles to the field and the *Bev* force pushes it round in a circle. The fact that the *Bev* force acts as a centripetal force gives us a clue as to how we can calculate the radius of the orbit of a charged particle in a uniform magnetic field. The centripetal force on the charged particle is given by:

$$\text{centripetal force} = \frac{mv^2}{r}$$

The centripetal force is provided by the magnetic force *Bev*. Therefore:

$$Bev = \frac{mv^2}{r}$$

Cancelling and rearranging to find r gives:

$$r = \frac{mv}{Be}$$

You can also write this equation in terms of the momentum p of the particle, that is:

$$p = Ber$$

The equation $r = \frac{mv}{Be}$ shows that:

- faster-moving particles move in bigger circles ($r \propto v$)
- particles with greater masses also move in bigger circles (they have more inertia: $r \propto m$)
- a stronger field makes the particles move in tighter circles ($r \propto \frac{1}{B}$).

This is made use of in a variety of scientific applications, such as particle accelerators and mass spectrometers. It can also be used to find the charge-to-mass ratio $\frac{e}{m_e}$ of an electron.

The charge-to-mass ratio of an electron

Experiments to find the mass of an electron first involve finding the charge-to-mass ratio $\frac{e}{m_e}$. This is known as the specific charge on the electron – the word 'specific' here means 'per unit mass'.

Using the equation for an electron travelling in a circle in a magnetic field, we have $\frac{e}{m_e} = \frac{v}{Br}$. Clearly, measurements of v, B and r are needed to measure $\frac{e}{m_e}$.

There are difficulties in measuring B and r. For example, it is difficult to measure r with a rule outside the tube in Figure 27.8 because of parallax error. Also, v must be measured, and you need to know how this is done. One way is to use the cathode–anode voltage V_{ca}. This p.d. causes each electron to accelerate as it moves from the cathode to the anode. If an individual electron has charge $-e$ then an amount of work $e \times V_{ca}$ is done on each electron. This is its kinetic energy as it leaves the anode:

$$eV_{ca} = \tfrac{1}{2} m_e v^2$$

where m_e is electron mass and v is the speed of the electron.

Eliminating v from the two equations $eV_{ca} = \frac{1}{2} m_e v^2$ and $r = \frac{m_e v}{Be}$ gives:

$$\frac{e}{m_e} = \frac{2V_{ca}}{r^2 B^2}$$

Hence, if we measure V_{ca}, r and B, we can calculate $\frac{e}{m_e}$. As we shall see shortly, the electron charge e can be measured more directly, and this allows us to calculate the electron mass m_e from the value of $\frac{e}{m_e}$.

WORKED EXAMPLE

1 An electron is travelling at right angles to a uniform magnetic field of flux density 1.2 mT. The speed of the electron is $8.0 \times 10^6\,\text{m s}^{-1}$. Calculate the radius of circle described by this electron. (For an electron, charge $e = 1.60 \times 10^{-19}\,\text{C}$ and mass $m_e = 9.11 \times 10^{-31}\,\text{kg}$.)

Step 1 Calculate the magnetic force on the electron.

$F = Bev = 1.2 \times 10^{-3} \times 1.60 \times 10^{-19} \times 8.0 \times 10^6$

$F = 1.536 \times 10^{-15}\,\text{N}$

Step 2 Use your knowledge of motion in a circle to determine the radius r.

$$F = \frac{m_e v^2}{r}$$

Therefore:

$$r = \frac{m_e v^2}{F} = \frac{9.11 \times 10^{-31} \times (8.0 \times 10^6)^2}{1.536 \times 10^{-15}}$$

$r \approx 3.8 \times 10^{-2}\,\text{m}$ (3.8 cm)

Note: The same result could have been obtained simply by using the equation:

$$r = \frac{m_e v}{Be}$$

QUESTIONS

4 Look at the photograph of the electron beam in the fine-beam tube (Figure 27.8). In which direction is the magnetic field (into or out of the plane of the photograph)?

5 The particles in the circular beam shown in Figure 27.8 all travel round in the same orbit. What can you deduce about their mass, charge and speed?

6 An electron beam in a vacuum tube is directed at right angles to a magnetic field, so that it travels along a circular path. Predict the effect on the size and shape of the path that would be produced (separately) by each of the following changes:

a increasing the magnetic flux density

b reversing the direction of the magnetic field

c slowing down the electrons

d tilting the beam, so that the electrons have a component of velocity along the magnetic field.

BOX 27.2: The deflection tube

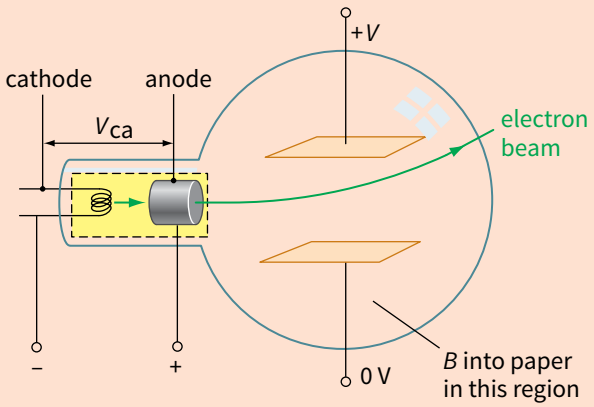


Figure 27.9 The path of an electron beam in a deflection tube.

A deflection tube (Figure 27.9) is designed to show a beam of electrons passing through a combination of electric and magnetic fields. By adjusting the strengths of the electric and magnetic fields, you can balance the two forces on the electrons, and the beam will remain horizontal. The magnetic field is provided by two coils, called Helmholtz coils (Figure 27.10), which give a very uniform field in the space between them.

If the electron beam remains straight, it follows that the electric and magnetic forces on each electron must have the same magnitude and act in opposite directions. Therefore:

electric force = magnetic force

(upward) (downward)

$$eE = Bev$$

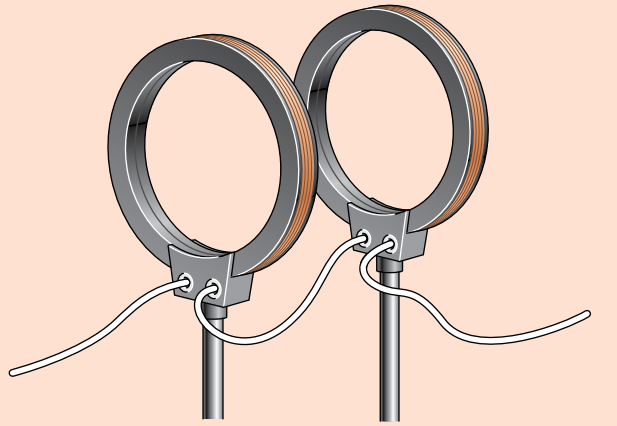


Figure 27.10 A pair of Helmholtz coils is used to give a uniform magnetic field.

where E is the electric field strength between the parallel plates with a p.d of V. The speed v of the electrons is simply related to the strengths of the two fields. That is:

$$v = \frac{E}{B}$$

The electric field strength is given by:

$$E = \frac{V}{d}$$

therefore:

$$v = \frac{V}{Bd}$$

Electric and magnetic fields

Now we will consider what happens when an electron beam passes through an electric field and a magnetic field at the same time.

Velocity selection

Balancing the effects of electric and magnetic fields is also used in a device called a **velocity selector**. This is used in devices such as mass spectrometers where it is desired to produce a beam of charged particles all moving with the same velocity. The construction of a velocity selector is shown in Figure 27.11.

The apparatus is very similar to the deflection tube in Figure 27.9. Two parallel plates are situated in an evacuated chamber. They provide a uniform electric field of strength E.

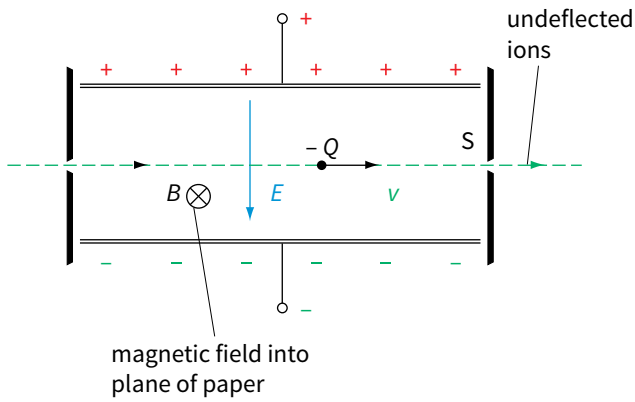


Figure 27.11 A velocity selector – only particles with the correct combination of charge, mass and velocity will emerge through the slit S.

The region between the plates is also occupied by a uniform magnetic field of flux density B which is at right angles to the electric field. Charged particles (electrons or ions) enter from the left. They all have the same charge and mass but are travelling at different speeds. The electric force Ee will be the same on all particles as it does not depend on their speed; however, the magnetic force Bev will be greater on those particles which are travelling faster. Hence, for particles travelling at the desired speed v, the electric and magnetic forces balance and they emerge undeflected from the slit S. If a negative ion has a speed greater than $\frac{V}{Bd}$ the downward magnetic force on it will be greater than the upward electric force. Thus it will be deflected downwards and it will hit below slit S.

Note that we do not have to concern ourselves with the gravitational force mg acting on the charged particles as this will be much smaller than the electric and magnetic forces.

QUESTION

7 This question is about the velocity selector shown in Figure 27.11.

a State the directions of the magnetic and electric forces on a positively charged ion travelling towards the slit S.

b The speed of the ion is given by the equation:

$$v = \frac{E}{B}$$

Calculate the speed of an ion emerging from the slit S when the magnetic flux density is 0.30 T and the electric field strength is $1.5 \times 10^3\ \text{V m}^{-1}$.

c Explain why ions travelling at a speed greater than your answer to **b** will not emerge from the slit.

The Hall effect

In Chapter 26, you saw how to use a Hall probe to measure magnetic flux density. The **Hall effect** is another mechanism in which the magnetic and electric forces on a moving charged particle are balanced.

A Hall probe works as follows. The probe itself is made of semiconductor (Figure 27.12). This material is used because the electrons move much faster in a semiconductor than in a metal for a given current, and so the effect is much greater. (Recall from Chapter 9 that the mean drift velocity of free electrons in a semiconductor is perhaps a million times greater than in a metal because there are many fewer free electrons in a semiconductor.)

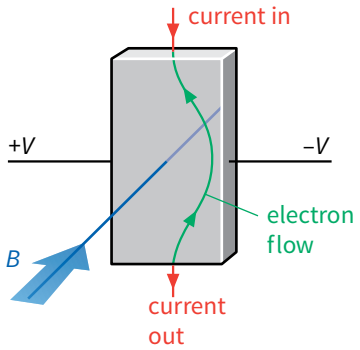


Figure 27.12 Electrons are deflected as they move through the Hall probe.

A small current flows through the probe from one end to the other. When a magnetic field is applied, the electrons are pushed sideways by the magnetic force, so that they accumulate along one side of the probe (the right-hand side in Figure 27.12). This is the Hall effect. The charge is detected as a small voltage across the probe, known as the **Hall voltage**. The greater the flux density of the field, the greater the Hall voltage. The control box amplifies the voltage and it is displayed by the meter. If the direction of the magnetic field is reversed, the electrons are pushed in the opposite direction and so the Hall voltage is reversed.

An equation for the Hall voltage

Using what we know about electric current and the forces on electric charges produced by electric and magnetic fields, we can derive an expression for the Hall voltage V_H. Figure 27.13 shows the situation. The Hall voltage is the voltage which appears between the two opposite sides of the slice.

As we have seen, this voltage arises because electrons accumulate on one side of the Hall probe. There is a corresponding lack of electrons on the opposite side, i.e. a positive charge. As a result, there is an electric field between the two sides. The electric field strength E is related to the Hall voltage V_H by:

$$E = \frac{V_H}{d}$$

where d is the width of the slice. Now, picture a single electron as it travels with drift velocity v through the slice. It will experience a force to the right, caused by the magnetic field; the magnitude of this force is Bev. It will also experience a second force to the left, due to the electric field; this force has magnitude Ee.

When the current first starts to flow, there is no Hall voltage and so an electron will be pushed to the right by the magnetic force. However, as the charge on the right-hand side builds up, so does the electric field and this pushes

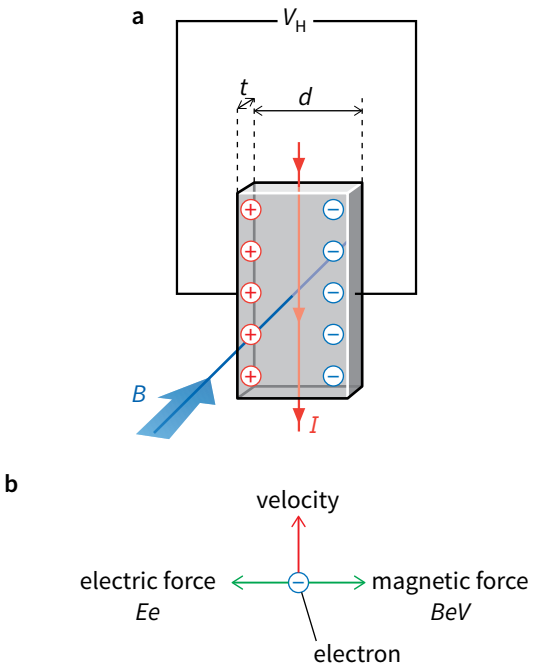


Figure 27.13 **a** The Hall voltage is measured across the slice of semiconductor. **b** The forces on an electron when the electric and magnetic forces on it are balanced.

the electron in the opposite direction to the magnetic force. Soon an equilibrium is reached. The resultant force on this moving electron is zero so that no more charge accumulates. Now we can equate the two forces:

$$eE = Bev$$

Substituting for E we have:

$$\frac{eV_H}{d} = Bev$$

Now recall from Chapter 9 that the current I is related to the mean drift velocity of electrons by $I = nAve$, where A is the cross-sectional area of the conductor and n is the number density of conducting particles (in this case, electrons). So we can substitute for v to get:

$$\frac{eV_H}{d} = \frac{BeI}{nAe}$$

Making V_H the subject of the equation (and cancelling e) gives:

$$V_H = \frac{BId}{nAe}$$

But the area of the side face of the conductor $A = d \times t$, where t is the thickness of the slice. Substituting and cancelling gives:

$$V_H = \frac{BI}{nte}$$

This equation for the Hall voltage shows that V_H is directly proportional to the magnetic flux density B. That is what makes the Hall effect so useful for measuring B fields. To get a large voltage, we want the denominator in this expression to be as small as possible. This is why the Hall probe uses a slice of semiconductor (n is small) and why the slice is thin (t is small).

In some materials, the charges moving are not electrons – for example, they may be positively-charged 'holes'. Consequently we can write a more general equation for the Hall voltage replacing e with q, where q is the charge of an individual charge carrier. This gives

$$V_H = \frac{BI}{ntq}$$

Positive charges will be deflected in the opposite direction to negative charges, and so we can determine whether the charge carriers are positive or negative by the sign of the Hall voltage.

QUESTIONS

8 A Hall probe is designed to operate with a steady current of 0.020 A flowing through a semiconductor slice of thickness 0.05 mm. The number density of electrons in the semiconductor is $1.5 \times 10^{23}\,m^{-3}$.

a Determine the Hall voltage which will result when the probe is placed in a magnetic field of flux density 0.10 T. (Electron charge = 1.60×10^{-19} C.)

b Explain why the current in the Hall probe must be maintained at a constant value.

9 Suggest how the Hall effect could be used to determine the number density of conducting charges in a semiconducting material.

Discovering the electron

Today, a great deal is known about electrons and we use the idea of electrons to explain all sorts of phenomena, including electric current and chemical bonding. However, at the end of the 19th century, physicists were only just beginning to identify the tiny particles which make up matter.

One of the leaders in this field was the English physicist J.J. Thomson (Figure 27.14). In the photograph he is shown with the deflection tube which he used in his discovery of the electron. His tube was similar in construction to the deflection tube shown in Figure 27.9. At one end was an electron gun that produced a beam of electrons (which he called 'cathode rays'). Two metal plates allowed him to



Figure 27.14 J.J. Thomson – in 1897 he discovered the electron using the vacuum tube shown here.

apply an electric field to deflect the beam, and he could place magnets outside the tube to apply a magnetic force to the beam. Here is a summary of his observations and what he concluded from them:

- The beam in his tube was deflected towards a positive plate and away from a negative plate, so the particles involved must have negative charge. This was confirmed by the deflection of the beam by a magnetic field.
- When the beam was deflected, it remained as a tight, single beam rather than spreading out into a broad beam. This showed that, if the beam consisted of particles, they must all have the same mass, charge and speed. (Lighter particles would have been deflected more than heavier ones; particles with greater charge would be deflected more; and faster particles would be deflected less.)
- By applying both electric and magnetic fields, Thomson was able to balance the electric and magnetic forces so that the beam in the tube remained straight. He could then calculate the charge-to-mass ratio $\frac{e}{m_e}$ for the particles he had discovered. Although he did not know the value of either e or m_e individually, he was able to show that the particles concerned must be much lighter than atoms. They were the particles which we now know as electrons. In fact, for a while, Thomson thought that atoms were made up of thousands of electrons, although his ideas could not explain how so many negatively charged particles could combine to produce a neutral atom.

The charge e of an electron is very small (1.60×10^{-19} C) and difficult to measure. The American physicist Robert Millikan devised an ingenious way to do it. He observed electrically charged droplets of oil as they moved in electric and gravitational fields and found that they all had a charge which was a small integer multiple of a particular value, which he took to be the charge on a single electron, e. Having established a value for e, he could easily combine this with Thomson's value for e/m_e to calculate the electron mass m_e.

QUESTION

10 If the electron charge is 1.60×10^{-19} C and the charge-to-mass ratio $\frac{e}{m_e}$ is 1.76×10^{11} C kg^{-1}, calculate the electron mass.

Summary

- The magnetic force on a moving charged particle is given by the equation $F = BQv$. For an electron the equation is $F = Bev$.
- A charged particle entering at right angles to a uniform magnetic field describes a circular path because the magnetic force is perpendicular to the velocity.
- The equation for an electron travelling in a uniform magnetic field is:

  $$\frac{m_e v^2}{r} = Bev$$

- The velocity of an undeflected charged particle in a region where electric and magnetic fields are at right angles is given by the equation:

  $$v = \frac{E}{B}$$

- The Hall voltage is given by

  $$V_H = \frac{BI}{ntq}$$

End-of-chapter questions

1 The magnetic force BQv causes an electron to travel in a circle in a uniform magnetic field. Explain why this force does not cause an increase in the speed of the electron. [3]

2 An electron beam is produced from an electron gun in which each electron is accelerated through a p.d. of 1.6 kV. When these electrons pass at right angles through a magnetic field of flux density 8.0 mT, the radius of curvature of the electron beam is 0.017 m. Determine the specific charge of the electron, $\frac{e}{m_e}$. [4]

3 Two particles, an α-particle and a β^--particle, are travelling through a uniform magnetic field. They have the same velocity and their velocities are at right angles to the field. Determine the ratio of:

a the masses of the two particles [2]

b the charges on the two particles [2]

c the forces created by the magnetic field on the two particles [2]

d the radii of the circular orbits of the two particles in the magnetic field. [2]

4 A moving charged particle experiences a force in an electric field and also in a magnetic field. State **two** differences between the forces experienced in the two types of field. [2]

5 Figure 27.15 shows the path of an electron as it travels in air. The electron rotates clockwise around a uniform magnetic field into the plane of the paper, but the radius of the orbit decreases in size.

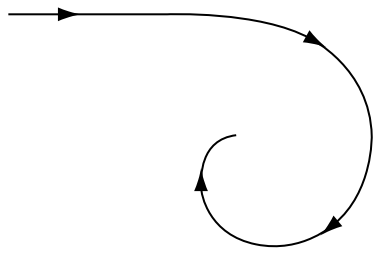


Figure 27.15 For End-of-chapter Question 5.

a i Explain the origin of the force that causes the electron to spiral in this manner. [2]

ii Explain why the radius of the circle gradually decreases. [2]

b At one point in the path, the speed of the electron is $1.0 \times 10^7\ \text{m s}^{-1}$ and the magnetic flux density is 0.25 T. Calculate:

i the force on an electron at this point due to the magnetic field [2]

ii the radius of the path at this point. [2]

6 Figure 27.16 shows an arrangement to deflect protons from a source to a detector using a magnetic field. A uniform magnetic field exists only within the area shown. Protons move from the source to the detector in the plane of the paper.

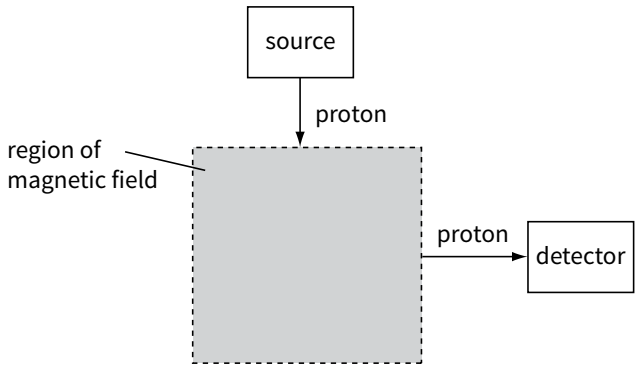


Figure 27.16 For End-of-chapter Question 6.

a **i** Copy the diagram and sketch the path of a proton from the source to the detector. Draw an arrow at **two** points on the path to show the direction of the force on the proton produced by the magnetic field. [3]

ii State the direction of the magnetic field within the area shown. [1]

b The speed of a proton as it enters the magnetic field is $4.0 \times 10^6\,\text{m s}^{-1}$. The magnetic flux density is 0.25 T. Calculate:

i the magnitude of the force on the proton caused by the magnetic field [1]

ii the radius of curvature of the path of the proton in the magnetic field. [2]

c Two changes to the magnetic field in the area shown are made. These changes allow an electron with the same speed as the proton to be deflected along the same path as the proton. State the **two** changes made. [2]

7 Figure 27.17 shows a thin strip of semiconductor material carrying a current in a magnetic field at right angles to the current.

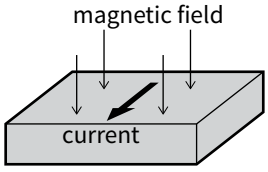


Figure 27.17 For End-of-chapter Question 7.

a The current in the strip is due to the movement of free electrons.

i Add + and – signs to the diagram to show the charge separation caused by the Hall effect. Explain why the charges separate. [3]

ii Explain how an electron is able to move in a straight line along the strip. [1]

b The Hall voltage is measured using the same slice of semiconductor, the same current and the same magnetic field, but with the laboratory at two temperatures, one significantly higher than the other. Describe and explain the changes in the magnitude of the number density, the drift velocity of the charge carriers and the Hall voltage in the two experiments. [6]

8 Figure 27.18 shows an electron tube. Electrons emitted from the cathode accelerate towards the anode and then pass into a uniform electric field created by two oppositely charged parallel metal plates.

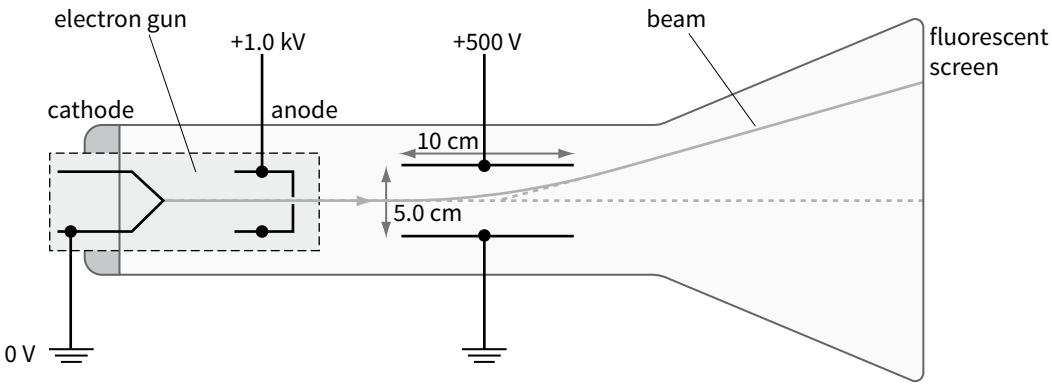


Figure 27.18 For End-of-chapter Question 8.

a **i** Explain why the beam curves upwards. [2]

ii Explain how the pattern formed on the fluorescent screen shows that all the electrons have the same speed as they leave the anode. [2]

b Write down an equation relating the speed of the electrons v to the potential difference V_{ac} between the anode and the cathode. [1]

c The deflection of the beam upwards can be cancelled by applying a suitable uniform magnetic field in the space between the parallel plates.

i State the direction of the magnetic field for this to happen. [1]

ii Write down an equation relating the speed of the electrons v, the electric field E that exists between the plates and the magnetic flux density B needed to make the electrons pass undeflected between the plates. [2]

iii Calculate the value of B required, using the apparatus shown in the diagram, given that the specific charge on an electron $\frac{e}{m_e}$ is $1.76 \times 10^{11}\,\text{C kg}^{-1}$. [2]

9 Protons and helium nuclei from the Sun pass into the Earth's atmosphere above the poles, where the magnetic flux density is $6.0 \times 10^{-5}\,\text{T}$. The particles are moving at a speed of $1.0 \times 10^{6}\,\text{m s}^{-1}$ at right angles to the magnetic field in this region. The magnetic field can be assumed to be uniform.

a Calculate the radius of the path of a proton as it passes above the Earth's pole. [3]

b Draw a diagram to show the deflection caused by the magnetic field to the paths of a proton and of a helium nucleus which both have the same initial velocity as they enter the magnetic field. State on the diagram the radius of the path of each particle. [2]

Mass of a helium nucleus = $6.8 \times 10^{-27}\,\text{kg}$

Charge on a helium nucleus = $3.2 \times 10^{-19}\,\text{C}$

10 Figure 27.19 shows a thin strip of metal of thickness t and width d. The metal strip is in a magnetic field of flux density B and carries a current I, as shown.

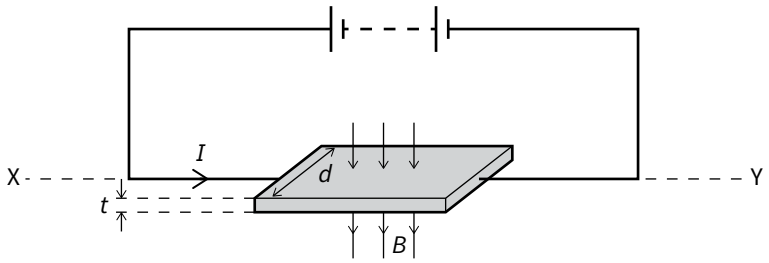


Figure 27.19 For End-of-chapter Question 10.

a Copy Figure 27.19 and mark on your diagram:

i the side of the strip that becomes negative because of the Hall effect [1]

ii where a voltmeter needs to be placed to measure the Hall voltage. [1]

b Derive an expression for the Hall voltage in terms of I, B, t, the number density n of free electrons in the metal and the charge e on an electron. [3]

c Given that $I = 40\,\text{mA}$, $d = 9.0\,\text{mm}$, $t = 0.030\,\text{mm}$, $B = 0.60\,\text{T}$, $e = 1.6 \times 10^{-19}\,\text{C}$ and $n = 8.5 \times 10^{28}\,\text{m}^{-3}$, calculate:

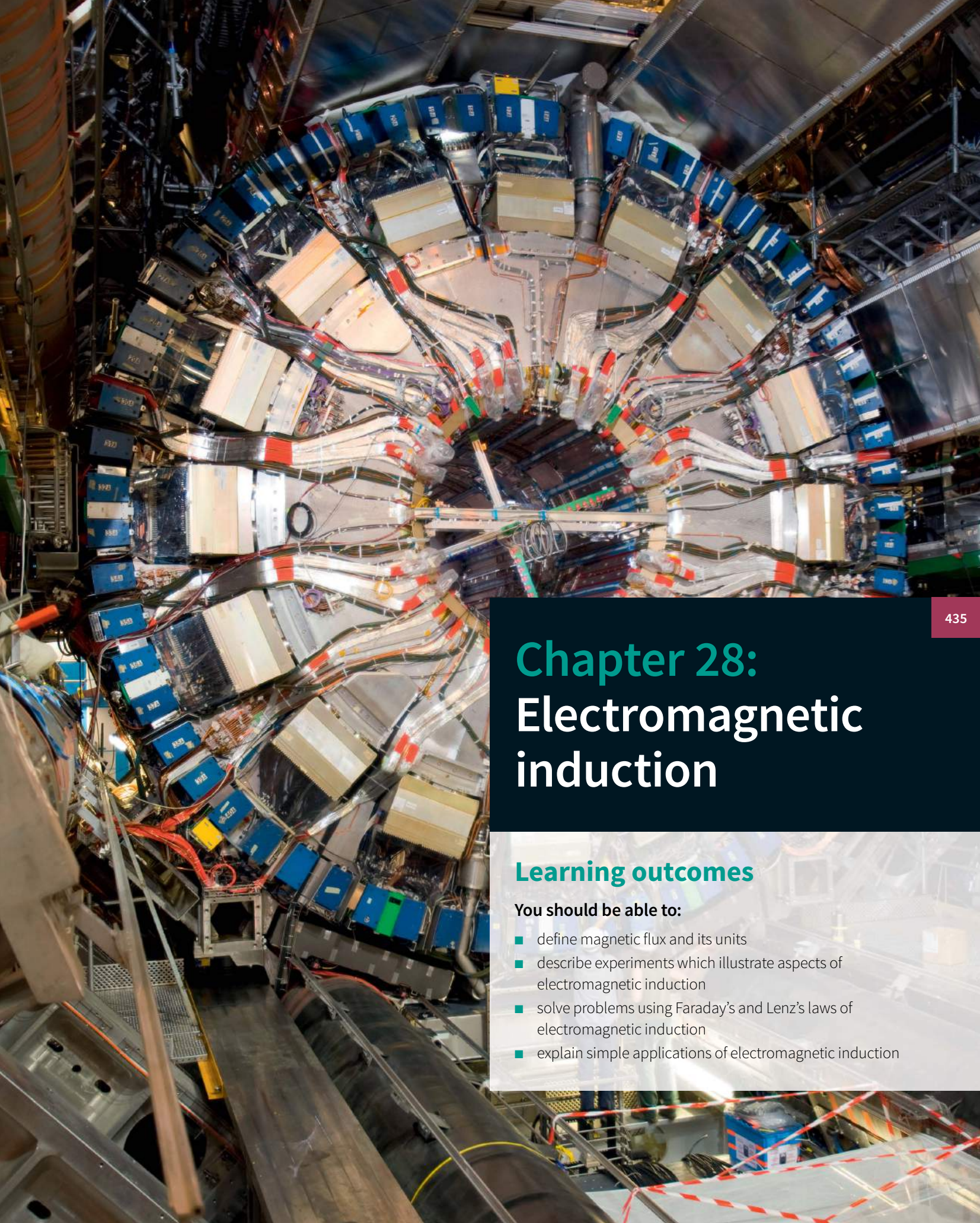
i the mean drift velocity v of the free electrons in the metal [2]

ii the Hall voltage across the metal strip. [2]

d **i** Explain why, in terms of the movement of electrons, the Hall voltage increases when I increases. [2]

ii A Hall probe which measures the strength of a magnetic field uses a thin strip of a semiconductor rather than metal. Explain why a semiconductor is used. [2]

e Explain why, when the strip of metal is rotated about the horizontal axis XY, the Hall voltage varies between a maximum positive value and a maximum negative value. [2]



Chapter 28: Electromagnetic induction

Learning outcomes

You should be able to:

- define magnetic flux and its units
- describe experiments which illustrate aspects of electromagnetic induction
- solve problems using Faraday's and Lenz's laws of electromagnetic induction
- explain simple applications of electromagnetic induction

Generating electricity

Most of the electricity we use is generated by electromagnetic induction. This process goes on in the generators at work in power stations, in wind turbines (Figure 28.1) and, on a much smaller scale, in bicycle dynamos. It is the process whereby a conductor and a magnetic field are moved relative to each other to induce, or generate, a current or electromotive force (e.m.f.).



Figure 28.1 This giant wind turbine uses electromagnetic induction to produce electricity. Look for the two engineers at work. (You can identify them by their white helmets.) This gives you an idea of the size of the generator.

Observing induction

You can carry out some simple experiments in which you can observe some of the features of electromagnetic induction. These are described in Box 28.1.

BOX 28.1: Observing induction

For each experiment, try to predict what you will observe before you try the experiment.

Experiment 1

Connect a small electric motor to a moving-coil voltmeter (Figure 28.2). Spin the shaft of the motor and observe the deflection of the voltmeter. What happens when you spin the motor more slowly? What happens when you stop? Usually, we connect a motor to a power supply and it turns. In this experiment, you have turned the motor and it generates a voltage across its terminals. A generator is like a motor working in reverse.

Experiment 2

Connect a coil to a sensitive microammeter (Figure 28.3). Move a bar magnet in towards the coil. Hold it still, and then remove it. How does the deflection on the meter change? Try different speeds, and the opposite pole of the magnet. Try weak and strong magnets.

With the same equipment, move the coil towards the magnet and observe the deflection of the meter.

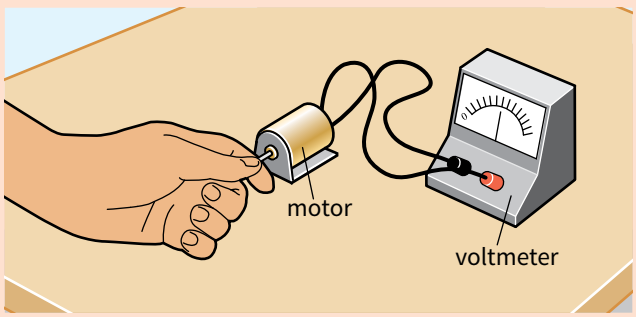


Figure 28.2 A motor works in reverse as a generator.

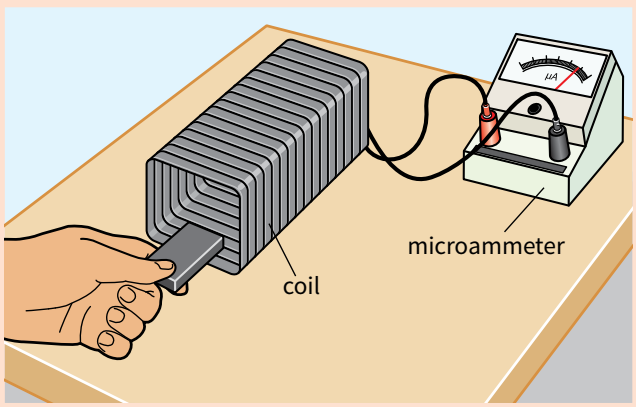


Figure 28.3 A magnet moving near a coil generates a small current.

BOX 28.1: Observing induction (continued)

Experiment 3

Connect a long wire to a sensitive microammeter. Move the middle section of the wire up and down through the magnetic field between the magnets (Figure 28.4). Double up the wire so that twice as much of it passes through the magnetic field. What happens to the meter reading? How can you form the wire into a loop to give twice the deflection on the meter?

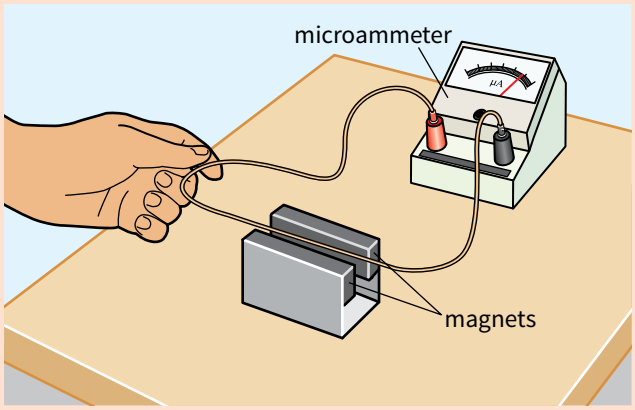


Figure 28.4 Investigating the current induced when a wire moves through a magnetic field.

Factors affecting induced current

In all the experiments described in Box 28.1, you have seen an electric current or an e.m.f. induced. In each case, there is a magnetic field and a conductor. When you move the magnet or the conductor, there is an induced current. When you stop, the current stops.

From the three experiments, you should see that the size of the induced current or e.m.f. depends on several factors.

For a straight wire, the induced current or e.m.f. depends on:

- the magnitude of the magnetic flux density
- the length of the wire in the field
- the speed of movement of the wire.

For a coil of wire, the induced current or e.m.f. depends on:

- the magnitude of the magnetic flux density
- the cross-sectional area of the coil
- the number of turns of wire
- the rate at which the coil turns in the field.

Explaining electromagnetic induction

You have seen that relative movement of a conductor and a magnetic field induces a current in the conductor when it is part of a complete circuit. (In the experiments in Box 28.1, the meter was used to complete the circuit.) Now we need to think about how to explain these observations, using what we know about magnetic fields.

Cutting magnetic field lines

Start by thinking about a simple bar magnet. It has a magnetic field in the space around it. We represent this field by magnetic field lines. Now think about what happens when a wire is moved into the magnetic field (Figure 28.5). As it moves, it **cuts across** the magnetic field. Remove the wire from the field, and again it must cut across the field lines, but in the opposite direction.

We think of this cutting of a magnetic field by a conductor as the effect that gives rise to an induced current in the conductor. It doesn't matter whether the conductor is moved through the field or the magnet is moved past the conductor, the result is the same – there will be an induced current.

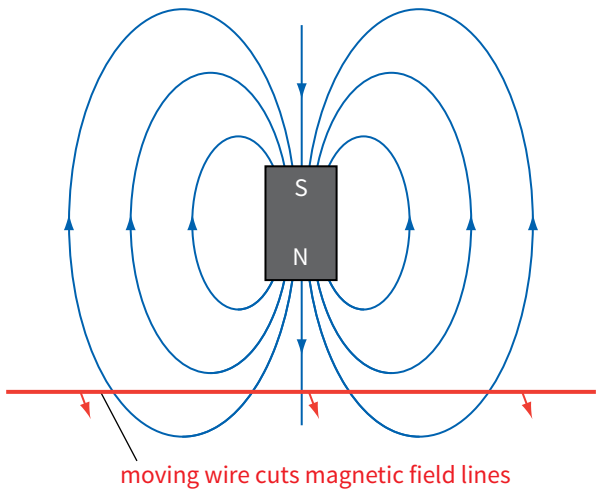


Figure 28.5 Inducing a current by moving a wire through a magnetic field.

The effect is magnified if we use a coil of wire. For a coil of N turns, the effect is N times greater than for a single turn of wire. With a coil, it is helpful to imagine the number of field lines **linking** the coil. If there is a change in the number of field lines which pass through the coil, an e.m.f. will be induced across the ends of the coil (or there will be an induced current if the coil forms part of a complete circuit).

Figure 28.6 shows a coil near a magnet. When the coil is outside the field, there are no magnetic field lines linking the coil. When it is inside the field, field lines link the coil. Moving the coil into or out of the field changes this linkage, and this induces an e.m.f. across the ends of the coil.

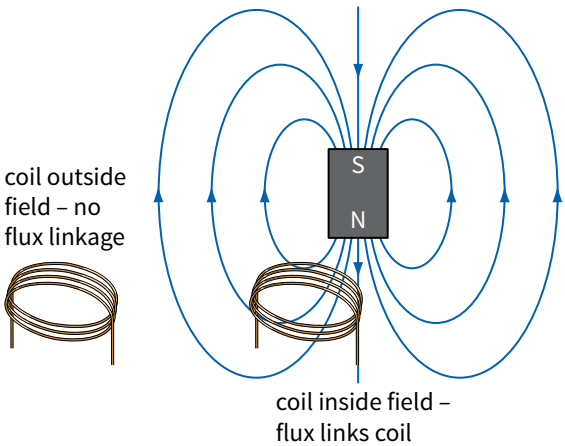


Figure 28.6 The flux passing through a coil changes as it is moved into and out of a magnetic field.

QUESTION

1 Use the idea of a conductor cutting magnetic field lines to explain how a current is induced in a bicycle generator (Figure 28.7).

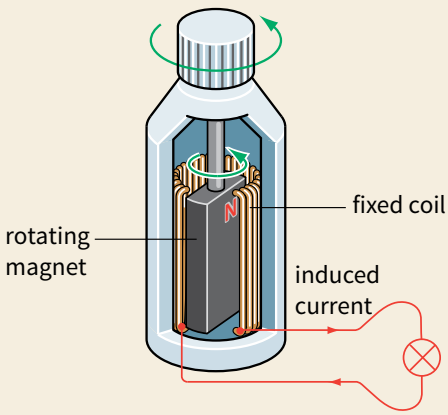


Figure 28.7 In a bicycle generator, a permanent magnet rotates inside a fixed coil of wire. For Question 1.

Current direction

How can we predict the direction of the induced current? For the motor effect in Chapter 27, we used Fleming's left-hand (motor) rule. Electromagnetic induction is like the mirror image of the motor effect. Instead of a current producing a force on a current-carrying conductor in a magnetic field, we provide an external force on a conductor by moving it through a magnetic field and this induces a current in the conductor. So you should not be too surprised to find that we use the mirror image of the left-hand rule: **Fleming's right-hand (generator) rule**.

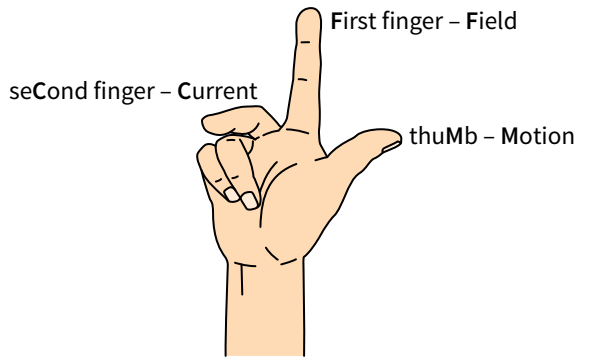


Figure 28.8 Fleming's right-hand (generator) rule.

The three fingers represent the same things again (Figure 28.8):

- thuMb – direction of Motion
- First finger – direction of external magnetic Field
- seCond finger – direction of (conventional) induced Current

In the example shown in Figure 28.9, the conductor is being moved downwards across the magnetic field. There is an induced current in the conductor as shown. Check this with your own right hand. You should also check that reversing the movement or the field will result in the current flowing in the opposite direction.

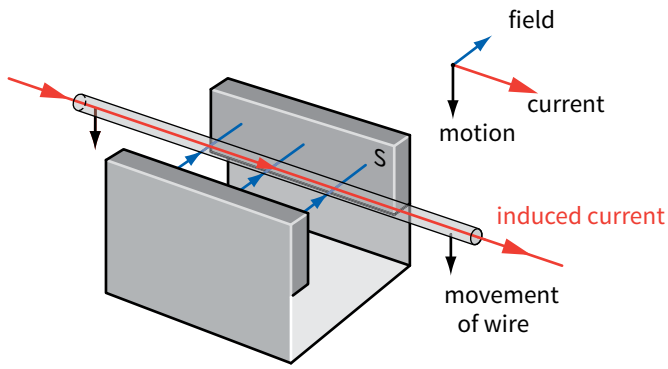


Figure 28.9 Deducing the direction of the induced current using Fleming's right-hand rule.

Induced e.m.f.

When a conductor is not part of a complete circuit, there cannot be an induced current. Instead, negative charge will accumulate at one end of the conductor, leaving the other end positively charged. We have induced an e.m.f. across the ends of the conductor.

Is e.m.f. the right term? Should it be voltage? In Chapter 9 you saw the distinction between voltage and e.m.f. The latter is the correct term here because, by pushing the wire through the magnetic field, work is done and this is transformed into electrical energy. Think of this in another way. Since we could connect the ends of the conductor so that there is a current in some other component, such as a lamp, which would light up, it must be an e.m.f. – a source of electrical energy.

Figure 28.10 shows how the induced current gives rise to an induced e.m.f. Notice that, within the conductor, conventional current is from negative to positive, in the same way as inside a battery or any other source of e.m.f. In reality, the free electrons within the conductor travel from right to left, making the left-hand side of the conductor negative. What causes these electrons to move? Moving the conductor is equivalent to giving an electron within the conductor a velocity in the direction of this motion. This electron is in an external magnetic field and hence experiences a magnetic force Bev from right to left. Check this out for yourself.

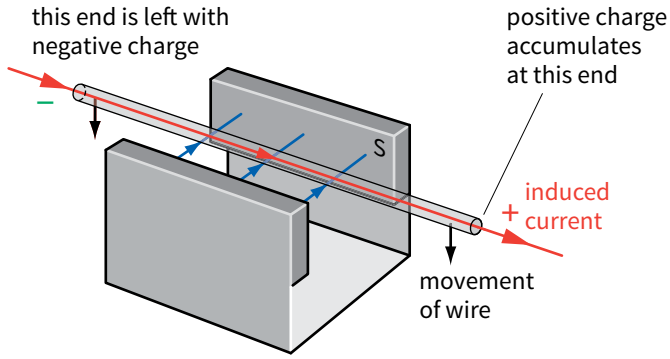


Figure 28.10 An e.m.f. is induced across the ends of the conductor.

QUESTIONS

2 The coil in Figure 28.11 is rotating in a uniform magnetic field. Deduce the direction of the induced current in sections AB and CD. State which terminal, X or Y, will become positive.

3 When an aircraft flies from east to west, its wings are an electrical conductor cutting across the Earth's magnetic flux. In the northern hemisphere, which wingtip will become positively charged? Why will this wingtip be negative in the southern hemisphere?

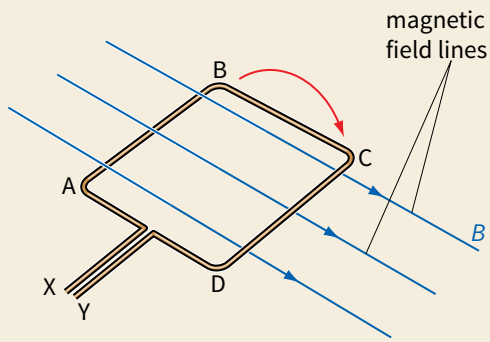


Figure 28.11 A coil rotating in a magnetic field.

Magnetic flux and magnetic flux linkage

So far in this chapter we have looked at the ideas of electromagnetic induction in a descriptive way. Now we will see how to calculate the value of the induced e.m.f. and look at a general way of determining its direction.

In Chapter 26, we saw how magnetic flux density B is defined by the equation

$$B = \frac{F}{IL}$$

Now we can go on to define **magnetic flux** as a quantity. We picture **magnetic flux density** B as the number of magnetic field lines passing through a region **per unit area**. Similarly, we can picture magnetic flux as the total number of magnetic field lines passing through an area A. For a magnetic field normal to A, the magnetic flux Φ (Greek letter phi) must therefore be equal to the product of magnetic flux density and the area A (Figure 28.12a).

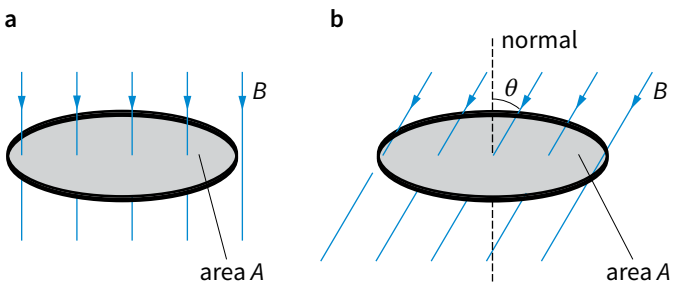


Figure 28.12 **a** The magnetic flux is equal to BA when the field is normal to the area. **b** The magnetic flux becomes $BA\cos\theta$ when the field is at an angle θ to the normal of the area.

The magnetic flux Φ through area A is defined as:

$$\Phi = BA$$

where B is the component of the magnetic flux density perpendicular to the area.

How can we calculate the magnetic flux when B is not perpendicular to A? You can easily see that when the field is parallel to the plane of the area, the magnetic flux through A is zero. To find the magnetic flux in general, we need to find the component of the magnetic flux density perpendicular to the area. Figure 28.12b shows a magnetic field at an angle θ to the normal. In this case:

$$\text{magnetic flux} = (B\cos\theta) \times A$$

or simply:

$$\text{magnetic flux} = BA\cos\theta$$

(Note that, when $\theta = 90°$, flux = 0 and when $\theta = 0°$, flux = BA.)

For a coil with N turns, the **magnetic flux linkage** is defined as the product of the magnetic flux and the number of turns; that is:

$$\text{magnetic flux linkage} = N\Phi$$

or

$$\text{magnetic flux linkage} = BAN\cos\theta$$

The unit for magnetic flux or flux linkage is the weber (Wb).

One weber (1 Wb) is the flux that passes through an area of 1 m^2 when the magnetic flux density is 1 T.
1 Wb = 1 T m^2.

An e.m.f. is induced in a circuit whenever there is a **change** in the magnetic flux linking the circuit. Since magnetic flux is equal to $BA\cos\theta$, there are three ways an e.m.f. can be induced:

- changing the magnetic flux density B
- changing the area A of the circuit
- changing the angle θ.

Now look at Worked example 1.

WORKED EXAMPLE

1 Figure 28.13 shows a solenoid with a cross-sectional area 0.10 m^2. It is linked by a magnetic field of flux density 2.0×10^{-3} T and has 250 turns. Calculate the magnetic flux and flux linkage for this solenoid.

Step 1 We have $B = 2.0 \times 10^{-3}$ T, $A = 0.10\ m^2$, $\theta = 0°$ and $N = 250$ turns. Hence we can calculate the flux Φ.

$\Phi = BA$

$\Phi = 2.0 \times 10^{-3} \times 0.10 = 2.0 \times 10^{-4}$ Wb

Step 2 Now calculate the flux linkage.

magnetic flux linkage = $N\Phi$

magnetic flux linkage = $2.0 \times 10^{-4} \times 250$

$= 5.0 \times 10^{-2}$ Wb

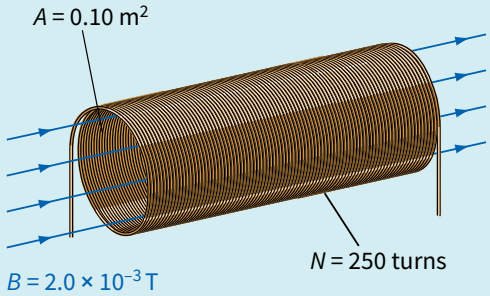


Figure 28.13 A solenoid in a magnetic field.

QUESTIONS

4 Use the idea of magnetic flux linkage to explain why, when a magnet is moved into a coil, the e.m.f. induced depends on the strength of the magnet and the speed at which it is moved.

5 In an experiment to investigate the factors that affect the magnitude of an induced e.m.f., a student moves a wire back and forth between two magnets, as shown in Figure 28.14. Explain why the e.m.f. generated in this way is much smaller than if the wire is moved up and down in the field.

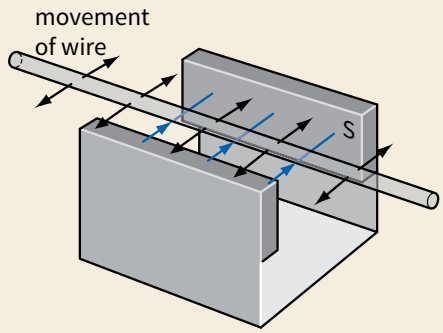


Figure 28.14 A wire is moved horizontally in a horizontal magnetic field. For Question 5.

QUESTIONS

6 In the type of generator found in a power station (Figure 28.15), a large electromagnet is made to rotate inside a fixed coil. An e.m.f. of 25 kV is generated; this is an alternating voltage of frequency 50 Hz. What factor determines the frequency? What factors do you think would affect the magnitude of the e.m.f.?



Figure 28.15 The generators of this power station produce electricity at an induced e.m.f. of 25 kV. For Question 6.

7 A bar magnet produces a uniform flux density of 0.15 T at the surface of its north pole. The pole measures 1.0 cm × 1.5 cm. Calculate the magnetic flux at this pole.

8 A solenoid has diameter 5.0 cm and length 25 cm (Figure 28.16). There are 200 turns of wire. A current of 2.0 A creates a magnetic field of flux density 2.0×10^{-5} T through the core of this solenoid. Calculate the magnetic flux linkage for this solenoid.

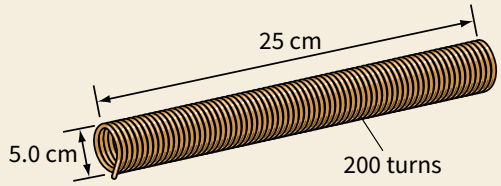


Figure 28.16 A solenoid. For Question 8.

9 A rectangular coil, 5.0 cm × 7.5 cm, and having 120 turns, is at right angles to a magnetic field of flux density 1.2 T. Calculate the magnetic flux linkage for this coil.

Faraday's law of electromagnetic induction

Earlier in this chapter, we saw that electromagnetic induction occurs whenever a conductor cuts across lines of magnetic flux – for example, when a coil is rotated in a magnetic field so that the magnetic flux linking the coil changes. We can use **Faraday's law of electromagnetic induction** to determine the magnitude of the induced e.m.f. in a circuit:

> The magnitude of the induced e.m.f. is proportional to the rate of change of magnetic flux linkage.

We can write this mathematically as:

$$E \propto \frac{\Delta(N\Phi)}{\Delta t}$$

where $\Delta(N\Phi)$ is the change in the flux linkage in a time Δt. Working in SI units, the constant of proportionality is equal to 1. Therefore:

$$E = \frac{\Delta(N\Phi)}{\Delta t}$$

The equation above is a mathematical statement of Faraday's law. Note that it allows us to calculate the *magnitude* of the induced e.m.f.; its *direction* is given by Lenz's law, which is explained in the next section on page 442.

Now look at Worked examples 2 and 3.

WORKED EXAMPLES

2 A straight wire of length 0.20 m moves at a steady speed of 3.0 m s^{-1} at right angles to a magnetic field of flux density 0.10 T. Use Faraday's law to determine the e.m.f. induced across the ends of the wire.

Step 1 With a single conductor, $N = 1$. To determine the e.m.f. E, we need to find the rate of change of magnetic flux; in other words, the change in magnetic flux per second.

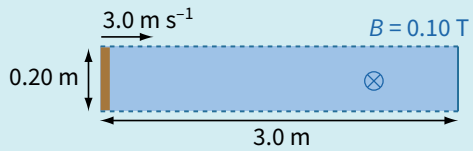


Figure 28.17 A moving wire cuts across the magnetic field.

Figure 28.17 shows that, in 1.0 s, the wire travels 3.0 m. Therefore:

change in magnetic flux = B × change in area

change in magnetic flux = $0.10 \times (3.0 \times 0.20)$

$= 6.0 \times 10^{-2}$ Wb

WORKED EXAMPLES (continued)

Step 2 Use Faraday's law to determine the e.m.f.

$$E = \frac{\Delta(N\Phi)}{\Delta t} \qquad (N = 1)$$

$\Delta\Phi = 6.0 \times 10^{-2}$ Wb and $\Delta t = 1.0$ s

$$E = \frac{6.0 \times 10^{-2}}{1.0} = 0.06\,\text{V}$$

The induced e.m.f. across the ends of the wire is about 60 mV.

3 This example illustrates one way in which the flux density of a magnetic field can be measured, shown in Figure 28.18.

A search coil of wire having 2500 turns and of area 1.2 cm^2 is placed between the poles of a magnet so that the magnetic flux passes perpendicularly through the coil. The flux density of the field is 0.50 T. The coil is pulled rapidly out of the field in a time of 0.10 s. What average e.m.f. is induced across the ends of the coil?

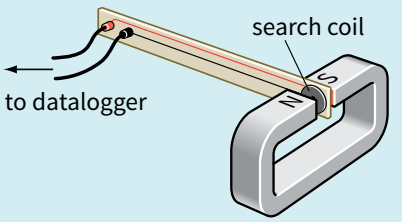


Figure 28.18 A search coil can be moved into and out of a magnetic field to detect magnetic flux.

Step 1 When the coil is pulled from the field, the flux linking it falls to zero. We have to calculate the magnetic flux linking the coil when it is in the field.

To convert cm^2 into m^2, multiply by a factor of 10^{-4}. Hence $A = 1.2 \times 10^{-4}$ m^2.

magnetic flux linkage $= N\Phi = BAN$

$= 0.50 \times 1.2 \times 10^{-4} \times 2500$

magnetic flux linkage = 0.15 Wb

Step 2 Now calculate the induced e.m.f. using Faraday's law of electromagnetic induction.

$\Delta(N\Phi) = 0.15$ Wb and $\Delta t = 0.10$ s

magnitude of induced e.m.f. = rate of change of flux linkage

$$E = \frac{\Delta(N\Phi)}{\Delta t} = \frac{0.15}{0.10} = 1.5\,\text{V}$$

Note that, in this example, we have assumed that the flux linking the coil falls steadily to zero during the time interval of 0.10 s. Our answer is thus the average value of the e.m.f.

QUESTIONS

10 A conductor of length L moves at a steady speed v at right angles to a uniform magnetic field of flux density B. Show that the e.m.f. E across the ends of the conductor is given by the equation:

$E = BLv$

(You can use Worked example 2 to guide you through Question 11.)

11 A wire of length 10 cm is moved through a distance of 2.0 cm in a direction at right angles to its length in the space between the poles of a magnet, and perpendicular to the magnetic field. The flux density is 1.5 T. If this takes 0.50 s, calculate the average e.m.f. induced across the ends of the wire.

12 Figure 28.19 shows a search coil, having 2000 turns and of area 1.2 cm^2, placed between the poles of a strong magnet. The ends of the coil are connected to a voltmeter. The coil is then pulled out of the magnetic field, and the voltmeter records an average e.m.f. of 0.40 V over a time interval of 0.20 s. Calculate the magnetic flux density between the poles of the magnet.

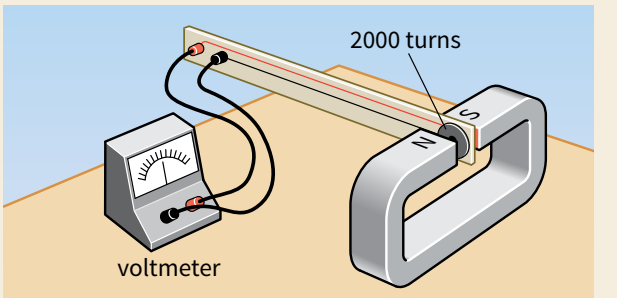


Figure 28.19 Using a search coil to measure flux.

Lenz's law

We use Faraday's law to calculate the magnitude of an induced e.m.f. Now we can go on to think about the direction of the e.m.f. – in other words, which end of a wire or coil moving in a magnetic field becomes positive, and which becomes negative.

Fleming's right-hand rule gives the direction of an induced current. This is a particular case of a more general law, Lenz's law, which will be explained in this section. First, we will see how the motor effect and the generator effect are related to each other.

The origin of electromagnetic induction

So far, we have not given an explanation of electromagnetic induction. You have seen, from the experiments at the beginning of this chapter, that it does occur, and you know the factors that affect it. But what is the origin of the induced current?

Figure 28.20 gives an explanation. A straight wire XY is being pushed downwards through a horizontal magnetic field of flux density *B*. Now, think about the free electrons in the wire. They are moving downwards, so they are in effect an electric current. Of course, because electrons are negatively charged, the conventional current is flowing upwards.

We now have a current flowing across a magnetic field, and the motor effect will therefore come into play. Each electron experiences a force of magnitude *Bev*. Using Fleming's left-hand rule, we can find the direction of the force on the electrons. The diagram shows that the electrons will be pushed in the direction from X to Y. So a current has been induced to flow in the wire; the direction of the conventional current is from Y to X.

Now we can check that Fleming's right-hand rule gives the correct directions for motion, field and current, which indeed it does.

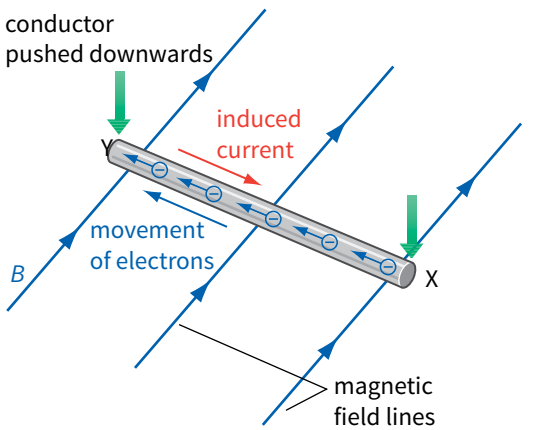


Figure 28.20 Showing the direction of the induced current.

So, to summarise, there is an induced current because the electrons are pushed by the motor effect. Electromagnetic induction is simply a consequence of the motor effect.

In Figure 28.20, electrons are found to accumulate at Y. This end of the wire is thus the negative end of the e.m.f. and X is positive. If the wire was connected to an external circuit, electrons would flow out of Y, round the circuit, and back into X. Figure 28.21 shows how the moving wire is equivalent to a cell (or any other source of e.m.f.).

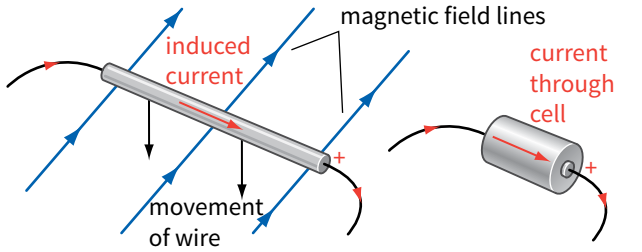


Figure 28.21 A moving conductor in a magnetic field is a source of e.m.f., equivalent to a cell.

Forces and movement

Electromagnetic induction is how we generate most of our electricity. We turn a coil in a magnetic field, and the mechanical energy we put in is transferred to electrical energy. By thinking about these energy transfers, we can deduce the direction of the induced current.

Figure 28.22 shows one of the experiments from earlier in this chapter. The north pole of a magnet is being pushed towards a coil of wire. There is an induced current in the coil, but what is its direction? The diagram shows the two possibilities.

The current in the coil turns it into an electromagnet. One end becomes the north pole, the other the south pole. In Figure 28.22a, if the induced current is in this direction, the coil end nearest the approaching north pole of the

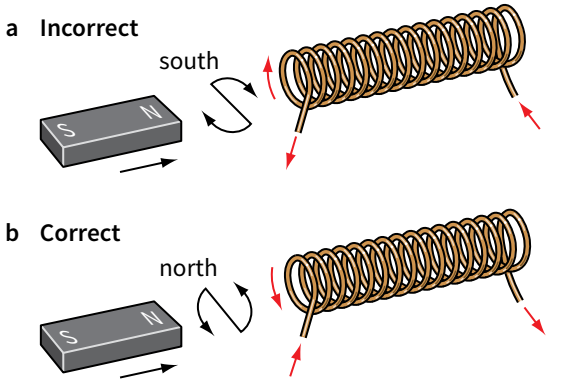


Figure 28.22 Moving a magnet towards a coil: the direction of the induced current is as shown in **b**, not **a**.

magnet would be a south pole. These poles will attract one another, and you could let go of the magnet and it would be dragged into the coil. The magnet would accelerate into the coil, the induced current would increase further, and the force of attraction between the two would also increase.

In this situation, we would be putting no energy into the system, but the magnet would be gaining kinetic energy, and the current would be gaining electrical energy. A nice trick if you could do it, but against the principle of conservation of energy!

It follows that Figure 28.22b must show the correct situation. As the north pole of the magnet is pushed towards the coil, the induced current makes the end of the coil nearest the magnet become a north pole. The two poles repel one another, and you have to do work to push the magnet into the coil. The energy transferred by your work is transferred to electrical energy of the current. The principle of energy conservation is not violated.

QUESTION

13 Use these ideas to explain what happens if **a** you stop pushing the magnet towards the coil, and **b** you pull the magnet away from the coil.

Figure 28.23 shows how we can apply the same reasoning to a straight wire being moved in a downward direction through a magnetic field. There will be an induced current in the wire, but in which direction? Since this is a case of a current across a magnetic field, a force will act on it (the motor effect), and we can use Fleming's left-hand rule to deduce its direction.

First we will consider what happens if the induced current is in the wrong direction. This is shown in Figure 28.23a. The left-hand rule shows that the force that results would be downward – in the direction in which we are trying to move the wire. The wire would thus be accelerated, the current would increase, and again we would be getting both kinetic and electrical energy for no energy input.

The induced current must be as shown in Figure 28.23b. The force that acts on it due to the motor effect pushes against you as you try to move the wire through the field. You have to do work to move the wire, and hence to generate electrical energy. Once again, the principle of energy conservation is not violated.

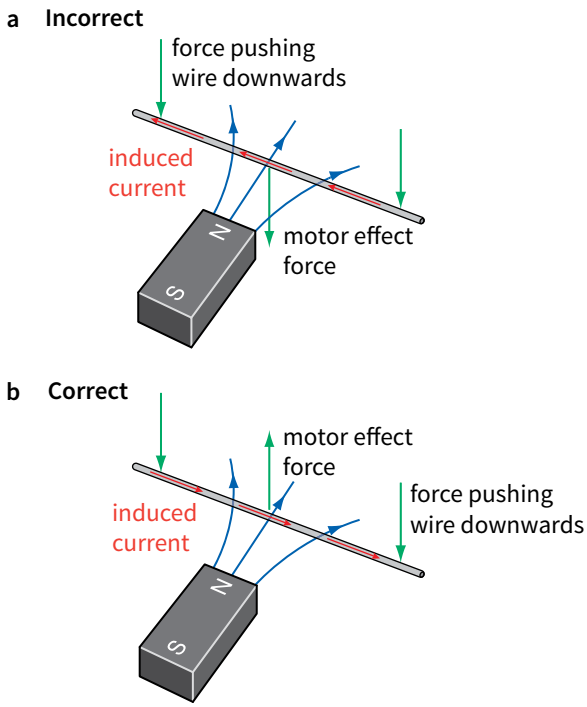


Figure 28.23 Moving a conductor through a magnetic field: the direction of the induced current is as shown in **b**, not **a**.

QUESTION

14 Draw a diagram to show the directions of the induced current and of the opposing force if you now try to move the wire shown in Figure 28.23 upwards through the magnetic field.

A general law for induced e.m.f.

Lenz's law summarises this general principle of energy conservation. The direction of an induced current is such that it always produces a force that opposes the motion that is being used to produce it. If the direction of the current were opposite to this, we would be getting energy for nothing. Here is a statement of **Lenz's law**:

> Any induced current or induced e.m.f. will be established in a direction so as to produce effects which oppose the change that is producing it.

This law can be shown to be correct in any experimental situation. For example, in Figure 28.3, a sensitive ammeter connected in the circuit shows the direction of the current as the magnet is moved in and out. If a battery is later connected to the coil to make a larger and constant current

in the same direction, a compass will show what the poles are at the end of the solenoid. If a north pole is moved into the solenoid, then the solenoid itself will have a north pole at that end. If a north pole is moved out of the solenoid, then the solenoid will have a south pole at that end.

QUESTIONS

15 A bar magnet is dropped vertically downwards through a long solenoid, which is connected to an oscilloscope (Figure 28.24). The oscilloscope trace shows how the e.m.f. induced in the coil varies as the magnet accelerates downwards.

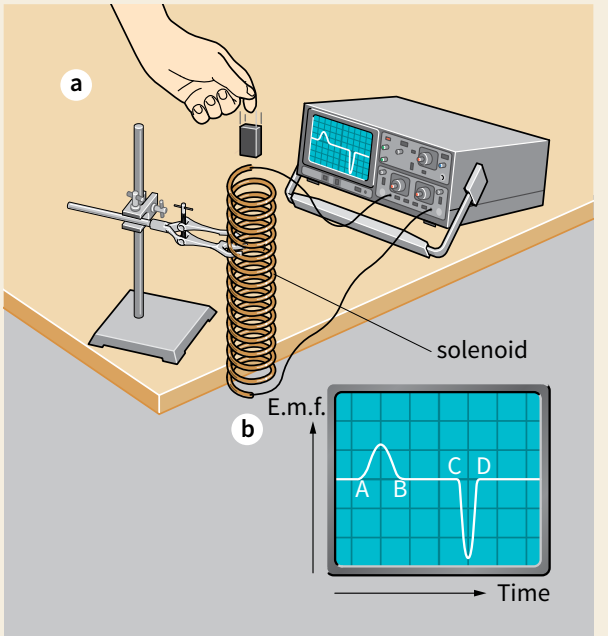


Figure 28.24 **a** A bar magnet falls through a long coil. **b** The oscilloscope trace shows how the induced e.m.f. varies with time.

- **a** Explain why an e.m.f. is induced in the coil as the magnet enters it (section AB of the trace).
- **b** Explain why no e.m.f. is induced while the magnet is entirely inside the coil (section BC).
- **c** Explain why section CD shows a negative trace, why the peak e.m.f. is greater over this section, and why CD represents a shorter time interval than AB.

16 You can turn a bicycle dynamo by hand and cause the lamps to light up. Use the idea of Lenz's law to explain why it is easier to turn the dynamo when the lamps are switched off than when they are on.

Using induction: eddy currents, generators and transformers

An induced e.m.f. can be generated in a variety of ways. What they all have in common is that a conductor is cutting across magnetic field lines (in some cases, the conductor moves; in others, the field lines move). The alternative way to look at any change is to say that the flux linking an area changes.

Eddy currents

Induced e.m.f.s are formed in some unexpected places. Consider the demonstration shown in Figure 28.25. A metal disc on the end of a rod swings freely between two opposite magnetic poles.

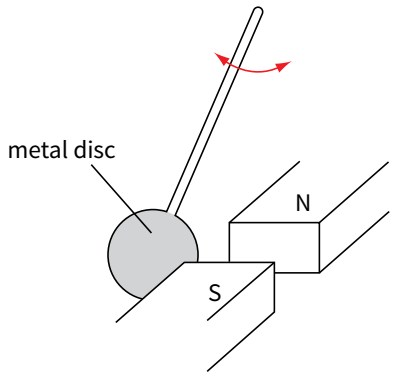


Figure 28.25 Demonstrating eddy current damping.

Without the magnets, the disc oscillates from side to side for a long time. This is because air resistance is small and it takes a long time for the energy of the disc to be lost. When the magnets are present, the oscillation of the disc dies away quickly. As the disc enters the magnetic field, one side of the disc is cutting the magnetic field lines and so an induced e.m.f. is created in that side but not in the side that has not yet entered. Since the disc is a conductor, the induced e.m.f. creates currents in the disc itself. These currents are known as eddy currents. They flow in a circular fashion inside the disc. Lenz's law predicts that the induced currents that flow in the disc will produce a force that opposes the motion, just as in Figure 28.23. Eddy currents, like other electrical currents, cause heating and the energy of the oscillation dies away quickly. The oscillation is damped by the eddy currents.

This principle can be used in some types of electromagnetic or eddy-current braking systems. For example, a large electromagnet suspended under a train can cause eddy currents in the rails and slow the train down. Better still, if the train has an electric motor, then

the kinetic energy of the train can be used to turn the electric motor to generate an induced e.m.f. With the appropriate electronics the energy from the induced current can be passed back to the power supply that runs the train. This is an example of regenerative braking.

Generators

We can generate electricity by spinning a coil in a magnetic field. This is equivalent to using an electric motor backwards. Figure 28.26 shows such a coil in three different orientations as it spins. Notice that the rate of change of flux linkage is maximum when the coil is moving through the horizontal position – one side is cutting rapidly downwards through the field lines, the other is cutting rapidly upwards. In this position, we get a large induced e.m.f. As the coil moves through the vertical position, the rate of change of flux is zero – the sides of the coil are moving parallel to the field lines, not cutting them, so that there is hardly any change in the flux linkage.

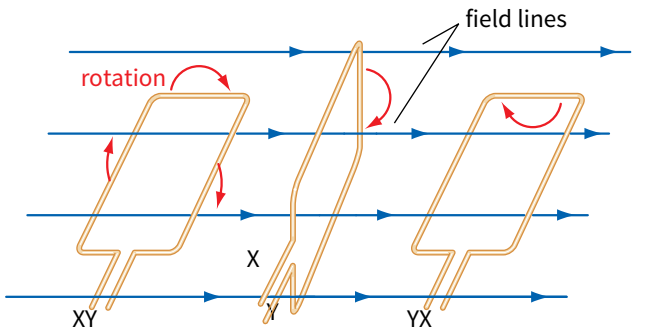


Figure 28.26 A coil rotating in a magnetic field.

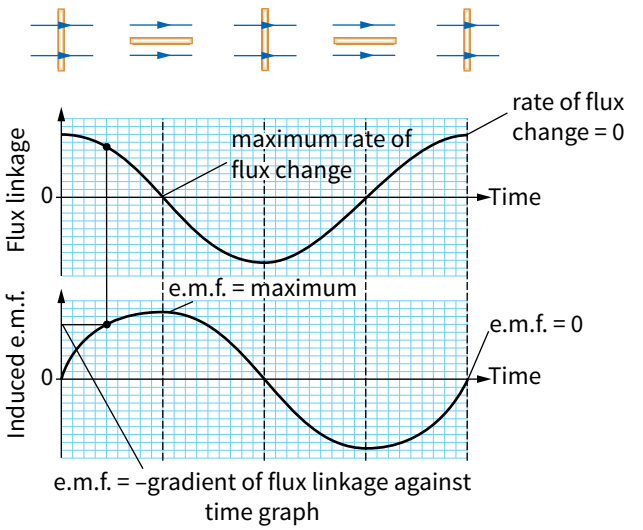


Figure 28.27 The magnetic flux linking a rotating coil as it changes. This gives rise to an alternating e.m.f. The orientation of the coil is shown above the graphs.

Figure 28.27 shows how the flux linkage varies with time for a rotating coil. According to Faraday's law, the induced e.m.f. is equal to minus the gradient of the flux linkage against time graph.

- When the flux linking the coil is maximum, the rate of change of flux is zero and hence the induced e.m.f. is zero.
- When the flux linking the coil is zero, the rate of change of flux is maximum (the graph is steepest) and hence the induced e.m.f. is also maximum.

Hence, for a coil like this we get a varying e.m.f. – this is how alternating current is generated. In practice, it is simpler to keep the large coil fixed and spin an electromagnet inside it (Figure 28.28). A bicycle generator (see Figure 28.7) is similar, but in this case a permanent magnet is made to spin inside a fixed coil. This makes for a very robust device.

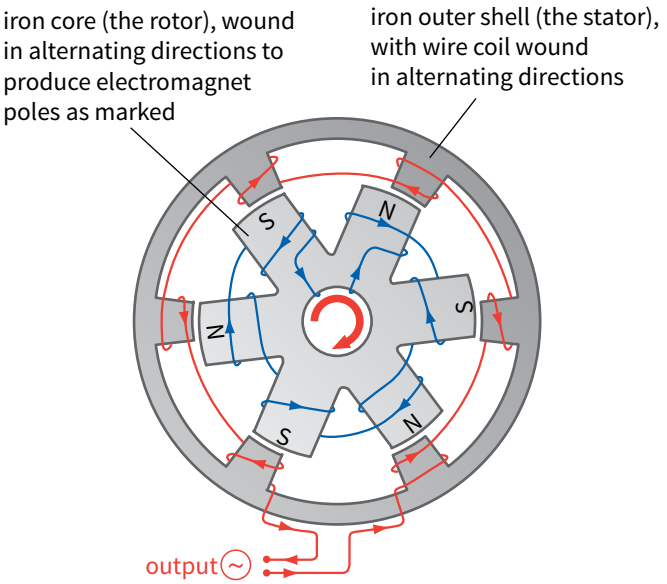


Figure 28.28 In a generator, an electromagnet rotates inside a coil.

Transformers

Another use of electromagnetic induction is in transformers. An alternating current is supplied to the primary coil and produces a varying magnetic field in the soft iron core (Figure 28.30). The secondary coil is also wound round this core, so the magnetic flux linking the secondary coil is constantly changing. Hence, according to Faraday's law, a varying e.m.f. is induced across the secondary coil. The core is laminated – it is made up of thin sheets of soft iron. Using soft iron in the core increases the amount of the magnetic flux and, hopefully, all of the magnetic flux from the primary coil passes to the secondary coil. The thin sheets of iron in the core are

separated by a non-conductor so eddy currents cannot flow from one sheet to the next. This reduces the eddy currents and the thermal energy that they create in the core.

There is much more about transformers in Chapter 29, where we will look at how they are used to change voltages in circuits which make use of alternating currents.

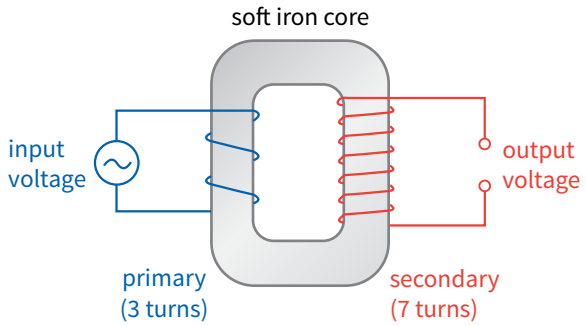


Figure 28.30 The construction of a transformer.

QUESTIONS

17 Figure 28.29 represents a coil of wire ABCD being rotated in a uniform horizontal magnetic field. Copy and complete the diagram to show the direction of the induced current in the coil, and the directions of the forces on sides AB and CD that oppose the rotation of the coil.

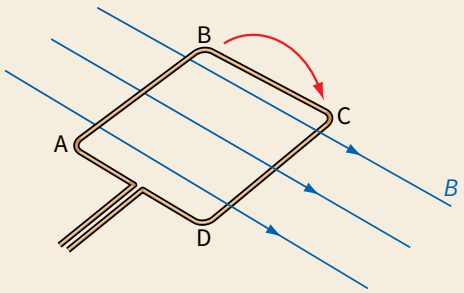


Figure 28.29 A coil rotating in a magnetic field.

18 Does a bicycle generator (Figure 28.7) generate alternating or direct current? Justify your answer.

19 The peak e.m.f. induced in a rotating coil in a magnetic field depends on four factors: magnetic flux density B, area of the coil A, number of turns N, and frequency f of rotation. Use Faraday's law to explain why the e.m.f. must be proportional to each of these quantities.

20 Explain why, if a transformer is connected to a steady (d.c.) supply, no e.m.f. is induced across the secondary coil.

Summary

- In a magnetic field of magnetic flux density B, the magnetic flux passing through an area A is given by $\Phi = BA$.
- The magnetic flux linking a coil of N turns is the magnetic flux linkage, $N\Phi$.
- Flux and flux linkage are measured in webers (Wb). $1\,\text{Wb} = 1\,\text{T m}^2$.
- When a conductor moves so that it cuts across a magnetic field, an e.m.f. is induced across its ends. When the magnetic flux linking a coil changes, an e.m.f. is induced in the coil.
- Faraday's law states that the magnitude of the induced e.m.f. is equal to the rate of change of magnetic flux linkage:

$$E = \frac{\Delta(N\Phi)}{\Delta t}$$

- Lenz's law states that the induced current or e.m.f. is in a direction so as to produce effects which oppose the change that is producing it.
- In an a.c. generator, an e.m.f. is induced because the rotating coil changes the magnetic flux linking the coil.

End-of-chapter questions

1 A student thinks that electrical current passes through the core in a transformer to the secondary coil. Describe how you might demonstrate that this is not true and explain how an electrical current is actually induced in the secondary coil. Use Faraday's law in your explanation. [3]

2 A square coil of 100 turns of wire has sides of length 5.0 cm. It is placed in a magnetic field of flux density 20 mT, so that the flux is perpendicular to the plane of the coil.

a Calculate the flux through the coil. [2]

b The coil is now pulled from the magnetic field in a time of 0.10 s. Calculate the average e.m.f. induced in it. [3]

3 An aircraft of wingspan 40 m flies horizontally at a speed of 300 m s^{-1} in an area where the vertical component of the Earth's magnetic field is 5.0×10^{-5} T. Calculate the e.m.f. generated between the aircraft's wingtips. [3]

4 What is an **eddy current**? State **one** example where eddy currents are useful and one where they are a disadvantage. [3]

5 Figure 28.27 shows the magnetic flux linkage and induced e.m.f. as a coil rotates. Explain why the induced e.m.f. is a maximum when there is no flux linkage and the induced e.m.f. is zero when the flux linkage is a maximum. [4]

6 **a** Explain what is meant by a **magnetic flux linkage** of 1 Wb. [2]

b Figure 28.31 shows how the magnetic flux density through a 240 turn coil with a cross-sectional area 1.2×10^{-4} m^2 varies with time.

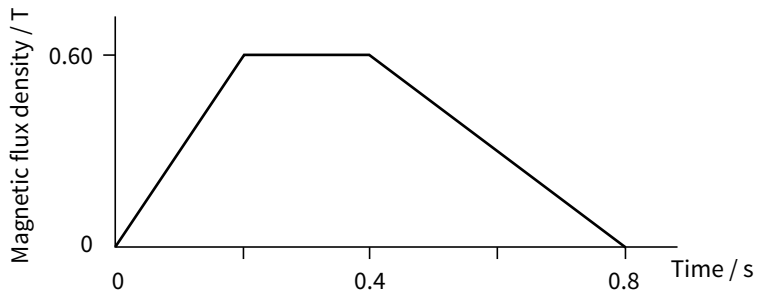


Figure 28.31 For End-of-chapter Question 6.

i Determine the maximum rate of change of flux in the coil. [2]

ii Determine the induced e.m.f. in the coil. [2]

iii Sketch a diagram to show the induced e.m.f. varies with time. Mark values on both the e.m.f. and time axes. [2]

7 Figure 28.32 shows a square coil about to enter a region of uniform magnetic field of magnetic flux density 0.30 T. The magnetic field is at right angles to the plane of the coil. The coil has 150 turns and each side is 2.0 cm in length. The coil moves at a constant speed of 0.50 m s^{-1}.

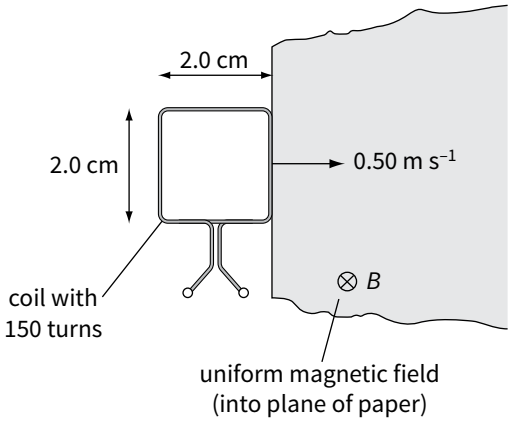


Figure 28.32 For End-of-chapter Question 7.

a **i** Calculate the time taken for the coil to completely enter the region of magnetic field. [1]

ii Determine the magnetic flux linkage through the coil when it is all within the region of magnetic field. [2]

b Explain why the induced e.m.f. is constant while the coil is entering the magnetic field. [1]

c Use your answer to **a** to determine the induced e.m.f. across the ends of the coil. [4]

d What is the induced e.m.f. across the ends of the coil when it is completely within the magnetic field? Explain your answer. [2]

e Sketch a graph to show the variation of the induced e.m.f. with time from the instant that the coil enters the magnetic field. Your time axis should go from 0 to 0.08 s. [2]

8 **a** State **Faraday's law** of electromagnetic induction. [2]

b A circular coil of diameter 200 mm has 600 turns (Figure 28.33). It is placed with its plane perpendicular to a horizontal magnetic field of uniform flux density 50 mT. The coil is then rotated through 90° about a vertical axis in a time of 120 ms.

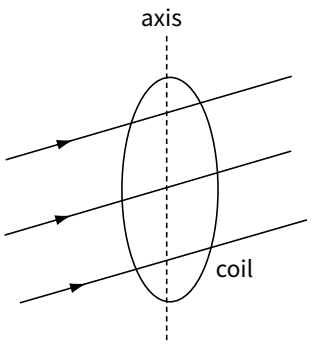


Figure 28.33 For End-of-chapter Question 8.

Calculate:

i the magnetic flux passing through the coil before the rotation [2]

ii the change of magnetic flux linkage produced by the rotation [2]

iii the average e.m.f. induced in the coil during the rotation. [2]

9 **a** State **Lenz's law** and explain how you would use a coil and a magnet to demonstrate the law. Make clear any other apparatus that you use. [4]

b A vehicle brake consists of an aluminium disc attached to a car axle (Figure 28.34). Electromagnets cause an e.m.f. to be induced in the disc.

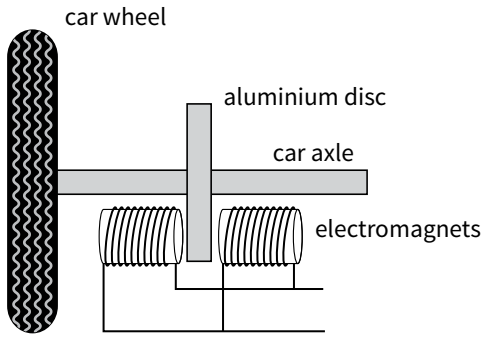


Figure 28.34 For End-of-chapter Question 9.

i Explain how the induction of an e.m.f. causes the vehicle to slow down. [3]

ii Explain why the braking effect increases when the speed of the car increases. [2]

10 A bicycle wheel is mounted vertically on a metal axle in a horizontal magnetic field (Figure 28.35). Sliding connections are made to the metal edge of the wheel and to the metal axle.

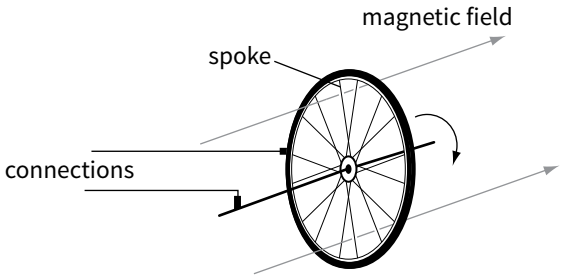


Figure 28.35 For End-of-chapter Question 10.

a **i** Explain why an e.m.f. is induced when the wheel rotates. [2]

ii State and explain **two** ways in which this e.m.f. can be increased. [2]

b The wheel rotates five times per second and has a radius of 15 cm. The magnetic flux density may be assumed to be uniform and of value 5.0×10^{-3} T.

Calculate:

i the area swept out each second by one spoke [2]

ii the induced e.m.f. between the contacts. [2]



Chapter 29: Alternating currents

Learning outcomes

You should be able to:

- measure frequency and voltage using a cathode-ray oscilloscope
- describe an alternating current or voltage in terms of period, frequency, peak value and r.m.s. value
- relate r.m.s. and peak values for sinusoidal currents
- solve problems involving transformers
- explain the benefits of transmission of electrical energy at high voltages
- explain how diodes and capacitors can be used to produce rectified, smoothed currents and voltages

Describing alternating current

In developed countries, mains electricity is a supply of alternating current (a.c.). The first mains electricity supplies were developed towards the end of the 19th century; at that time, a great number of different voltages and frequencies were used in different places. In some places, the supply was direct current (d.c.). Nowadays this has been standardised across much of the world, with standard voltages of 110 V or 230 V (or similar), and frequencies of 50 Hz or 60 Hz.

In this chapter we will look at some of the reasons why a.c. has been chosen as standard. First, however, we must take a close look at the nature of alternating currents.



Figure 29.1 Public electricity supplies made possible new forms of street lighting and advertising.

Sinusoidal current

An alternating current can be represented by a graph such as that shown in Figure 29.2. This shows that the current varies cyclically. During half of the cycle, the current is positive, and in the other half it is negative. This means that the current flows alternately one way and then the other in the wires in which it is travelling. Whenever you use a mains appliance, current flows backwards and forwards in the wires between you and the power station where it is being generated. At any instant in time, the current has a particular magnitude and direction given by the graph.

The graph has the same shape as the graphs used to represent simple harmonic motion (see Chapter 19), and it can be interpreted in the same way. The electrons in a wire carrying a.c. thus move back and forth with s.h.m. The current varies like a sine wave and so it is described as **sinusoidal**. (In principle, any current whose direction changes between positive and negative can be described as **alternating**, but we will only be concerned with those which have a regular, sinusoidal pattern.)

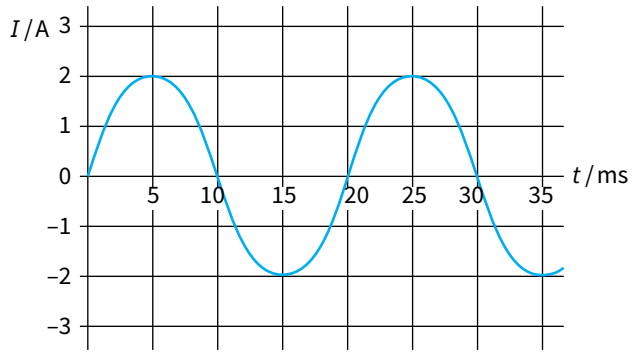


Figure 29.2 A graph to represent a sinusoidal alternating current.

QUESTION

1 The following questions relate to the graph of Figure 29.2.

- **a** What is the value of the current I when time $t = 5$ ms? In which direction is it flowing?
- **b** At what time does the current next have the same value, but negative?
- **c** What is the time T for one complete cycle?
- **d** What is the frequency of the alternating current?

An equation for a.c.

As well as drawing a graph, we can write an equation to represent alternating current. This equation tells us the value of the current I at any time t:

$$I = I_0 \sin \omega t$$

where ω is the angular frequency of the supply measured in rad s^{-1} (radians per second). This is related to the frequency f in the same way as for s.h.m.:

$$\omega = 2\pi f$$

and the frequency and period are related by:

$$f = \frac{1}{T}$$

The quantity I_0 is known as the peak value of the alternating current, found from the highest point on the graph. (It is the amplitude of the varying current.)

QUESTIONS

2 The following questions relate to the graph of Figure 29.2.

a What are the values of I_0 and ω?

b Write an equation to represent this current.

3 An alternating current (measured in amps, A) is represented by the equation:

$I = 5.0 \sin(120\pi t)$

a For this current what are the values of I_0, ω and f? What is the period T of the oscillation?

b Sketch a graph to represent the current.

Alternating voltages

Alternating current is produced in power stations by large generators like those shown in Figure 29.3.



Figure 29.3 Generators in the generating hall of a large power station.

In principle, a generator consists of a coil rotating in a magnetic field. An e.m.f. is induced in the coil according to the laws of electromagnetic induction. This e.m.f. V varies sinusoidally, and so we can write an equation to represent it which has the same form as the equation for alternating current:

$$V = V_0 \sin \omega t$$

where V_0 is the peak value of the voltage. We can also represent this graphically, as shown in Figure 29.4.

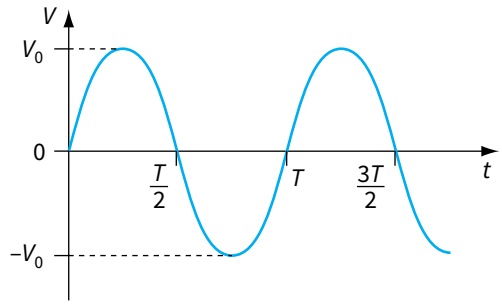


Figure 29.4 An alternating voltage.

QUESTION

4 An alternating voltage V (in V) is represented by the equation:

$V = 300 \sin(100\pi t)$

a What are the values of V_0, ω and f for this voltage?

b What is the value of V when $t = 0.002$ s? (Recall that $100\pi t$ is in radians when you calculate this.)

c Sketch a graph to show **two** complete cycles of this voltage.

Measuring frequency and voltage

An oscilloscope can be used to measure the frequency and voltage of an alternating current. Box 29.1 explains how to do this. There are two types of oscilloscope. The traditional cathode-ray oscilloscope (c.r.o.) uses an electron beam. The alternative is a digital oscilloscope, which is likely to be much more compact and which can store data and display the traces later.

BOX 29.1: Measurements using an oscilloscope

A c.r.o. is an electron beam tube, as shown in Figure 27.4 (page 424), but with an extra set of parallel plates to produce a horizontal electric field at right angles to the beam (Figure 29.5).

The principles of a c.r.o.

The signal into the c.r.o. is a repetitively varying voltage. This is applied to the y-input, which deflects the beam up and down using the parallel plates Y_1 and Y_2 shown in Figure 29.5. The time-base produces a p.d. across the other set of parallel plates X_1 and X_2 to move the beam from left to right across the screen.

BOX 29.1: Measurements using an oscilloscope (continued)

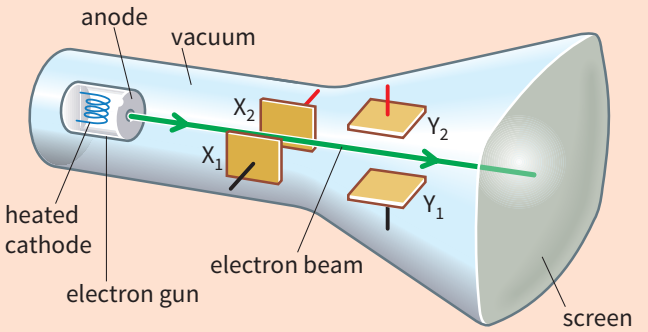


Figure 29.5 The construction of a cathode-ray oscilloscope. Cathode rays (beams of electrons) are produced in the electron gun and then deflected by electric fields before they strike the screen.

When the beam hits the screen of the c.r.o. it produces a small spot of light. If you look at the screen and slow the movement down you can see the spot move from left to right, while the applied signal moves the spot up and down. When the spot reaches the right side of the screen it flies back very quickly and waits for the next cycle of the signal to start before moving to the right once again. In this way the signal is displayed as a stationary trace on the screen. There may be many controls on a c.r.o., even more than those shown on the c.r.o. illustrated in Figure 29.6.

The controls

The X-shift and the Y-shift controls move the whole trace in the *x*-direction and the *y*-direction, respectively. The two controls that you must know about are the time-base and the Y-gain, or Y-sensitivity.

You can see in Figure 29.6 that the time-base control has units marked alongside. Let us suppose that this reads 5 ms/cm, although it might be 5 ms/division. This shows that 1 cm (or 1 division) on the *x*-axis represents 5 ms. Varying the time-base control alters the speed with which the spot moves across the screen. If the time-base is changed to 1 ms/cm, then the spot moves faster and each centimetre represents a smaller time.

The Y-gain control has a unit marked in volts/cm, or sometimes volts/division. If the actual marking is 5 V/cm, then each centimetre on the *y*-axis represents 5 V in the applied signal.

It is important to remember that the *x*-axis represents time and the *y*-axis represents voltage.

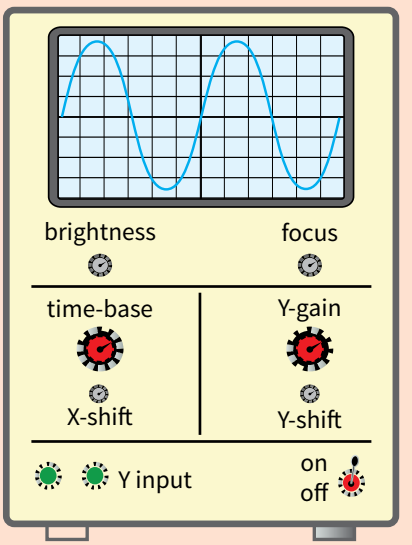


Figure 29.6 The controls of a typical c.r.o.

Determining amplitude and frequency

If you look at the c.r.o trace shown in Figure 29.7, you can see that the amplitude of the waveform is 2 cm and the distance along the *x*-axis for one complete wave is 4 cm.

If the Y-gain or Y-sensitivity setting is 2 V/cm, then the highest voltage is $2 \times 2 = 4$ V. If the time-base setting is 5 ms/cm, then the time for one wave (the period) is $4 \times 5 = 20$ ms.

Since the *x*-axis measures time, the c.r.o. trace can be used to measure frequency. In the above example, since

$$\text{period} = \frac{1}{\text{frequency}}$$

$$\text{frequency} = \frac{1}{0.02} = 50\ \text{Hz}$$

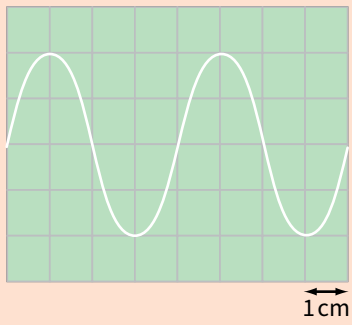


Figure 29.7 A c.r.o. trace when a sinusoidal alternating current is applied to the Y-plates.

QUESTIONS

5 If the Y-sensitivity and time-base for the trace shown in Figure 29.7 are 5 V/cm and 10 ms/cm, what are the amplitude, period and frequency of the signal to the Y-input?

6 Draw the c.r.o. trace for a sinusoidal voltage of frequency 100 Hz and amplitude 10 V, when the time-base is 10 ms/cm and the Y-sensitivity is 10 V/cm.

Power and a.c.

We use mains electricity to supply us with energy. If the current and voltage are varying all the time, does this mean that the power is varying all the time too? The answer to this is yes. You may have noticed that some fluorescent lamps flicker continuously, especially if you observe them out of the corner of your eye. A tungsten filament lamp would flicker too, but the frequency of the mains has been chosen so that the filament does not have time to cool down noticeably between peaks in the supply.

Root-mean-square values

There is a mathematical relationship between the peak value V_0 of the alternating voltage and a d.c. voltage which delivers the same average electrical power. The d.c. voltage is about 70% of V_0. (You might have expected it to be about half, but it is more than this, because of the shape of the sine graph.) This steady d.c. voltage is known as the root-mean-square (r.m.s.) value of the alternating voltage. In the same way, we can think of the root-mean-square value of an alternating current, I_{rms}:

The root-mean-square value of an alternating current is that steady current which delivers the same average power as the a.c. to a resistive load.

(The lamps in Box 29.2 are the 'resistive loads'.) A full analysis, which we will come to shortly, shows that I_{rms} is related to I_0 by:

$$I_{rms} = I_0/\sqrt{2} \quad \text{or} \quad I_{rms} \approx 0.707 \times I_0$$

This is where the factor of 70% comes from. Note that this factor only applies to sinusoidal alternating currents.

BOX 29.2: Comparing a.c. and d.c

Because power supplied by an alternating current is varying all the time, we need to have some way of describing the **average power** which is being supplied. To do this, we compare an alternating current with a direct current, and try to find the direct current that supplies the same average power as the alternating current.

Figure 29.8 shows how this can be done in practice. Two lamps are placed side by side; one is connected to an a.c. supply (on the right) and the other to a d.c. supply (the batteries on the left). The a.c. supply is adjusted so that the two lamps are equally bright, indicating that the two supplies are providing energy at the same average rate. The output voltages are then compared on the double-beam oscilloscope.

A typical trace is shown in Figure 29.9. This shows that the a.c. trace sometimes rises above the steady d.c. trace, and sometimes falls below it. This makes sense: sometimes the a.c. is delivering more power than the d.c., and sometimes less, but the average power is the same for both.

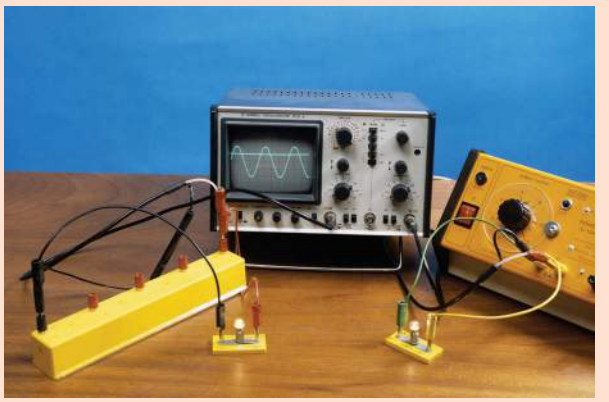


Figure 29.8 Comparing direct and alternating currents that supply the same power. The lamps are equally bright.

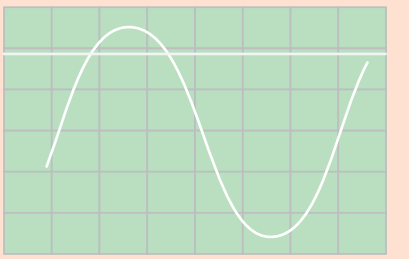


Figure 29.9 The oscilloscope trace from the experiment shown in Figure 29.8.

QUESTIONS

7 What is the r.m.s. value of an alternating current represented (in amps) by the equation $I = 2.5 \sin(100\pi t)$?

8 The mains supply to domestic consumers in many European countries has an r.m.s. voltage V_{rms} of 230 V. (Note that it is the r.m.s. value which is generally quoted, not the peak value.) What is the peak value of the supply?

Calculating power

The importance of r.m.s. values is that they allow us to apply equations from our study of direct current to situations where the current is alternating. So, to calculate the average power dissipated in a resistor, we can use the usual formulae for power:

$$P = I^2R = IV = \frac{V^2}{R}$$

Remember that it is essential to use the r.m.s. values of I and V, as in Worked example 1. If you use peak values, your answer will be too great by a factor of 2.

Where does this factor of 2 come from? Recall that r.m.s. and peak values are related by $I_0 = 2I_{rms}$. So, if you calculate I^2R using I_0 instead of I_{rms}, you will introduce a factor of $\sqrt{2}^2$, i.e. a factor of 2. (The same is true if you calculate power using V_0 instead of V_{rms}.) It follows that, for a sinusoidal alternating current, peak power is twice average power.

WORKED EXAMPLE

1 A sinusoidal p.d. of peak value 25 V is connected across a 20 Ω resistor. What is the average power dissipated in the resistor?

Step 1 Calculate the r.m.s. value of the p.d.:

$$V_{rms} = \frac{V_0}{\sqrt{2}} = \frac{25}{\sqrt{2}} = 17.7\,\text{V}$$

Step 2 Now calculate the average power dissipated:

$$P = \frac{V^2}{R} = \frac{17.7^2}{20} = 15.6\,\text{W}$$

(Note that, if we had used V_0 rather than V_{rms}, we would have found $P = \frac{25^2}{20} = 31.3\,\text{W}$, which is double the correct answer.)

QUESTIONS

9 What is the average power dissipated when a sinusoidal alternating current with a peak value of 3.0 A flows through a 100 Ω resistor?

10 A sinusoidal voltage of peak value 325 V is connected across a 1 kΩ resistor.

a What is the r.m.s. value of this voltage?

b Use $V = IR$ to calculate the r.m.s. current which flows through the resistor.

c What is the average power dissipated in the resistor?

d What is the peak power dissipated in the resistor?

Explaining root-mean-square

We will now briefly consider the origin of the term root-mean-square and show how the factor of $\sqrt{2}$ comes about. The equation $P = I^2R$ tells us that the power P is proportional to the square of the current I. Figure 29.10 shows how we can calculate I^2 for an alternating current. The current I varies sinusoidally, and during half of each cycle it is negative. However, I^2 is always positive (because the square of a negative number is positive). Notice that I^2 varies up and down, and that it has twice the frequency of the current.

Now, if we consider $<I^2>$, the average (mean) value of I^2, we find that its value is half the peak value (because the graph is symmetrical):

$$<I^2> = \tfrac{1}{2}I^2$$

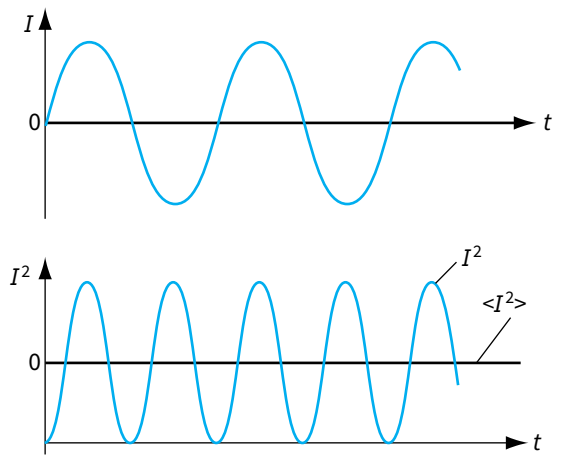


Figure 29.10 An alternating current I is alternately positive and negative, while I^2 is always positive.

To find the r.m.s. value of I, we now take the square root of $<I^2>$. This introduces a factor of the square root of $\frac{1}{2}$, or $\frac{1}{\sqrt{2}}$.

Summarising this process: to find the r.m.s. value of the current, we find the root of mean of the square of the current – hence r.m.s.

Why use a.c. for electricity supply?

There are several reasons for preferring alternating voltages for a national electricity supply system. The most important reason is that a.c. can be transformed to high voltages, so that the current flowing is reduced, and this leads to lower power losses in the transmission lines. Typically, the generators at a power station produce electrical power at a voltage of 25 kV. This is transformed up to a voltage of perhaps 400 kV (and as much as 1 MV in some countries). The power is then transmitted along many kilometres of high-voltage power lines (Figure 29.11) before being transformed down to a lower voltage for supply to the millions of consumers. The transformers used for increasing and decreasing the voltage are discussed in detail in the next section.

This high voltage brings problems: the lines must be suspended high above the ground between pylons, and high-quality insulators are needed to prevent current passing from the cables to the pylons. As current flows through transmission lines (wires), it loses power because of the resistance of the lines. The wires become warm; this is resistive or ohmic heating, as discussed in Chapter 9. The smaller the current, the smaller the losses. This is illustrated in Worked example 2.



Figure 29.11 Power lines carry electricity from power station to consumer.

WORKED EXAMPLE

2 A power station generates electrical power at a rate of 10 MW. This power is to be transmitted along cables whose total resistance is 10 Ω. Calculate the power losses in the cable if the power is transmitted at 50 kV and at 250 kV.

Step 1 Using $I = \frac{P}{V}$, calculate the current flowing in each case:

for 50 kV: $I = \frac{10 \times 10^6}{5 \times 10^4} = 200\,\text{A}$

for 250 kV: $I = \frac{10 \times 10^6}{25 \times 10^4} = 40\,\text{A}$

Step 2 Using $P = I^2R$, calculate the power losses in each case:

for 50 kV: $P = 200^2 \times 10 = 4 \times 10^5 = 400\,\text{kW}$

for 250 kV: $P = 40^2 \times 10 = 1.6 \times 10^4 = 16\,\text{kW}$

Hint: *Take care! Note that we have two quantities for which we are using the symbol P: the total power being transmitted, and the power lost in the wires. Notice that using a higher voltage does not change the resistance of the cables.*

We have shown that, by increasing the voltage by a factor of 5, we have reduced the power losses by a factor of 25.

Economic savings

The resistive heating of power lines is a waste of money, in two ways. Firstly, it costs money to generate power because of the fuel needed. Secondly, more power stations are required, and power stations are expensive. The use of transformers to transform power to high voltages saves a few per cent of a national bill for electrical power, and means that fewer expensive power stations are needed.

It is claimed that having a few, very large power stations gives economies of scale, but this is debated by many environmentalists who would prefer to see many small, local power stations. It is also the case that new developments in technology are making it easier to transform direct current to high voltages. This is more compatible with sustainable electricity generating systems such as photovoltaics (solar cells), so we may see the development of d.c. grid systems in the near future.

Transformers

Figure 29.12 shows the construction of a simple transformer. The primary coil of N_p turns of wire is wound around an iron core. The secondary coil of N_s turns is wound on the opposite side of the core. (Many different configurations are possible, with different shapes of core and with the coils wound separately, or one on top of the other.)

The p.d. V_p across the primary coil causes an alternating current I_p to flow. This produces an alternating magnetic field in the soft iron core. The secondary coil is thus in a changing magnetic field, and an alternating current I_s is induced in it. There is thus an alternating e.m.f. V_s across the secondary coil.

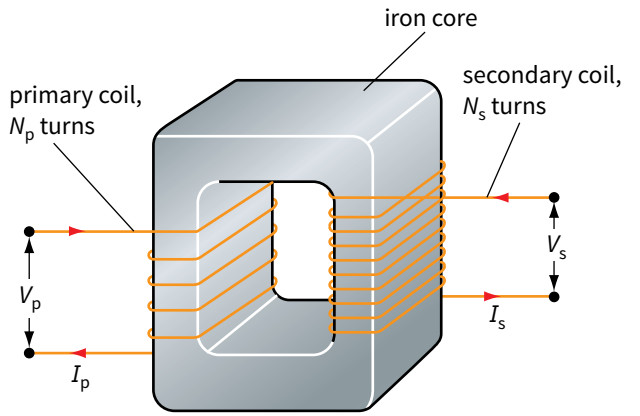


Figure 29.12 Defining quantities for a simple iron-cored transformer.

Note that there is no **electrical** connection between the primary coil and the secondary coil. Energy is transferred from one to the other via the magnetic field in the core.

Step-up, step-down

The transformer represented in Figure 29.12 has 5 turns on its primary coil and 10 on its secondary coil. It is described as a **step-up transformer** because the output voltage is greater than the input voltage (the voltage has been 'stepped up').

How does this happen? We have 5 turns producing magnetic flux. This flux links the 10 turns of the secondary coil. Because flux linkage $N\Phi$ is proportional to the number of turns, it follows that there is more magnetic flux – twice as much – linking the secondary coil than the primary. As the magnetic flux changes (because we are using alternating voltages), the e.m.f. induced in the secondary coil is greater than the voltage across the primary coil.

We can write an equation relating the voltages across the coils to the number of turns in each coil:

$$\frac{V_s}{V_p} = \frac{N_s}{N_p}$$

The equation above is known as the **turns-ratio equation** for a transformer.

In words, the ratio of the voltages is equal to the ratio of the number of turns of the transformer. For the transformer in Figure 29.12, a voltage of 1.0 V applied to the primary coil will result in an output of 2.0 V across the secondary coil; 50 V will give 100 V, and so on.

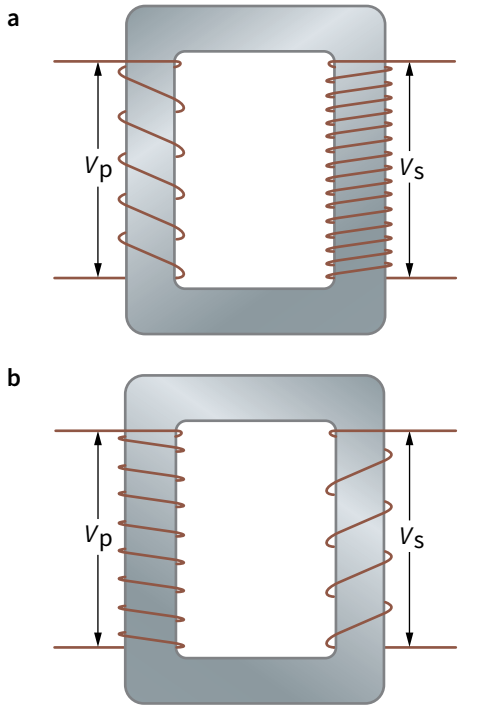


Figure 29.13 Two transformers: **a** step-up, and **b** step-down.

A transformer with fewer turns on the secondary coil than on the primary coil is described as a **step-down transformer**. It lowers the voltage at the primary coil. Figure 29.13 shows both types of transformer. Worked example 3 shows how to use the turns-ratio equation.

QUESTIONS

11 **a** What is the turns ratio of the transformer shown in Figure 29.13a?

b What is the turns ratio of the transformer shown in Figure 29.13b?

c If an alternating p.d. of value 10.0 V is connected across the primary coil of each, what will be the induced e.m.f. across each secondary?

12 A power station generates electricity at a voltage of 25 kV. This must be transformed for onward transmission at 400 kV. If the primary coil of the transformer used has 2000 turns, how many turns must the secondary coil have?

WORKED EXAMPLE

3 A radio requires a 6.0 V supply but has to be operated from a 230 V mains supply. It is fitted with a transformer to reduce the mains voltage. Calculate the turns ratio for this transformer. If the primary coil has 5000 turns, how many turns must the secondary have?

Step 1 The turns ratio $\frac{N_s}{N_p}$ is given by:

$$\frac{V_s}{V_p} = \frac{N_s}{N_p}$$

Hence:

$$\frac{N_s}{N_p} = \frac{V_s}{V_p} = \frac{6.0}{230} = 0.026$$

This ratio is less than one because we are reducing the voltage.

Step 2 We know that $N_p = 5000$, so we can calculate N_s.

$N_s = N_p \times 0.026 = 5000 \times 0.026 = 130$

So the secondary coil must have 130 turns. Check that this seems reasonable: the voltage has to be reduced by a factor of about 40, so the number of turns must be reduced by the same factor.

Voltage, current, power

If there is no power lost in a transformer, it follows that the quantity $I \times V$ is the same for both primary and secondary coils:

$$I_p V_p = I_s V_s \quad \text{or} \quad \frac{V_s}{V_p} = \frac{I_p}{I_s}$$

In other words, the ratio of the voltages is the **inverse** ratio of the currents. If the voltage is stepped up, the current is stepped down, and vice versa. This explains the function of transformers in power transmission, as discussed earlier.

Note that this relationship assumes that no power is lost in the transformer. In practice, some power is lost because of the resistance of the transformer coil windings, so that the windings become warm. In addition, the changing magnetic flux can induce currents in the core. These eddy currents heat the core because of its resistance. Large transformers such as the one shown in Figure 29.14 handle a large amount of power. A small percentage is wasted, and the resulting heat is carried away by cooling fluid pumped around the transformer and through the fins which are visible in the photograph.

The transformers used in the electricity supply industry must be designed with great care to minimise energy losses. The electricity supply may pass through as many as ten transformers between the generator and the consumer. If each transformer wasted just 1% of the power, that would give an overall loss of 10%. Since there are roughly 100 big power stations in the UK, that would require ten power stations just to cope with the losses in transformers. Today's well-designed transformers have losses of under 0.1%. This contributes greatly to energy savings in the power transmission industry.



Figure 29.14 Installing a new transformer in a gold mine in Ghana.

QUESTIONS

13 A step-up transformer has a turns ratio of 10 : 1. When an alternating p.d. of 20 V is connected across the primary coil, a current of 50 mA flows in it.

a Calculate the values of the p.d. across the secondary coil and the current flowing in it, assuming no power is lost in the transformer.

b In practice the secondary p.d. is found to be 180 V and the secondary current is 4.5 mA. What percentage of the power is wasted in the transformer?

14 Table 29.1 shows information about three transformers. Copy and complete the table. (Assume no power is lost in the transformers.)

Transformer	N_p	N_s	V_p/V	V_s/V	I_p/A	I_s/A	P/W
A	100	500	230		1.0		
B	500	100	230		1.0		
C	100		12	240	0.2		

Table 29.1 Details of three transformers for Question 14.

Rectification

Many electrical appliances work with alternating current. Some, like electrical heaters, will work equally well with d.c. or a.c. However, there are many appliances such as electronic equipment which require d.c. For these, the alternating mains must be converted to d.c. by the process of **rectification**.

A simple way to do this is to use a diode, which is a component that will only allow current to flow in one direction. Figure 29.15 shows a circuit for doing this, together with a graph to show the effect. You will see that the output voltage is always positive, but it goes up and down. This is still technically direct current, because the current only flows in one direction.

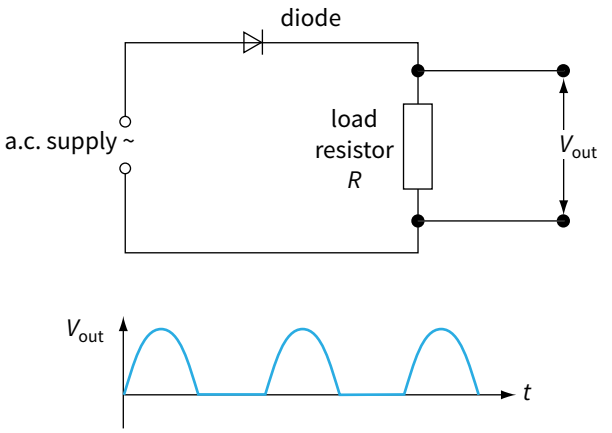


Figure 29.15 Half-wave rectification of a.c. requires a single diode.

This type of rectification is known as half-wave rectification. For one-half of the time the voltage is zero, and this means that the power available from a half-wave rectified supply is reduced.

The bridge rectifier

To overcome this problem of reduced power, a bridge rectifier circuit is used. This consists of four diodes connected across the alternating voltage, as shown in Figure 29.16. The resulting output voltage across the load resistor *R* is full-wave rectified.

The way in which this works is shown in Figure 29.17.

- During the first half of the a.c. cycle, terminal A is positive. Current flows through diode 2, downwards through *R* and through diode 3 to terminal B. In this half of the cycle, current cannot flow through diodes 1 or 4 because they are pointing the wrong way.
- In the second half of the cycle, terminal B is positive. Current flows through diode 4, downwards through *R*, and through diode 1 to terminal A. Diodes 2 and 3 do not conduct because they are pointing the wrong way.

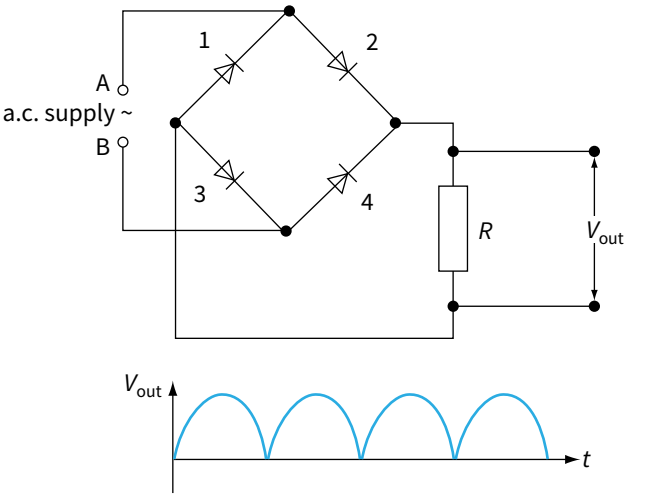


Figure 29.16 Full-wave rectification of a.c. using a diode bridge.

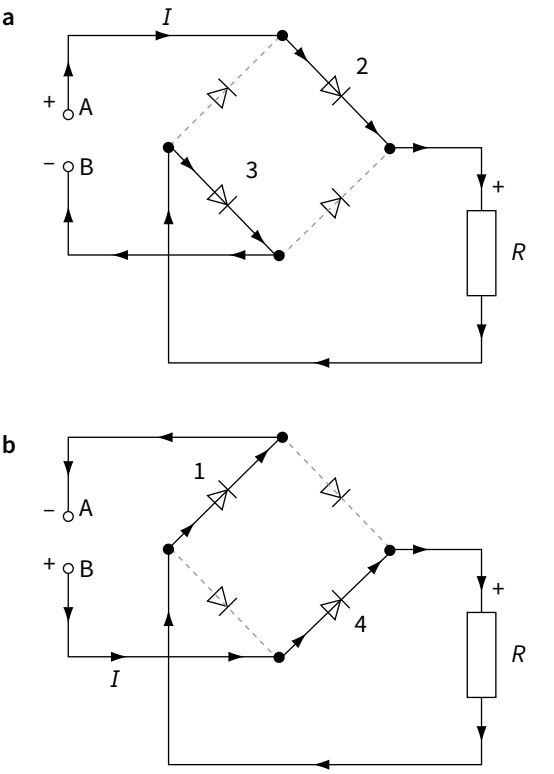


Figure 29.17 Current flow during full-wave rectification.

Note that in both halves of the cycle, current flows the same way (downwards) through *R*, so the top end of *R* must be positive.

You can construct a bridge rectifier using light-emitting diodes (LEDs) which light up when current flows through them. By connecting this bridge to a slow a.c. supply (for instance 1 Hz from a signal generator), you can see the sequence in which the diodes conduct during rectification.

QUESTION

15 Explain why, when terminal B in Figure 29.17 is positive (during the second half of the cycle), the current flows through diodes 1 and 4, but not through diodes 2 and 3.

Smoothing

In order to produce steady d.c. from the 'bumpy' d.c. that results from rectification, a smoothing capacitor must be incorporated in the circuit, in parallel with the load resistor *R*. This is shown in Figure 29.18. The idea is that the capacitor charges up and maintains the voltage at a high level. It discharges gradually when the rectified voltage drops, but the voltage soon rises again and the capacitor charges up again. The result is an output voltage with 'ripple'.

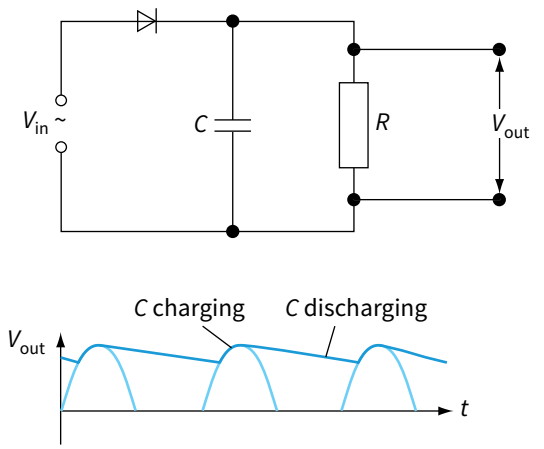


Figure 29.18 A smoothing capacitor is connected across (in parallel with) the load resistor.

The amount of ripple can be controlled by careful choice of the values of *C* and *R*. A large capacitor discharges more slowly than a small capacitor, so will give less ripple. Similarly, if *R* has a large value, *C* will discharge more slowly. In practice, the greater the value of the quantity $R \times C$, the smoother the rectified a.c. However, if *R* and *C* have large values, it will be difficult to change the value of the voltage quickly.

Note that in Figures 29.15 to 29.18 we have represented the load on the supply by the load resistor *R*. This represents any components that are connected to the supply. For example, *R* could be replaced by a mobile phone battery that is being charged, or by one of the electronic circuits discussed in Chapter 25.

QUESTIONS

16 Sketch the following voltage patterns:

- **a** a sinusoidal alternating voltage
- **b** the same voltage as **a**, but half-wave rectified
- **c** the same voltage as **b**, but smoothed
- **d** the same voltage as **a**, but full-wave rectified
- **e** the same voltage as **d**, but smoothed.

17 A student wires a bridge rectifier incorrectly as shown in Figure 29.19. Explain what you would expect to observe when an oscilloscope is connected across the load resistor *R*.

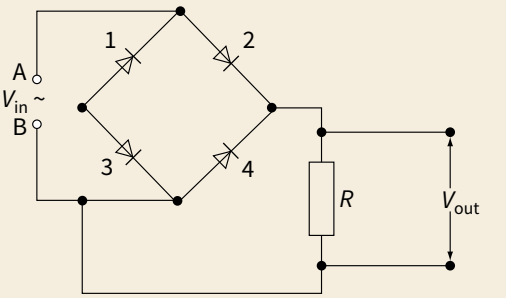


Figure 29.19 A bridge rectifier circuit that is wired incorrectly – see Question 17.

18 A bridge rectifier circuit is used to rectify an alternating current through a resistor *R*. A smoothing capacitor *C* is connected across *R*. Figure 29.20 shows how the current varies. Use sketches to show the changes you would expect:

- **a** if *R* is increased
- **b** if *C* is decreased.

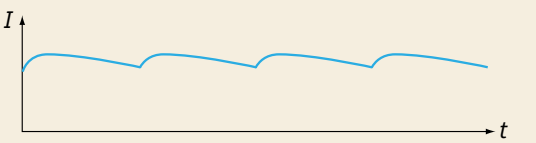


Figure 29.20 A smoothed, rectified current – see Question 18.

Summary

- A sinusoidal alternating current can be represented by $I = I_0 \sin \omega t$, where I_0 is the peak value of the current.
- The root-mean-square value of an alternating current is that steady current which delivers the same average power as the a.c. to a resistive load; for a sinusoidal a.c., $I_{rms} = \frac{I_0}{\sqrt{2}}$.
- Electrical power is usually transmitted at high voltages; this allows the current to be reduced, and so resistive losses are lower.
- Transformers are used to change an alternating voltage. The voltage is stepped up or down in proportion to the turns ratio of the transformer.
- For a transformer, $\frac{V_s}{V_p} = \frac{N_s}{N_p}$. If it is 100% efficient, then $V_p I_p = V_s I_s$.
- Diodes are used to convert a.c. to d.c. A single diode gives half-wave rectification. A bridge of four diodes gives full-wave rectification. A capacitor smoothes the rectified voltage.

End-of-chapter questions

1 Write down a general expression for the sinusoidal variation with time t of:

a an alternating voltage V [1]

b an alternating current I (you may assume that I and V are in phase) [1]

c the power P dissipated due to this current and voltage. [1]

2 The value in amps of an alternating current is represented by the equation $I = 2 \sin (50\pi t)$.

a What is the peak value of the current? [1]

b What is the frequency of the supply? [2]

c Sketch a graph to show **two** cycles of the variation of current with time. Mark the axes with suitable values. [2]

d Calculate I_{rms}, the r.m.s. value of current, and mark this on your graph. [1]

e Find **two** values of t at which $I = I_{rms}$. [3]

3 The a.c. mains of 240 V r.m.s. is connected to the primary coil of a transformer, which contains 1200 turns. The r.m.s. output of the transformer is 6.0 V.

a Calculate the number of turns on the secondary coil. [1]

b A resistance of 6.0 Ω is connected across the secondary coil. Calculate:

i the average power dissipated in the resistor [1]

ii the peak current in the primary coil. [3]

4 The graph in Figure 29.21 shows the sinusoidal variation of current in the primary coil of a transformer. The current in the secondary coil is zero.

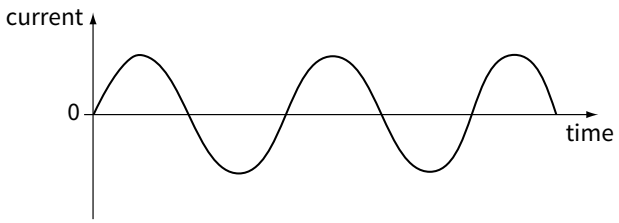


Figure 29.21 For End-of-chapter Question 4.

a Copy the graph. On your copy draw, on the same axes, the variation with time of:

i the magnetic flux in the core of the transformer – label this A [1]

ii the induced e.m.f. in the primary coil of the transformer – label this B. [1]

b State how the two graphs in **a i** and **ii** are related to each other. [1]

5 An oscilloscope is used to measure the voltage waveform across a 200 Ω resistor. The waveform is shown in Figure 29.22. The time-base of the oscilloscope is set at 5 ms div^{-1} and the Y-gain at 0.5 V div^{-1}.

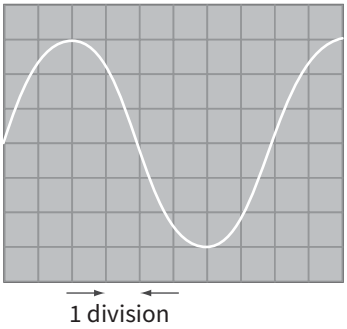


Figure 29.22 For End-of-chapter Question 5.

Determine:

a the period and hence the frequency of the waveform [2]

b the peak voltage and hence the r.m.s. voltage [2]

c the r.m.s. current in the resistor [1]

d the mean power dissipated in the resistor. [2]

6 **a** State the relationship between the peak current I_0 and the r.m.s. current I_{rms} for a sinusoidally varying current. [1]

b The current in a resistor connected to a steady d.c. supply is 2.0 A. When the same resistor is connected to an a.c. supply, the current in it has a peak value of 2.0 A. The heating effects of the two currents in the resistor are different.

i Explain why the heating effects are different and state which heating effect is the greater. [2]

ii Calculate the ratio of the power dissipated in the resistor by the d.c. current to the power dissipated in the resistor by the a.c. current. [2]

c State **one** advantage of using alternating current in the home. [1]

7 A sinusoidal voltage of 6.0 V r.m.s. and frequency 50 Hz is connected to a diode and a resistor R of resistance 400 Ω as shown in Figure 29.23.

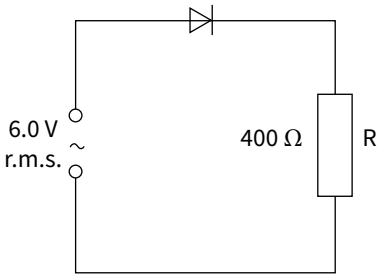


Figure 29.23 For End-of-chapter Question 7.

a Draw a sketch graph showing the variation with time of both the supply waveform (use a dotted line) and the voltage across R (use a solid line). Put numerical scales on both the voltage and time axes. [4]

b An uncharged capacitor C is connected across R. When the 6.0 V r.m.s. supply is switched on, the capacitor charges fully during the first quarter of a cycle. You may assume that the p.d. across the diode is zero when it conducts. For the next three-quarters of the first cycle, the diode stops conducting and the p.d. across R falls to one-half of the peak value. During this time the mean p.d. across R is 5.7 V.

For the last three-quarters of the first cycle, calculate:

i the time taken [1]

ii the mean current in R [2]

iii the charge flowing through R [2]

iv the capacitance of C. [2]

c Explain why the diode stops conducting during part of each cycle in **b**. [2]

8 The rectified output from a transformer is connected to a resistor R of resistance 1000 Ω. Graph A in Figure 29.24 shows the variation with time t of the p.d. V across the resistor. Graph B shows the variation of V when a capacitor is placed across R to smooth the output.

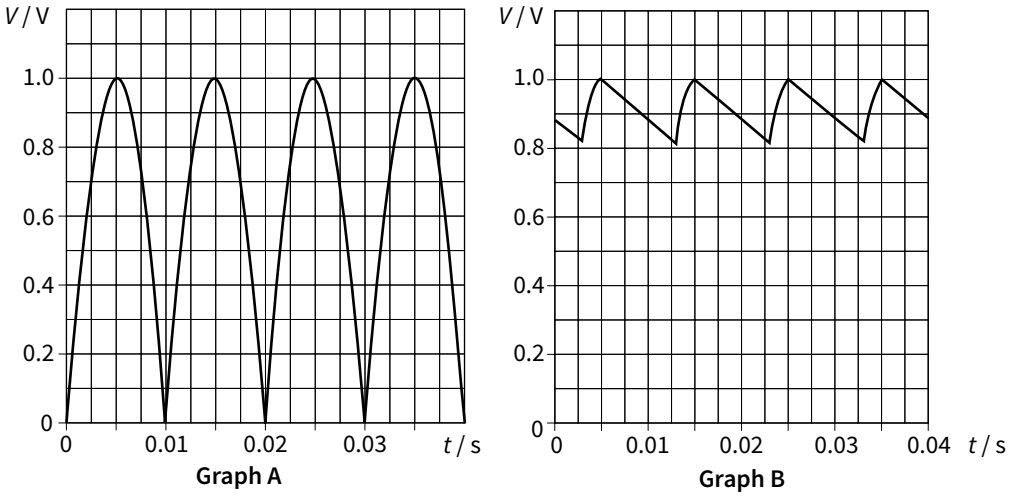


Figure 29.24 For End-of-chapter Question 8.

a Explain how the rectification is achieved. Draw a circuit diagram to show the components involved. [6]

b Explain the action of the capacitor in smoothing the output. [3]

c Using graph B between $t = 0.005$ and $t = 0.015$ s, determine:

i the time during which the capacitor is charging [1]

ii the mean value of the p.d. across R [1]

iii the average power dissipated in R. [2]

9 A transformer is used to produce an output of peak value 12 V from the 230 V r.m.s. mains supply.

a The primary coil of the transformer contains 2000 turns. Calculate the number of turns in the secondary coil. [3]

b The output from the secondary coil is half-wave rectified, and connected to a resistor R and capacitor C in parallel.

Sketch graphs on the same axes to show the variation with time of:

i the output p.d. of the secondary coil [1]

ii the p.d. across R. [2]

c State and explain what happens to the p.d. across R when another capacitor of equal value is placed in parallel with C. [3]

10 Electrical energy is supplied by a high-voltage power line which has a total resistance of 4.0 Ω. At the input to the line, the root-mean-square (r.m.s.) voltage has a value of 400 kV and the input power is 500 MW.

a i Explain what is meant by **root-mean-square voltage**. [2]

ii Calculate the minimum voltage that the insulators which support the line must withstand without breakdown. [2]

b **i** Calculate the value of the r.m.s. current in the power line. [2]

ii Calculate the power loss on the line. [2]

iii Explain why it is an advantage to transmit the power at a high voltage. [2]

c Power at 400 kV is converted to power at 124 kV by a step-down transformer. Describe the basic principle of a step-down transformer. [2]

Chapter 30: Quantum physics

Learning outcomes

You should be able to:

- appreciate the particulate nature of electromagnetic radiation
- interpret the photoelectric effect in terms of photons
- describe electron diffraction and the evidence it provides for the wave nature of matter
- explain line spectra in terms of discrete electron energy levels in atoms
- use simple band theory to explain the different electrical conductivities of metals, insulators and semiconductors

What is light?

When the first laser was made in 1960, it seemed like a clever idea, but it was a long time before it found any useful application. Today, lasers are everywhere – in CD and DVD machines, computer disc drives, supermarket barcode scanners – there are probably more lasers than people. Figure 30.1 shows a patient undergoing laser eye surgery.

The invention of the laser was only possible when scientists had cracked the mystery of the nature of light. Does light behave as particles or as waves? As we shall see in this chapter, the answer is – a bit of both.



Figure 30.1 This patient is undergoing laser eye surgery, which improves the focusing of the eye by modifying the shape of the surface of the eyeball.

Modelling with particles and waves

In this chapter, we will study two very powerful scientific models – particles and waves – to see how they can help us to understand more about both light and matter. First we will take a closer look at each of these models in turn.

Particle models

In order to explain the properties of matter, we often think about the particles of which it is made and the ways in which they behave. We imagine particles as being objects that are hard, have mass and move about according to the laws of Newtonian mechanics (Figure 30.2). When two particles collide, we can predict how they will move after the collision, based on knowledge of their masses and velocities before the collision. If you have played snooker or pool, you will have a pretty good idea of how particles behave.

Particles are a macroscopic model. Our ideas of particles come from what we observe on a macroscopic scale – when we are walking down the street, or observing the motion of stars and planets, or working with trolleys and balls in the laboratory. But what else can we explain using a particle model?

The importance of particle models is that we can apply them to the microscopic world, and explain more phenomena.

We can picture gas molecules as small, hard particles, rushing around and bouncing haphazardly off one another and the walls of their container. This is the kinetic model of a gas which we studied in depth in Chapter 22. We can explain the macroscopic (larger scale) phenomena of pressure and temperature in terms of the masses and speeds of the microscopic particles. This is a very powerful model, which has been refined to explain many other aspects of the behaviour of gases.



Figure 30.2 Pool balls provide a good model for the behaviour of particles on a much smaller scale.

Table 30.1 shows how, in particular areas of science, we can use a particle model to interpret and make predictions about macroscopic phenomena.

Area	Model	Macroscopic phenomena
electricity	flow of electrons	current
gases	kinetic theory	pressure, temperature and volume of a gas
solids	crystalline materials	mechanical properties
radioactivity	nuclear model of the atom	radioactive decay, fission and fusion reactions
chemistry	atomic structure	chemical reactions

Table 30.1 Particle models in science.

Wave models

Waves are something that we see on the sea. There are tidal waves, and little ripples. Some waves have foamy tops, others are breaking on the beach.

Physicists have an idealised picture of a wave – it is shaped like a sine graph. You will not see any waves quite this shape on the sea. However, it is a useful picture, because it can be used to represent some simple phenomena. More complicated waves can be made up of several simple waves, and physicists can cope with the mathematics of sine waves. (This is the principle of superposition, which we looked at in detail in Chapter 14.)

Waves are a way in which energy is transferred from one place to another. In any wave, something is changing in a regular way, while energy is travelling along. In water waves, the surface of the water moves up and down periodically, and energy is transferred horizontally.

Table 30.2 shows some phenomena that we explain in terms of waves.

Phenomenon	Varying quantity
sound	pressure (or density)
light (and other electromagnetic waves)	electric and magnetic field strengths
waves on strings	displacement

Table 30.2 Wave models in science.

The characteristic properties of waves are that they all show reflection, refraction, diffraction and interference. Waves themselves do not have mass or charge. Since particle models can also explain reflection and refraction, it is diffraction and interference that we regard as the defining characteristics of waves. If we can show diffraction and interference, we know that we are dealing with waves (Figure 30.3).



Figure 30.3 A diffraction grating splits up light into its component colours and can produce dramatic effects in photographs.

Waves or particles?

Wave models and particle models are both very useful. They can explain a great many different observations. But which should we use in a particular situation? And what if both models seem to work when we are trying to explain something?

This is just the problem that physicists struggled with for over a century, in connection with light. Does light travel as a wave or as particles?

For a long time, Newton's view prevailed – light travels as particles. This was set out in 1704 in his famous book *Opticks*. He could use this model to explain both reflection and refraction. His model suggested that light travels faster in water than in air. In 1801 Thomas Young, an English physicist, demonstrated that light showed diffraction and interference effects. Physicists were still very reluctant to abandon Newton's particle model of light. The ultimate blow to Newton's model came from the work carried out by the French physicist Léon Foucault in 1853. His experiments on the speed of light showed that light travelled more slowly in water than in air. Newton's model was in direct contradiction with experimental results. Most scientists became convinced that light travelled through space as a wave.

Particulate nature of light

We expect light to behave as waves, but can light also behave as particles? The answer is yes, and you are probably already familiar with some of the evidence.

If you place a Geiger counter next to a source of gamma radiation you will hear an irregular series of clicks. The counter is detecting γ-rays (gamma-rays). But γ-rays are part of the electromagnetic spectrum. They belong to the same family of waves as visible light, radio waves, X-rays, etc.

So, here are waves giving individual or discrete clicks, which are indistinguishable from the clicks given by α-particles (alpha-particles) and β-particles (beta-particles). We can conclude that γ-rays behave like particles when they interact with a Geiger counter.

This effect is most obvious with γ-rays, because they are at the most energetic end of the electromagnetic spectrum. It is harder to show the same effect for visible light.

Photons

The **photoelectric effect**, and Einstein's explanation of it, convinced physicists that light could behave as a stream of particles. Before we go on to look at this in detail, we need to see how to calculate the energy of photons.

Newton used the word **corpuscle** for the particles which he thought made up light. Nowadays, we call them **photons** and we believe that all electromagnetic radiation consists of photons. A photon is a 'packet of energy' or a quantum of electromagnetic energy. Gamma-photons (γ-photons) are the most energetic. According to Albert Einstein, who based his ideas on the work of another German physicist, Max Planck, the energy E of a photon in joules (J) is related to the frequency f in hertz (Hz) of the electromagnetic radiation of which it is part, by the equation:

$$E = hf$$

The constant h has an experimental value equal to 6.63×10^{-34} J s.

This constant h is called the **Planck constant**. It has units of joule seconds (J s), but you may prefer to think of this as 'joules per hertz'. The energy of a photon is directly proportional to the frequency of the electromagnetic waves, that is:

$$E \propto f$$

Hence, high-frequency radiation means high-energy photons.

Notice that the equation $E = hf$ tells us the relationship between a particle property (the photon energy E) and a wave property (the frequency f). It is called the **Einstein relation** and applies to all electromagnetic waves.

The frequency f and wavelength λ of an electromagnetic wave are related to the wave speed c by the wave equation $c = f\lambda$, so we can also write this equation as:

$$E = \frac{hc}{\lambda}$$

It is worth noting that the energy of the photon is inversely proportional to the wavelength. Hence the short-wavelength X-ray photon is far more energetic than the long-wavelength photon of light.

QUESTIONS

To answer questions 1–7 you will need these values:

speed of light in a vacuum $c = 3.00 \times 10^8$ m s^{-1}

Planck constant $h = 6.63 \times 10^{-34}$ J s

1 Calculate the energy of a high-energy γ-photon, of frequency 10^{26} Hz.

2 Visible light has wavelengths in the range 400 nm (violet) to 700 nm (red). Calculate the energy of a photon of red light and a photon of violet light.

3 Determine the wavelength of the electromagnetic waves for each photon below and hence use Figure 30.4 to identify the region of the electromagnetic spectrum to which each belongs.

The photon energy is:

a 10^{-12} J **b** 10^{-15} J **c** 10^{-18} J

d 10^{-20} J **e** 10^{-25} J

4 A 1.0 mW laser produces red light of wavelength 6.48×10^{-7} m. Calculate how many photons the laser produces per second.

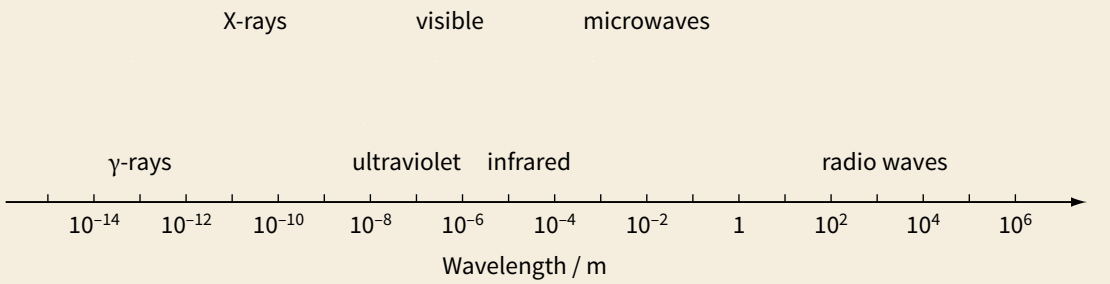


Figure 30.4 Wavelengths of the electromagnetic spectrum. The boundaries between some regions are fuzzy.

Now we can work out the energy of a γ-photon. Gamma-rays typically have frequencies greater than 10^{20} Hz. The energy of a γ-photon is therefore greater than $(6.63 \times 10^{-34} \times 10^{20}) \approx 10^{-13}$ J. This is a very small amount of energy on the human scale, so we don't notice the effects of individual γ-photons. However, some astronauts have reported seeing flashes of light as individual cosmic rays, high-energy γ-photons, passed through their eyeballs.

The electronvolt (eV)

The energy of a photon is extremely small and far less than a joule. Hence the joule is not a very convenient unit for measuring photon energies. You may remember from Chapter 16 that we use another energy unit, the **electronvolt**, when considering amounts of energy much smaller than a joule.

To recap from Chapter 16: when an electron travels through a potential difference, energy is transferred. If an electron, which has a charge of magnitude 1.60×10^{-19} C, travels through a potential difference of 1 V, its energy change W is given by:

$$W = QV = 1.60 \times 10^{-19} \times 1 = 1.60 \times 10^{-19} \text{ J}$$

We can use this the electronvolt:

> One electronvolt (1 eV) is the energy transferred when an electron travels through a potential difference of one volt.

Therefore:

$$1 \text{ eV} = 1.60 \times 10^{-19} \text{ J}$$

So when an electron moves through 1 V, 1 eV of energy is transferred. When one electron moves through 2 V, 2 eV of energy is transferred. When five electrons move through 10 V, a total of 50 eV is transferred, and so on.

- To convert from eV to J, multiply by 1.60×10^{-19}.
- To convert from J to eV, divide by 1.60×10^{-19}.

QUESTIONS

5 An electron travels through a cell of e.m.f. 1.2 V. How much energy is transferred to the electron? Give your answer in eV and in J.

6 Calculate the energy in eV of an X-ray photon of frequency 3.0×10^{18} Hz.

7 To which region of the electromagnetic spectrum (Figure 30.4) does a photon of energy 10 eV belong?

When a charged particle is accelerated through a potential difference V, its kinetic energy increases. For an electron (charge e), accelerated from rest, we can write:

$$eV = \tfrac{1}{2}mv^2$$

We need to be careful when using this equation. It does not apply when a charged particle is accelerated through a large voltage to speeds approaching the speed of light c. For this, we would have to take account of relativistic effects. (The mass of a particle increases as its speed gets closer to 3.00×10^8 m s^{-1}.)

Rearranging the equation gives the electron's speed:

$$v = \sqrt{\frac{2eV}{m}}$$

This equation applies to any type of charged particle, including protons (charge $+e$) and ions.

QUESTION

8 A proton (charge = $+1.60 \times 10^{-19}$ C, mass = 1.67×10^{-27} kg) is accelerated through a potential difference of 1500 V. Determine:

a its final kinetic energy in joules (J)

b its final speed.

BOX 30.1: Estimating the Planck constant

You can obtain an estimate of the value of the Planck constant h by means of a simple experiment. It makes use of light-emitting diodes (LEDs) of different colours (Figure 30.5). You may recall from Chapter 11 that an LED conducts in one direction only (the forward direction) and that it requires a minimum voltage, the **threshold voltage**, to be applied in this direction before it allows a current. This experiment makes use of the fact that LEDs of different colours require different threshold voltages before they conduct and emit light.

- A red LED emits photons that are of low energy. It requires a low threshold voltage to make it conduct.
- A blue LED emits higher-energy photons, and requires a higher threshold voltage to make it conduct.

What is happening to produce photons of light when an LED conducts? The simplest way to think of this is to say that the electrical energy lost by a single electron passing through the diode reappears as the energy of a single photon.

BOX 30.1: Estimating the Planck constant (continued)



Figure 30.5 Light-emitting diodes (LEDs) come in different colours. Blue (on the right) proved the trickiest to develop.

Hence we can write:

energy lost by electron = energy of photon

$$eV = \frac{hc}{\lambda}$$

where V is the threshold voltage for the LED. The values of e and c are known. Measurements of V and λ will allow you to calculate h. So the measurements required are:

- V – the voltage across the LED when it begins to conduct (its threshold voltage). It is found using a circuit like the one shown in Figure 30.6a
- λ – the wavelength of the light emitted by the LED. This is found by measurements using a diffraction grating or from the wavelength quoted by the manufacturer of the LED.

If several LEDs of different colours are available, V and λ can be determined for each and a graph of V against $\frac{1}{\lambda}$ drawn (see Figure 30.6b). The gradient of this graph will be $\frac{hc}{e}$ and hence h can be estimated.

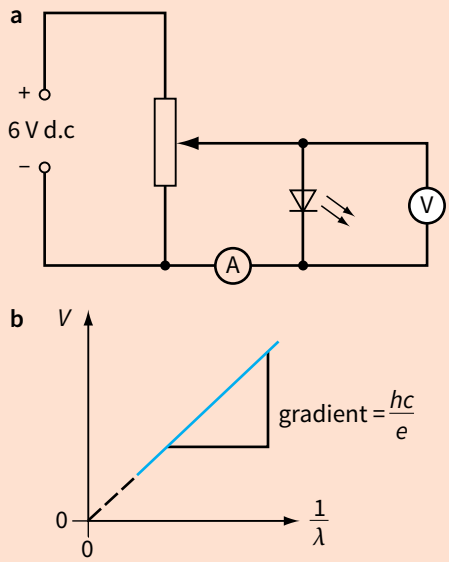


Figure 30.6 **a** A circuit to determine the threshold voltage required to make an LED conduct. An ammeter helps to show when this occurs. **b** The graph used to determine h from this experiment.

QUESTION

9 In an experiment to determine the Planck constant h, LEDs of different colours were used. The p.d. required to make each conduct was determined, and the wavelength of their light was taken from the manufacturer's catalogue. The results are shown in Table 30.3. For each LED, calculate the experimental value for h and hence determine an average value for the Planck constant.

Colour of LED	Wavelength / 10^{-9} m	Threshold voltage / V
infrared	910	1.35
red	670	1.70
amber	610	2.00
green	560	2.30

Table 30.3 Results from an experiment to determine h.

The photoelectric effect

In the photoelectric effect, light shines on a metal surface and electrons are released from it. The Greek word for light is photo, hence the word 'photoelectric'. The electrons removed from the metal plate in this manner are often known as photoelectrons.

The apparatus used to observe the photoelectric effect is shown in Box 30.2. Light from a lamp is shone onto a negatively charged metal plate and some of the electrons in the metal are emitted. A simple explanation is that light is a wave that carries energy and this energy releases electrons from the metal. However, detailed observations of the effect at first proved difficult to explain, in particular that there is a minimum threshold frequency of light below which no effect is observed.

BOX 30.2: Observing the photoelectric effect

You can observe the photoelectric effect yourself by fixing a clean zinc plate to the top of a gold-leaf electroscope (Figure 30.7). Give the electroscope a negative charge and the leaf deflects. Now shine electromagnetic radiation from a mercury discharge lamp on the zinc and the leaf gradually falls. (A mercury lamp strongly emits ultraviolet radiation.) Charging the electroscope gives it an excess of electrons. Somehow, the electromagnetic radiation from the mercury lamp helps electrons to escape from the surface of the metal.

Placing the mercury lamp closer causes the leaf to fall more rapidly. This is not very surprising. However, if you insert a sheet of glass between the lamp and the zinc, the radiation from the lamp is no longer effective. The gold leaf does not fall. Glass absorbs **ultraviolet** radiation and it is this component of the radiation from the lamp that is effective.

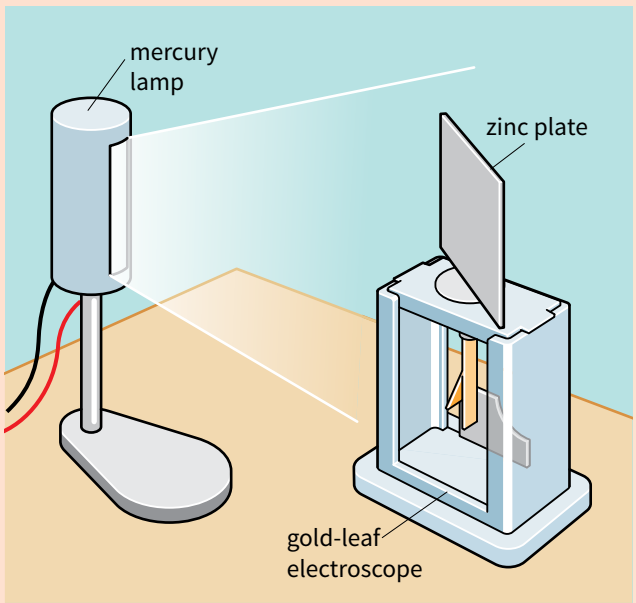


Figure 30.7 A simple experiment to observe the photoelectric effect.

Low frequency, high frequency

If you try the experiment described above with a bright filament lamp, you will find it has no effect. It does not produce ultraviolet radiation. There is a minimum frequency that the incident radiation must have in order to release electrons from the metal. This is called the **threshold frequency**. The threshold frequency is a property of the metal plate being exposed to electromagnetic radiation.

> The threshold frequency is defined as the minimum frequency required to release electrons from the surface of a metal.

Physicists found it hard to explain why weak ultraviolet radiation could have an immediate effect on the electrons in the metal, but very bright light of lower frequency had no effect. They imagined light waves arriving at the metal, spread out over its surface, and they could not see how weak ultraviolet waves could be more effective than the intense visible waves. In 1905, Albert Einstein came up with an explanation based on the idea of photons.

Metals (such as zinc) have electrons that are not very tightly held within the metal. These are the conduction electrons, and they are free to move about within the metal. When photons of electromagnetic radiation strike the metal, some electrons break free from the surface of the metal (Figure 30.8). They only need a small amount of energy (about 10^{-19} J) to escape from the metal surface.

We can picture the electrons as being trapped in an energy 'well' (Figure 30.9). A single electron requires a minimum energy Φ (Greek letter phi) to escape the surface of the metal. The **work function energy**, or simply **work function**, of a metal is the minimum amount of energy required by an electron to escape its surface. (Energy is needed to release the surface electrons because they are attracted by the electrostatic forces due to the positive metal ions.)

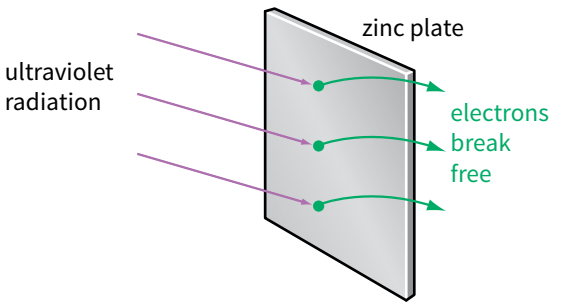


Figure 30.8 The photoelectric effect. When a photon of ultraviolet radiation strikes the metal plate, its energy may be sufficient to release an electron.

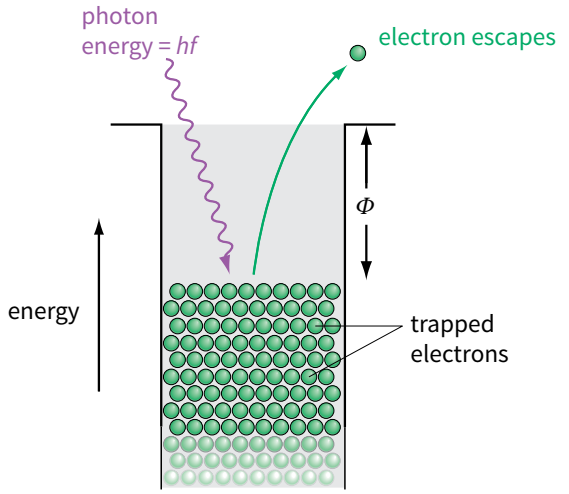


Figure 30.9 A single photon may interact with a single electron to release it.

Einstein did not picture electromagnetic waves interacting with all of the electrons in the metal. Instead, he suggested that a single photon could provide the energy needed by an individual electron to escape. The photon energy would need to be at least as great as Φ. By this means, Einstein could explain the threshold frequency. A photon of visible light has energy less than Φ, so it cannot release an electron from the surface of zinc.

When a photon arrives at the metal plate, it may be captured by an electron. The electron gains all of the photon's energy and the photon no longer exists. Some of the energy is needed for the electron to escape from the energy well; the rest is the electron's kinetic energy.

Now we can see that the photon model works because it models electromagnetic waves as concentrated 'packets' of energy, each one able to release an electron from the metal.

Here are some rules for the photoelectric effect:

- Electrons from the surface of the metal are removed.
- A single photon can only interact, and hence exchange its energy, with a single electron (one-to-one interaction).
- A surface electron is removed **instantaneously** from the metal surface when the energy of the incident photon is greater than, or equal to, the work function Φ of the metal. (The frequency of the incident radiation is greater than, or equal to, the threshold frequency of the metal.)
- Energy must be conserved when a photon interacts with an electron.
- Increasing the intensity of the incident radiation does not release a single electron when its frequency is less than the threshold frequency. The intensity of the incident radiation is proportional to the rate at which photons arrive at the plate. Each photon still has energy which is less than the work function.

Photoelectric experiments showed that the electrons released had a range of kinetic energies up to some maximum value, k.e._{max}. These fastest-moving electrons are the ones which were least tightly held in the metal.

Imagine a single photon interacting with a single surface electron and freeing it. According to Einstein:

energy of photon
= work function + maximum kinetic energy of electron

$$hf = \Phi + \text{k.e.}_{\text{max}}$$

or

$$hf = \Phi + \tfrac{1}{2}mv_{\text{max}}^2$$

This equation, known as Einstein's photoelectric equation, can be understood as follows:

- We start with a photon of energy hf.
- It is absorbed by an electron.
- Some of the energy (Φ) is used in escaping from the metal. The rest remains as kinetic energy of the electron.
- If the photon is absorbed by an electron that is lower in the energy well, the electron will have less kinetic energy than k.e._{max} (Figure 30.10).

What happens when the incident radiation has a frequency equal to the threshold frequency f_0 of the metal?

The kinetic energy of the electrons is zero. Hence, according to Einstein's photoelectric equation:

$$hf_0 = \Phi$$

Hence, the threshold frequency f_0 is given by the expression:

$$f_0 = \frac{\Phi}{h}$$

What happens when the incident radiation has frequency less than the threshold frequency? A single photon can still give up its energy to a single electron, but this electron

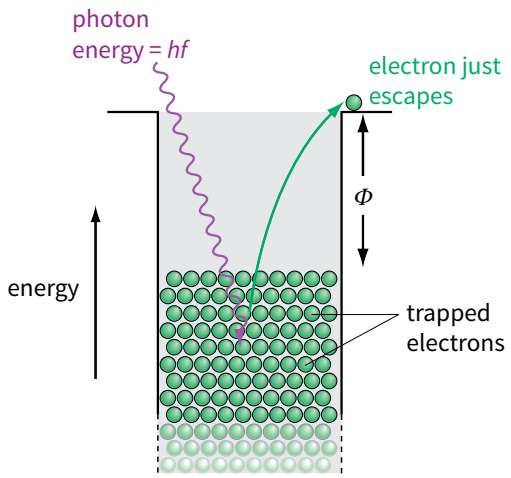


Figure 30.10 A more tightly bound electron needs more energy to release it from the metal.

Observation	Wave model	Photon model
Emission of electrons happens as soon as light shines on metal	Very intense light should be needed to have immediate effect	A single photon is enough to release one electron
Even weak (low-intensity) light is effective	Weak light waves should have no effect	Low-intensity light means fewer photons, not lower-energy photons
Increasing intensity of light increases rate at which electrons leave metal	Greater intensity means more energy, so more electrons are released	Greater intensity means more photons per second, so more electrons released per second
Increasing intensity has no effect on energies of electrons	Greater intensity should mean electrons have more energy	Greater intensity does not mean more energetic photons, so electrons cannot have more energy
A minimum threshold frequency of light is needed	Low-frequency light should work; electrons would be released more slowly	A photon in a low-frequency light beam has energy that is too small to release an electron
Increasing frequency of light increases maximum kinetic energy of electrons	It should be increasing intensity, not frequency, that increases energy of electrons	Higher frequency means more energetic photons; so electrons gain more energy and can move faster

Table 30.4 The success of the photon model in explaining the photoelectric effect.

cannot escape from the attractive forces of the positive metal ions. The energy absorbed from the photons appears as kinetic energy of the electrons. These electrons lose their kinetic energy to the metal ions when they collide with them. This warms up the metal. This is why a metal plate placed in the vicinity of a table lamp gets hot.

Different metals have different threshold frequencies, and hence different work functions. For example, alkali metals such as sodium, potassium and rubidium have threshold frequencies in the visible region of the electromagnetic spectrum. The conduction electrons in zinc are more tightly bound within the metal and so its threshold frequency is in the ultraviolet region of the spectrum.

Table 30.4 summarises the observations of the photoelectric effect, the problems a wave model of light has in explaining them, and how a photon model is more successful.

QUESTIONS

You will need these values to answer questions 10–13:

speed of light in a vacuum $c = 3.00 \times 10^8\,\text{m s}^{-1}$

Planck constant $h = 6.63 \times 10^{-34}\,\text{J s}$

mass of electron $m_e = 9.11 \times 10^{-31}\,\text{kg}$

elementary charge $e = 1.60 \times 10^{-19}\,\text{C}$

10 Photons of energies 1.0 eV, 2.0 eV and 3.0 eV strike a metal surface whose work function is 1.8 eV.

a State which of these photons could cause the release of an electron from the metal.

b Calculate the maximum kinetic energies of the electrons released in each case. Give your answers in eV and in J.

11 Table 30.5 shows the work functions of several different metals.

a Which metal requires the highest frequency of electromagnetic waves to release electrons?

b Which metal will release electrons when the lowest frequency of electromagnetic waves is incident on it?

c Calculate the threshold frequency for zinc.

d What is the longest wavelength of electromagnetic waves that will release electrons from potassium?

Metal	Work function Φ / J	Work function Φ / eV
caesium	3.0×10^{-19}	1.9
calcium	4.3×10^{-19}	2.7
gold	7.8×10^{-19}	4.9
potassium	3.2×10^{-19}	2.0
zinc	6.9×10^{-19}	4.3

Table 30.5 Work functions of several different metals.

QUESTIONS

12 Electromagnetic waves of wavelength 2.4×10^{-7} m are incident on the surface of a metal whose work function is 2.8×10^{-19} J.

- **a** Calculate the energy of a single photon.
- **b** Calculate the maximum kinetic energy of electrons released from the metal.
- **c** Determine the maximum speed of the emitted photoelectrons.

13 When electromagnetic radiation of wavelength 2000 nm is incident on a metal surface, the maximum kinetic energy of the electrons released is found to be 4.0×10^{-20} J. Determine the work function of the metal in joules (J).

Line spectra

We will now look at another phenomenon which we can explain in terms of light as photons. We rely a great deal on light to inform us about our surroundings. Using our eyes we can identify many different colours. Scientists take this further by analysing light, by breaking or splitting it up into a spectrum. (The technical term for the splitting of light into its components is **dispersion**.) You will be familiar with the ways in which this can be done, using a prism or a diffraction grating (Figure 30.11). The spectrum of white light shows that it consists of a range of wavelengths, from about 4×10^{-7} m (violet) to about 7×10^{-7} m (red), as in Figure 30.12a. This is a continuous spectrum.

It is more interesting to look at the spectrum from a hot gas. If you look at a lamp that contains a gas such as neon or sodium, you will see that only certain colours are present. Each colour has a unique wavelength. If the source is narrow and it is viewed through a diffraction grating, a **line spectrum** is seen.

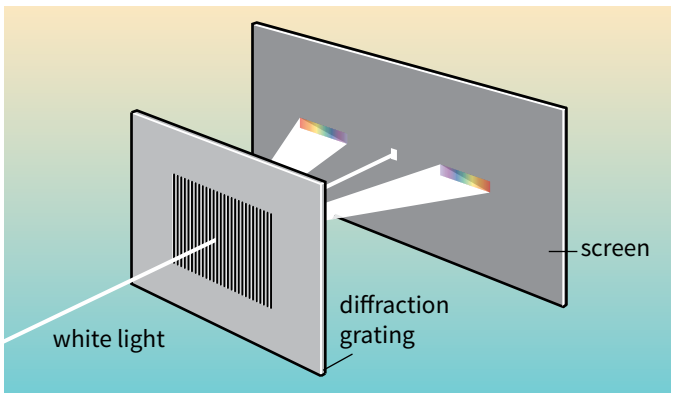


Figure 30.11 White light is split up into a continuous spectrum when it passes through a diffraction grating.

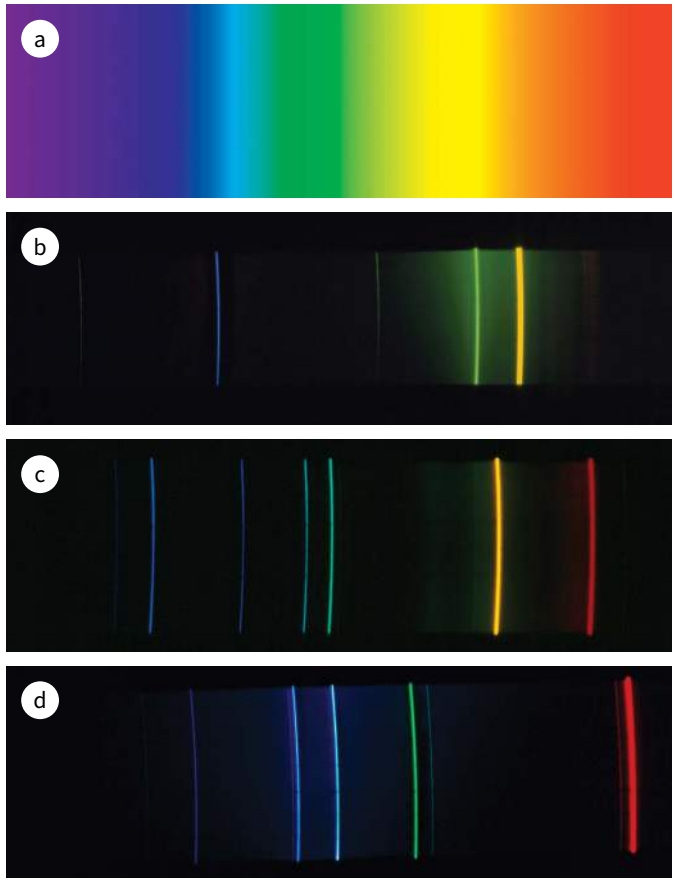


Figure 30.12 Spectra of **a** white light, and of light from **b** mercury, **c** helium and **d** cadmium vapour.

Figures 30.12b–d show the line spectra of hot gases of the elements mercury, helium and cadmium. Each element has a spectrum with a unique collection of wavelengths. Therefore line spectra can be used to identify elements. This is exactly what the British astronomer William Huggins did when he deduced which elements are the most common in the stars.

These line spectra, which show the composition of light emitted by hot gases, are called **emission line spectra**.

There is another kind of spectra, called **absorption line spectra**, which are observed when white light is passed through cool gases. After the light has passed through a diffraction grating, the continuous white light spectrum is

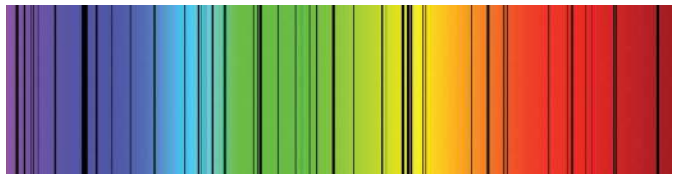


Figure 30.13 An absorption line spectrum formed when white light is passed through cool mercury vapour.

found to have black lines across it (Figure 30.13). Certain wavelengths have been absorbed as the white light passed through the cool gas.

Absorption line spectra are found when the light from stars is analysed. The interior of the star is very hot and emits white light of all wavelengths in the visible range. However, this light has to pass through the **cooler** outer layers of the star. As a result, certain wavelengths are absorbed. Figure 30.14 shows the spectrum for the Sun.

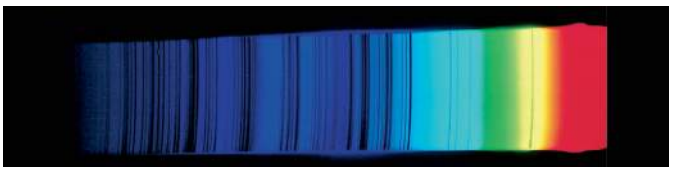


Figure 30.14 The Sun's spectrum shows dark lines. These dark lines arise when light of specific wavelengths coming from the Sun's hot interior is absorbed by its cooler atmosphere.

Explaining the origin of line spectra

From the description above, we can see that the atoms of a given element (e.g. helium) can only emit or absorb light of certain wavelengths.

Different elements emit and absorb different wavelengths. How can this be? To understand this, we need to establish two points:

- First, as with the photoelectric effect, we are dealing with light (an electromagnetic wave) interacting with matter. Hence we need to consider light as consisting of photons. For light of a single wavelength λ and frequency f, the energy E of each photon is given by the equation:

  $E = hf$ or $E = \frac{hc}{\lambda}$

- Secondly, when light interacts with matter, it is the electrons that absorb the energy from the incoming photons. When the electrons lose energy, light is emitted by matter in the form of photons.

What does the appearance of the line spectra tell us about electrons in atoms? They can only absorb or emit photons of certain energies. From this we deduce that electrons in atoms can themselves only have certain fixed values of energy. This idea seemed very odd to scientists a hundred years ago. Figure 30.15 shows diagrammatically the permitted **energy levels** (or **energy states**) of the electron of a hydrogen atom. An electron in a hydrogen atom can have only one of these values of energy. It cannot have an energy that is between these energy levels. The energy levels of the electron are analogous to the rungs of a ladder. The energy levels have **negative** values because external energy has

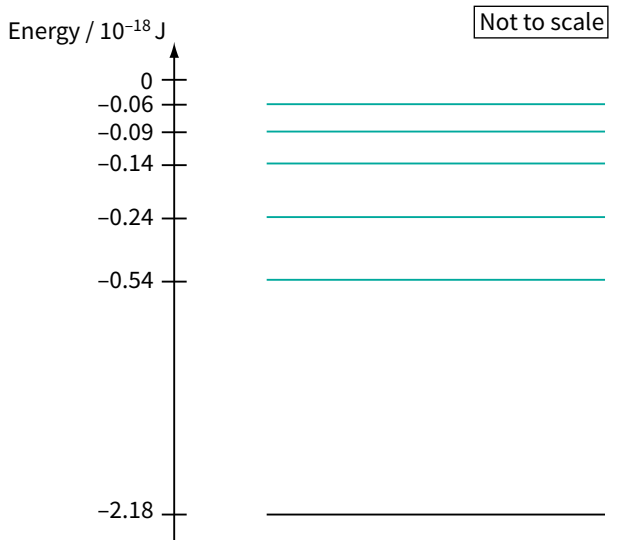


Figure 30.15 Some of the energy levels of the hydrogen atom.

to be supplied to remove an electron from the atom. The negative energy shows that the electron is trapped within the atom by the attractive forces of the atomic nucleus. An electron with zero energy is free from the atom.

The energy of the electron in the atom is said to be **quantised**. This is one of the most important statements of quantum physics.

Now we can explain what happens when an atom emits light. One of its electrons falls from a high energy level to a lower one (Figure 30.16a). The electron makes a **transition** to a lower energy level. The loss of energy of the electron leads to the emission of a single photon of light. The energy of this photon is exactly equal to the energy difference between the two energy levels. If the electron makes a transition from a higher energy level, the energy loss of the electron is larger and this leads to the emission of a more energetic photon. The distinctive energy levels of an atom mean that the energy of the photons emitted, and hence the wavelengths emitted, will be unique to that atom. This explains why only certain wavelengths are present in the emission line spectrum of a hot gas.

Atoms of different elements have different line spectra because they have different spacings between their energy levels. It is not within the scope of this book to discuss why this is.

Similarly, we can explain the origin of absorption line spectra. White light consists of photons of many different energies. For a photon to be absorbed, it must have exactly the right energy to lift an electron from one energy level to another (Figure 30.16b). If its energy is too little or too great, it will not be absorbed. This effect can also described as a form of resonance (Chapter 19) – the frequency of the photon must be such that its energy matches the gap between the two energy levels.

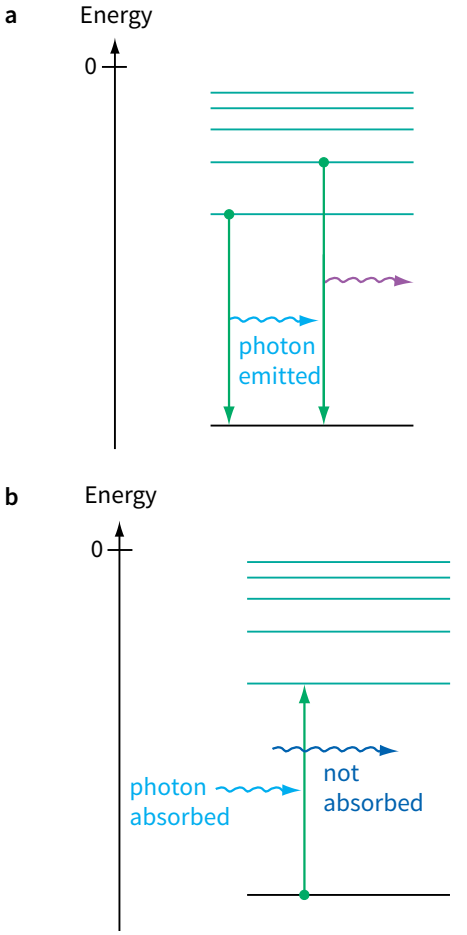


Figure 30.16 **a** When an electron drops to a lower energy level, it emits a single photon. **b** A photon must have just the right energy if it is to be absorbed by an electron.

Photon energies

When an electron changes its energy from one level E_1 to another E_2, it either emits or absorbs a **single** photon. The energy of the photon hf is simply equal to the **difference** in energies between the two levels:

$$\text{photon energy} = \Delta E$$

$$hf = E_1 - E_2$$

or

$$\frac{hc}{\lambda} = E_1 - E_2$$

Referring back to the energy level diagram for hydrogen (Figure 30.15), you can see that, if an electron falls from the second level to the lowest energy level (known as the **ground state**), it will emit a photon of energy:

$$\text{photon energy} = \Delta E$$

$$hf = [(-0.54) - (-2.18)] \times 10^{-18}\,\text{J}$$

$$hf = 1.64 \times 10^{-18}\,\text{J}$$

We can calculate the frequency f and wavelength λ of the emitted electromagnetic radiation.

The frequency is:

$$f = \frac{E}{h} = \frac{1.64 \times 10^{-18}}{6.63 \times 10^{-34}}$$

$$f = 2.47 \times 10^{15}\,\text{Hz}$$

The wavelength is:

$$\lambda = \frac{c}{f} = \frac{3.00 \times 10^{8}}{2.47 \times 10^{15}}$$

$$\lambda = 1.21 \times 10^{-7}\,\text{m} = 121\,\text{nm}$$

This is a wavelength in the ultraviolet region of the electromagnetic spectrum.

QUESTION

14 Figure 30.17 shows part of the energy level diagram of an imaginary atom. The arrows represent three transitions between the energy levels. For each of these transitions:

- **a** calculate the energy of the photon
- **b** calculate the frequency and wavelength of the electromagnetic radiation (emitted or absorbed)
- **c** state whether the transition contributes to an emission or an absorption spectrum.

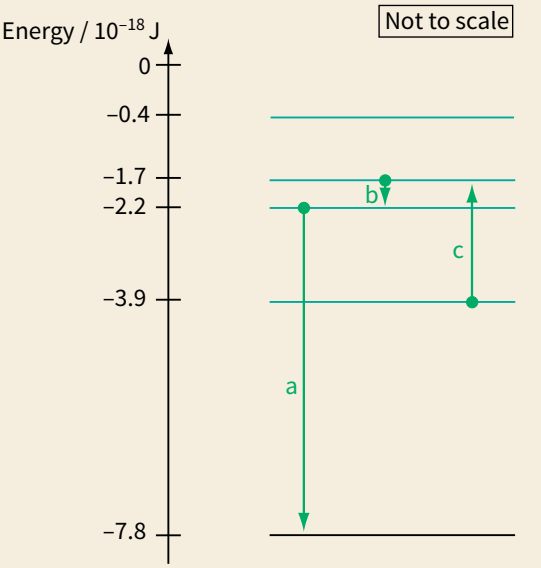


Figure 30.17 An atomic energy level diagram, showing three electron transitions between levels – see Question 14.

QUESTIONS

15 Figure 30.18 shows another energy level diagram. In this case, energies are given in electronvolts (eV). From the list below, state which photon energies could be absorbed by such an atom:
6.0 eV 9.0 eV 11 eV 20 eV 25 eV 34 eV 45 eV

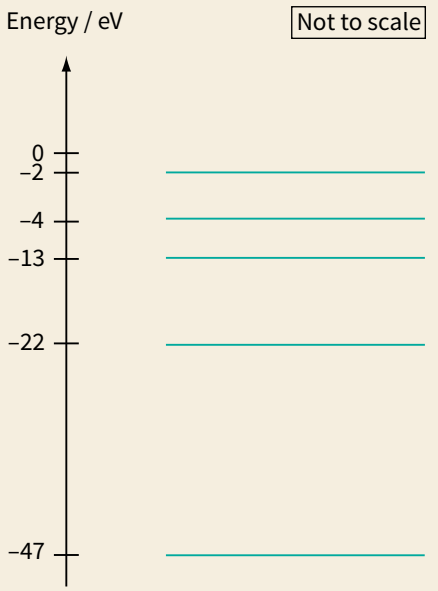


Figure 30.18 An energy level diagram – see Question 15.

16 The line spectrum for a particular type of atom is found to include the following wavelengths:
83 nm 50 nm 25 nm

a Calculate the corresponding photon energies in eV.

b Sketch the energy levels which could give rise to these photons. On the diagram indicate the corresponding electron transitions responsible for these three spectral lines.

Electron energies in solids

So far, we have only discussed the spectra of light from hot gases. In a gas, the atoms are relatively far apart, so they do not interact with one another very much. Gas atoms that exert negligible electrical forces on each other are known as **isolated atoms**. As a consequence, they give relatively simple line spectra. Similar spectra can be obtained from some gemstones and coloured glass. In these, the basic material is clear and colourless, but it gains its colour from impurity atoms, which are well separated from one another within the material.

In a solid or liquid, however, the atoms are close together. The electrons from one atom interact with those of neighbouring atoms. This has the effect of altering the energy level diagram, which becomes much more complicated, with a large number of closely spaced energy levels.

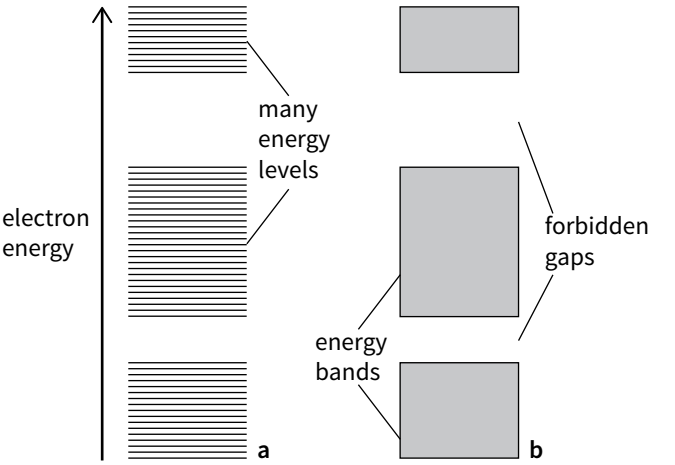


Figure 30.19 **a** In a solid, the electron energy levels are very close together. **b** The energy levels form bands with forbidden gaps between them.

Figure 30.19a shows the result. Instead of individual energy levels we have bands of many levels, all close together. In between the bands are gaps with no allowed energy levels.

Figure 30.19b shows a more conventional representation of these **energy bands** in a solid. An electron can have an energy at any level in one of the bands. However, it cannot have an energy value which lies in the **forbidden gap** between bands (just as an electron in an isolated atom cannot have an energy which lies between two energy levels).

Band theory and electrical conduction

We can use this **band theory** of solids to explain why some materials are better conductors than others. In Figure 30.20, the energy bands are shown with green shading where they are occupied and with grey shading where they are unoccupied.

- In a metal, one band, known as the **conduction band**, is only partially filled. The electrons in the conduction band are the conduction or free electrons which give the metal its conductivity, as discussed in Chapter 11.
- In an insulator, the conduction band is unoccupied. The band below this, known as the **valence band**, is fully occupied. An electron whose energy lies in the valence band is bound to an individual atom.

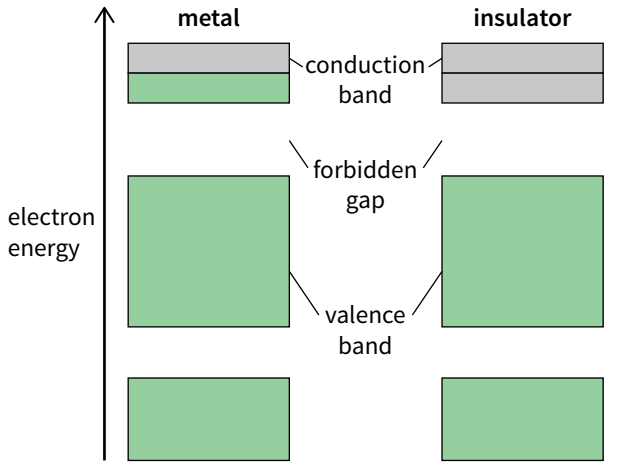


Figure 30.20 How the energy bands are filled in a metal and in an insulator.

How does this explain the difference between metals and insulators? When a piece of metal is connected to a cell, the electrons in the metal gain energy. For this to happen, there must be empty energy levels higher in the band into which these electrons can move. When they have moved upwards they are free to move through the metal, and so there is a current.

In an insulator, there is a large energy gap between the top of the valence band and the bottom of the conduction band. The voltage of a cell is insufficient to lift even the most energetic electrons across the gap and into the conduction band. This means that electrons are not free to move through the material – it is an insulator.

(In some metals, the top of the valence band overlaps with the bottom of the conduction band, so there are plenty of free electrons in the conduction band to carry a current.)

A semiconductor is a material which conducts electric current, but only very slightly. As shown in Figure 30.21, it is similar to an insulator. Its valence band is full and its conduction band is empty. However, the gap between the two is very small. At room temperature, a few electrons have enough energy to jump across the gap into the conduction band. These electrons are 'free' and can form a current.

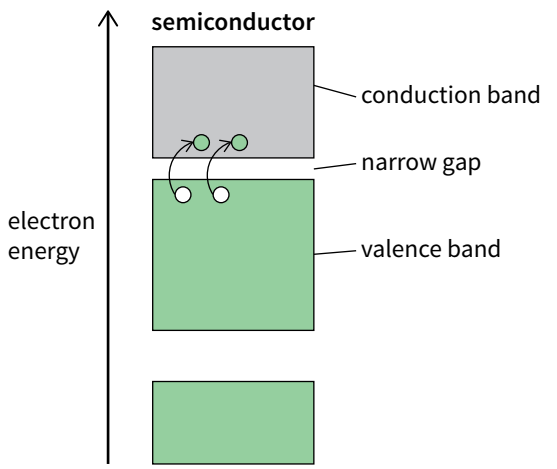


Figure 30.21 Electron energy bands in a semiconductor.

As we saw in Chapter 9 when discussing the drift velocity of electrons, a metal typically has one free electron per atom. A semiconductor has perhaps one free electron per million atoms at room temperature, and so its electrical conductivity will be of the order of one-millionth of that of a metal.

If a piece of semiconductor is heated, more electrons will gain the energy needed to jump up into the conduction band and the material will conduct better – its resistance decreases because of the increased number density of electrons in the conduction band. This is the opposite of a metal, whose resistance increases when it is heated. There is no increase in the number density of free electrons when a metal is heated. Instead, its atoms vibrate more, and the electrons collide more frequently with the vibrating atoms.

Silicon and germanium are examples of semiconductor materials like this. They are described as **intrinsic semiconductors** because their conductivity is a property of the pure material itself. Diodes, transistors and computer chips use semiconductors which have small amounts of other elements added to them to increase their conductivity.

There is another way to give electrons in a semiconductor the energy needed to jump across the forbidden gap. In a light-dependent resistor (LDR), photons of light are absorbed by electrons in the valence band so that they jump the gap and enter the conduction band. This is why an LDR has high resistance in the dark but its resistance decreases as the light level increases.

QUESTIONS

17 What materials are being described here?

- **a** Valence band full, conduction band empty, wide forbidden gap
- **b** Valence band full, conduction band partially filled
- **c** Number density of conduction electrons increases rapidly with increasing temperature.

18 Use band theory to explain why the resistivity of an intrinsic semiconductor increases as the temperature decreases.

The nature of light – waves or particles?

It is clear that, in order to explain the photoelectric effect, we must use the idea of light (and all electromagnetic radiation) as particles. Similarly, photons explain the appearance of line spectra. However, to explain diffraction, interference and polarisation of light, we must use the wave model. How can we sort out this dilemma?

We have to conclude that sometimes light shows wave-like behaviour; at other times it behaves as particles (photons). In particular, when light is absorbed by a metal surface, it behaves as particles. Individual photons are absorbed by individual electrons in the metal. In a similar way, when a Geiger counter detects γ-radiation, we hear individual γ-photons being absorbed in the tube.

So what is light? Is it a wave or a particle? Physicists have come to terms with the dual nature of light. This duality is referred to as the **wave–particle duality** of light. In simple terms:

- Light interacts with matter (e.g. electrons) as a particle – the photon. The evidence for this is provided by the photoelectric effect.
- Light travels through space as a wave. The evidence for this comes from the diffraction and interference of light using slits.

Electron waves

Light has a dual nature. Is it possible that particles such as electrons also have a dual nature? This interesting question was first contemplated by Louis de Broglie (pronounced 'de Broy') in 1924 (Figure 30.22).



Figure 30.22 Louis de Broglie provided an alternative view of how particles behave.

De Broglie imagined that electrons would travel through space as a wave. He proposed that the wave-like property of a particle like the electron can be represented by its wavelength λ, which is related to its momentum p by the equation:

$$\lambda = \frac{h}{p}$$

where h is the Planck constant. The wavelength λ is often referred to as the **de Broglie wavelength**. The waves associated with the electron are referred to as **matter waves**.

The momentum p of a particle is the product of its mass m and its velocity v. Therefore, the de Broglie equation may be written as:

$$\lambda = \frac{h}{mv}$$

The Planck constant h is the same constant that appears in the equation $E = hf$ for the energy of a photon. It is fascinating how the Planck constant h is entwined with the behaviour of both matter as waves (e.g. electrons) and electromagnetic waves as 'particles' (photons).

The wave property of the electron was eventually confirmed in 1927 by researchers in America and in England. The Americans Clinton Davisson and Edmund Germer showed experimentally that electrons were diffracted by single crystals of nickel. The diffraction of electrons confirmed their wave-like property. In England, George Thomson fired electrons into thin sheets of metal in a vacuum tube. He, too, provided evidence that electrons were diffracted by the metal atoms.

Louis de Broglie received the 1929 Nobel Prize in Physics. Clinton Davisson and George Thomson shared the Nobel Prize in Physics in 1937.

Electron diffraction

We can reproduce the same diffraction results in the laboratory using an electron diffraction tube; see Figure 30.23.

In an electron diffraction tube, the electrons from the heated filament are accelerated to high speeds by the large potential difference between the negative heater (cathode) and the positive electrode (anode). A beam of electrons passes through a thin sample of polycrystalline graphite. It is made up of many tiny crystals, each of which consists of large numbers of carbon atoms arranged in uniform atomic layers. The electrons emerge from the graphite film and produce diffraction rings on the phosphor screen. The diffraction rings are similar to those produced by light (a wave) passing through a small circular hole. The rings cannot be explained if electrons behaved as

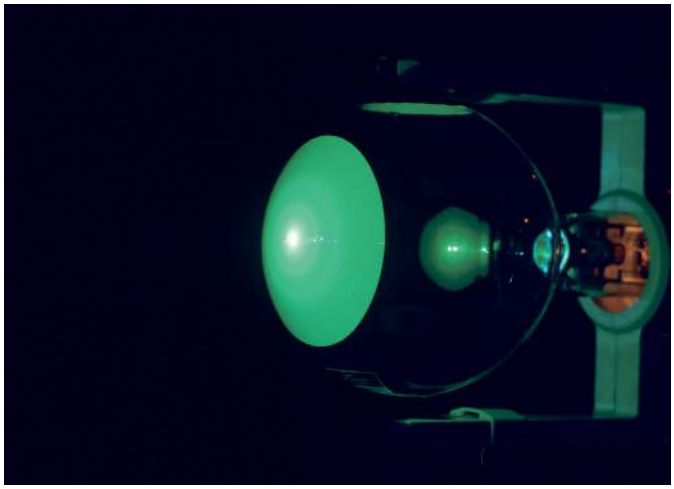


Figure 30.23 When a beam of electrons passes through a graphite film, as in this vacuum tube, a diffraction pattern is produced on the phosphor screen.

particles. Diffraction is a property of waves. Hence the rings can only be explained if the electrons pass through the graphite film as a wave. The electrons are diffracted by the carbon atoms and the spacing between the layers of carbon atoms. The atomic layers of carbon behave like a diffraction grating with many slits. The electrons show diffraction effects because their de Broglie wavelength λ is similar to the spacing between the atomic layers.

This experiment shows that electrons appear to travel as waves. If we look a little more closely at the results of the experiment, we find something else even more surprising. The phosphor screen gives a flash of light for each electron that hits it. These flashes build up to give the diffraction pattern (Figure 30.24). But if we see flashes at particular points on the screen, are we not seeing individual electrons – in other words, are we not observing particles?

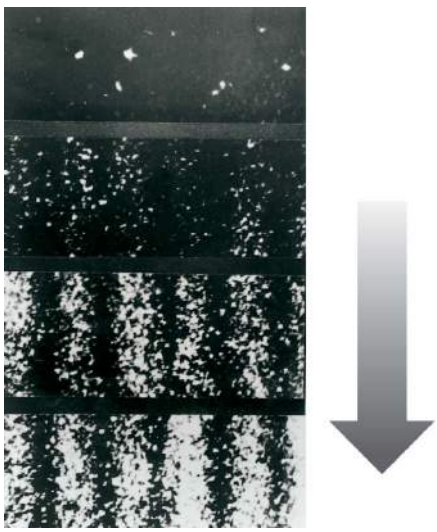


Figure 30.24 The speckled diffraction pattern shows that it arises from many individual electrons striking the screen.

BOX 30.3: Investigating electron diffraction

If you have access to an electron diffraction tube (Figure 30.25), you can see for yourself how a beam of electrons is diffracted. The electron gun at one end of the tube produces a beam of electrons. By changing the voltage between the anode and the cathode, you can change the energy of the electrons, and hence their speed. The beam strikes a graphite target, and a diffraction pattern appears on the screen at the other end of the tube.

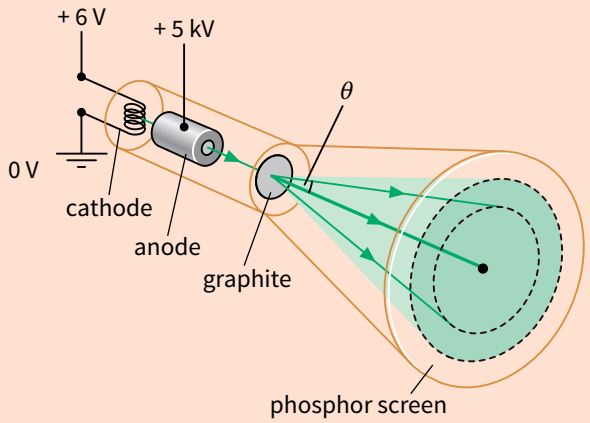


Figure 30.25 Electrons are accelerated from the cathode to the anode; they form a beam which is diffracted as it passes through the graphite film.

You can use an electron diffraction tube to investigate how the wavelength of the electrons depends on their speed. Qualitatively, you should find that increasing the anode–cathode voltage makes the pattern of diffraction rings shrink. The electrons have more kinetic energy (they are faster); the shrinking pattern shows that their wavelength has decreased. You can find the wavelength λ of the electrons by measuring the angle θ at which they are diffracted:

$$\lambda = 2d \sin\theta$$

where d is the spacing of the atomic layers of graphite.

You can find the speed of the electrons from the anode–cathode voltage V:

$$\tfrac{1}{2}mv^2 = eV$$

WORKED EXAMPLE

1 Calculate the de Broglie wavelength of an electron travelling through space at a speed of $10^7\,m\,s^{-1}$. State whether or not these electrons can be diffracted by solid materials (atomic spacing in solid materials ~ 10^{-10} m).

Step 1 According to the de Broglie equation, we have:

$$\lambda = \frac{h}{mv}$$

Step 2 The mass of an electron is 9.11×10^{-31} kg. Hence:

$$\lambda = \frac{6.63 \times 10^{-34}}{9.11 \times 10^{-31} \times 10^7} = 7.3 \times 10^{11}\,m$$

Electrons travelling at $10^7\,m\,s^{-1}$ have a de Broglie wavelength of order of magnitude 10^{-10} m. Hence they can be diffracted by matter.

QUESTION

19 X-rays are used to find out about the spacings of atomic planes in crystalline materials.

a Describe how beams of electrons could be used for the same purpose.

b How might electron diffraction be used to identify a sample of a metal?

People waves

The de Broglie equation applies to all matter; anything that has mass. It can also be applied to objects like golf balls and people!

Imagine a 65 kg person running at a speed of $3.0\,m\,s^{-1}$ through an opening of width 0.80 m. According to the de Broglie equation, the wavelength of this person is:

$$\lambda = \frac{h}{mv}$$

$$\lambda = \frac{6.63 \times 10^{-34}}{65} \times 3.0$$

$$\lambda = 3.4 \times 10^{-36}\,m$$

This wavelength is very small indeed compared with the size of the gap, hence no diffraction effects would be observed. People cannot be diffracted through everyday gaps. The de Broglie wavelength of this person is much smaller than any gap the person is likely to try to squeeze through! For this reason, we do not use the wave model to describe the behaviour of people; we get much better results by regarding people as large particles.

QUESTION

20 A beam of electrons is accelerated from rest through a p.d. of 1.0 kV.

a What is the energy (in eV) of each electron in the beam?

b Calculate the speed, and hence the momentum (mv), of each electron.

c Calculate the de Broglie wavelength of each electron.

d Would you expect the beam to be significantly diffracted by a metal film in which the atoms are separated by a spacing of 0.25×10^{-9} m?

Probing matter

All moving particles have a de Broglie wavelength. The structure of matter can be investigated using the diffraction of particles. Diffraction of slow-moving neutrons (known as thermal neutrons) from nuclear reactors is used to study the arrangements of atoms in metals and other materials. The wavelength of these neutrons is about 10^{-10} m, which is roughly the separation between the atoms.

Diffraction of slow-moving electrons is used to explore the arrangements of atoms in metals (Figure 30.26) and the structures of complex molecules such as DNA (Figure 30.27). It is possible to accelerate electrons to the right speed so that their wavelength is similar to the spacing between atoms, around 10^{-10} m.

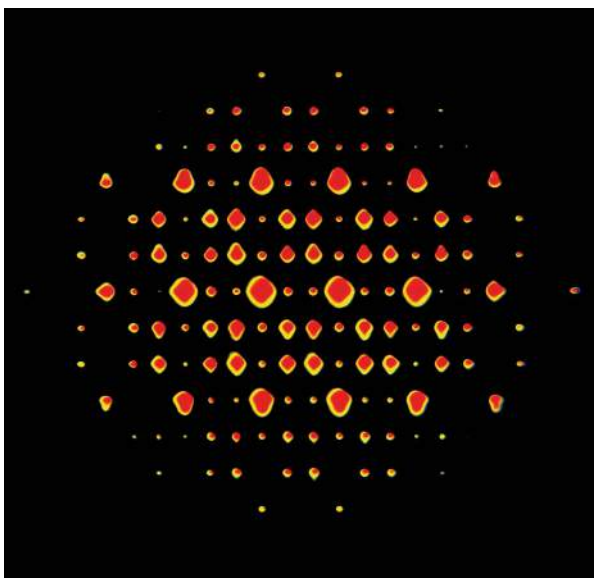


Figure 30.26 Electron diffraction pattern for an alloy of titanium and nickel. From this pattern, we can deduce the arrangement of the atoms and their separations.

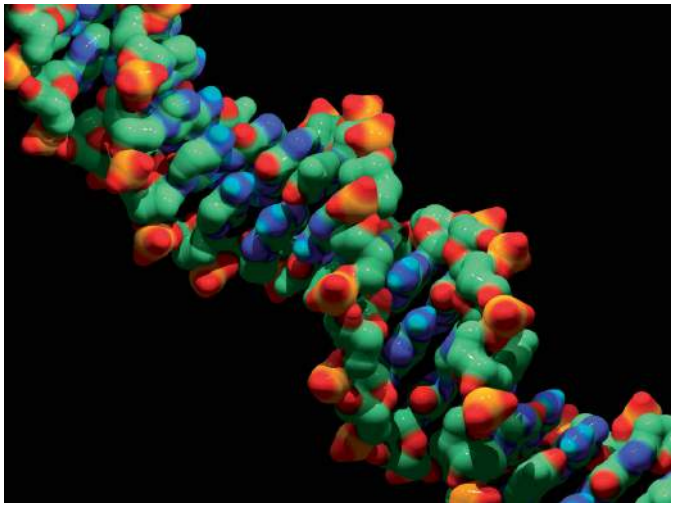


Figure 30.27 The structure of the giant molecule DNA, deduced from electron diffraction.

High-speed electrons from particle accelerators have been used to determine the diameter of atomic nuclei. This is possible because high-speed electrons have wavelengths of order of magnitude 10^{-15} m. This wavelength is similar to the size of atomic nuclei. Electrons travelling close to the speed of light are being used to investigate the internal structure of the nucleus. These electrons have to be accelerated by voltages up to 10^9 V.

The nature of the electron – wave or particle?

The electron has a dual nature, just like electromagnetic waves. This duality is referred to as the **wave–particle duality** of the electron. In simple terms:

- An electron interacts with matter as a particle. The evidence for this is provided by Newtonian mechanics.
- An electron travels through space as a wave. The evidence for this comes from the diffraction of electrons.

Summary

- For electromagnetic waves of frequency f and wavelength λ, each photon has energy E given by:

  $E = hf$ or $E = \frac{hc}{\lambda}$ where h is the Planck constant.

- One electronvolt is the energy transferred when an electron travels through a potential difference of 1 V.

  $1\,\text{eV} = 1.60 \times 10^{-19}\,\text{J}$

- A particle of charge e accelerated through a voltage V has kinetic energy given by:

  $eV = \frac{1}{2}mv^2$

- The photoelectric effect is an example of a phenomenon explained in terms of the particle-like (photon) behaviour of electromagnetic radiation.
- Einstein's photoelectric equation is:

  $hf = \Phi + \text{k.e.}_{\text{max}}$

  where Φ = work function = minimum energy required to release an electron from the metal surface.

- The threshold frequency is the minimum frequency of the incident electromagnetic radiation that will release an electron from the metal surface.
- Electron diffraction is an example of a phenomenon explained in terms of the wave-like behaviour of matter.
- The de Broglie wavelength λ of a particle is related to its momentum (mv) by the de Broglie equation:

  $\lambda = \frac{h}{mv}$

- Both electromagnetic radiation (e.g. light) and matter (e.g. electrons) exhibit wave–particle duality; that is, they show both wave-like and particle-like behaviours, depending on the circumstances. In wave–particle duality:
  - interaction is explained in terms of particles
  - travel through space is explained in terms of waves.
- Line spectra arise for isolated atoms (the electrical forces between such atoms is negligible).
- The energy of an electron in an isolated atom is quantised. The electron is allowed to exist in specific energy states known as energy levels.
- An electron loses energy when it makes a transition from a higher energy level to a lower energy level. A photon of electromagnetic radiation is emitted because of this energy loss. The result is an emission line spectrum.
- Absorption line spectra arise when electromagnetic radiation is absorbed by isolated atoms. An electron absorbs a photon of the correct energy to allow it to make a transition to a higher energy level.
- The frequency f and the wavelength λ of the emitted or absorbed radiation are related to the energy levels E_1 and E_2 by the equations:

  $hf = \Delta E = E_1 - E_2$ and $\frac{hc}{\lambda} = \Delta E = E_1 - E_2$

- In solids, electrons can exist in energy states within broad bands separated by forbidden gaps. This band theory can explain the different electrical behaviours of metals, insulators and semiconductors.

End-of-chapter questions

1 Calculate the energy of a photon of frequency 4.0×10^{18} Hz. [2]

2 The microwave region of the electromagnetic spectrum is considered to have wavelengths ranging from 5 mm to 50 cm. Calculate the range of energy of microwave photons. [3]

3 In a microwave oven photons of energy 1.02×10^{-5} eV are used to heat food.

a Express 1.02×10^{-5} eV in joules. [1]

b Calculate the frequency of the photons. [1]

c Calculate the wavelength of the photons. [1]

4 **a** Alpha-particles of energy 5 MeV are emitted in the radioactive decay of radium. Express this energy in joules. [1]

b Electrons in an cathode-ray tube are accelerated through a potential difference of 10 kV. State their energy:

i in electronvolts [1]

ii in joules. [1]

c In a nuclear reactor, neutrons are slowed to energies of 6×10^{-21} J. Express this in eV. [1]

5 A helium nucleus (charge = $+3.2 \times 10^{-19}$ C, mass = 6.8×10^{-27} kg) is accelerated through a potential difference of 7500 V.

Calculate:

a its kinetic energy in electronvolts [2]

b its kinetic energy in joules [1]

c its speed. [2]

6 Ultraviolet light with photons of energy 2.5×10^{-18} J is shone onto a zinc plate. The work function of zinc is 4.3 eV.

Calculate the maximum energy with which an electron can be emitted from the zinc plate.
Give your answer:

a in eV [3]

b in J. [1]

7 Calculate the minimum frequency of electromagnetic radiation that will cause the emission of photoelectrons from the surface of gold. [2]
(Work function for gold = 4.9 eV.)

8 Figure 30.28 shows five of the energy levels in a helium ion. The lowest energy level is known as the ground state.

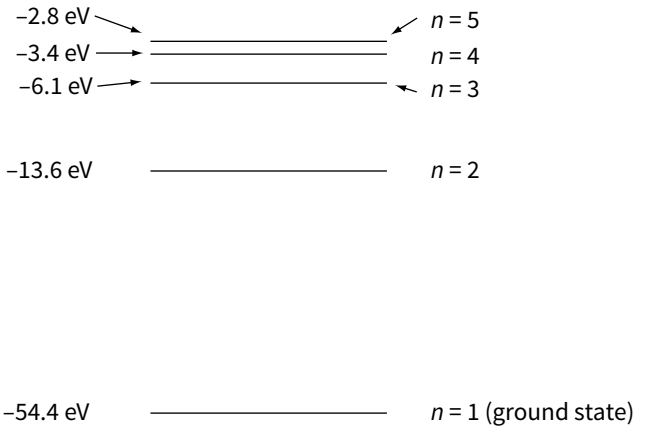


Figure 30.28 For End-of-chapter Question 8.

a Calculate the energy, in joules, that is required to completely remove the remaining electron, originally in its ground state, from the helium ion. [2]

b Calculate the frequency of the radiation which is emitted when the electron drops from the level $n = 3$ to $n = 2$. State the region of the electromagnetic spectrum in which this radiation lies. [2]

c Without further calculation, describe qualitatively how the frequency of the radiation emitted when the electron drops from the level $n = 2$ to $n = 1$ compares with the energy of the radiation emitted when it drops from $n = 3$ to $n = 2$. [2]

9 The spectrum of sunlight has dark lines. These dark lines are due to the absorption of certain wavelengths by the cooler gases in the atmosphere of the Sun.

a One particular dark spectral line has a wavelength of 590 nm. Calculate the energy of a photon with this wavelength. [3]

b Figure 30.29 shows some of the energy levels of an isolated atom of helium.

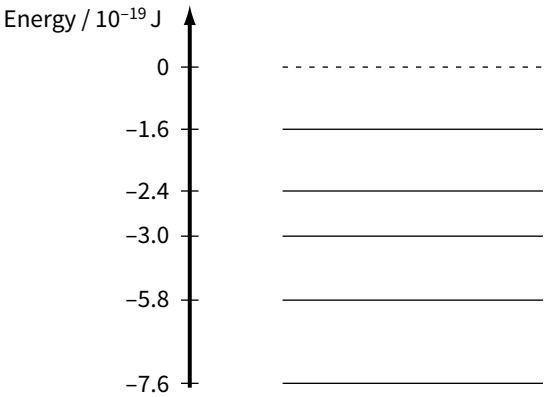


Figure 30.29 For End-of-chapter Question 9.

i Explain the significance of the energy levels having negative values. [1]

ii Explain, with reference to the energy level diagram above, how a dark line in the spectrum may be due to the presence of helium in the atmosphere of the Sun. [2]

iii All the light absorbed by the atoms in the Sun's atmosphere is re-emitted. Suggest why a dark spectral line of wavelength of 590 nm is still observed from the Earth. [1]

10 Figure 30.30 shows three of the energy levels in an isolated hydrogen atom. The lowest energy level is known as the ground state.

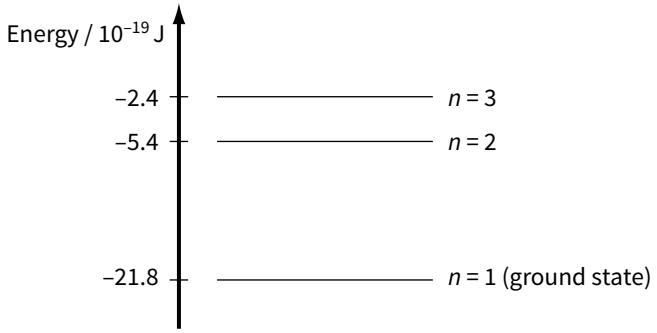


Figure 30.30 For End-of-chapter Question 10.

a Explain what happens to an electron in the ground state when it absorbs the energy from a photon energy 21.8×10^{-19} J. [1]

b **i** Explain why a photon is emitted when an electron makes a transition between energy levels $n = 3$ and $n = 2$. [2]

ii Calculate the wavelength of electromagnetic radiation emitted when an electron makes a jump between energy levels $n = 3$ and $n = 2$. [3]

iii In the diagram, each energy level is labeled with its 'principal quantum number' n. Use the energy level diagram to show that the energy E of an energy level is inversely proportional to n^2. [2]

11 **a** **i** Explain what is meant by the **wave–particle duality** of electromagnetic radiation. [2]

ii Explain how the photoelectric effect gives evidence for this phenomenon. [2]

Figure 30.31 shows the maximum kinetic energy of the emitted photoelectrons as the frequency of the incident radiation on a sodium plate is varied.

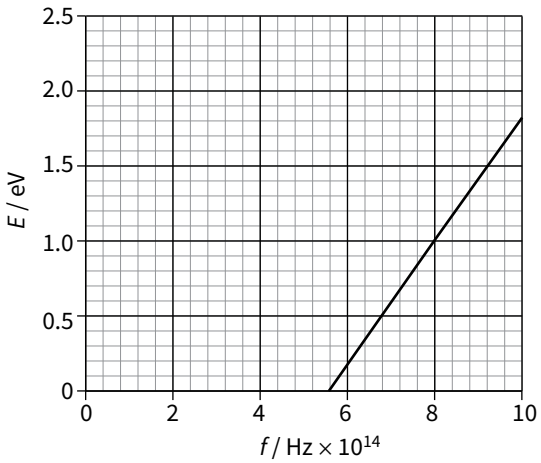


Figure 30.31 For End-of-chapter Question 11.

b Explain why there are no photoelectrons emitted when the frequency of the incident light is less than 5.6×10^{14} Hz. [2]

c Calculate the work function for sodium. [3]

d Use the graph to calculate the value of the Planck constant in J s. [2]

12 **a** Explain what is meant by the **de Broglie wavelength** of an electron. [2]

b Figure 30.32 shows the principles of an electron tube used to demonstrate electron diffraction.

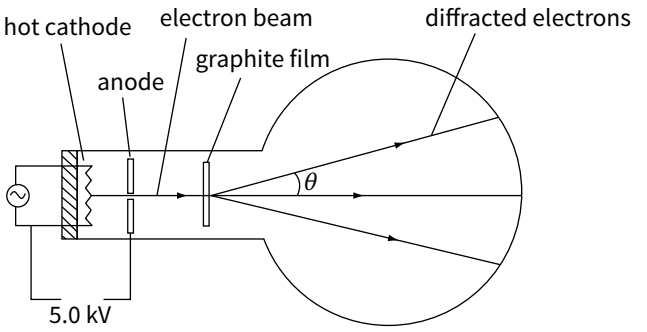


Figure 30.32 For End-of-chapter Question 12.

i Calculate the kinetic energy E (in joules) of the electrons incident on the graphite film. [1]

ii Show that the momentum of an electron is equal to $\sqrt{2Em_e}$ where m_e is the mass of an electron, and hence calculate the momentum of an electron. ($m_e = 9.11 \times 10^{-31}$ kg). [3]

iii Calculate the de Broglie wavelength of the electrons. [2]

c Discuss how the wavelengths of neutrons and electrons moving with the same energy would compare. [3]

13 Blue light of wavelength 450 nm is incident on a light-dependent resistor (LDR) made from a semiconductor material. The energy gap between the valence and conduction bands of the material of the LDR is 2.4 eV.

a Explain what is meant by the **valence band** and the **conduction band** in a semiconductor. You may include a diagram if you wish. [3]

b Show that the energy of a photon of the blue light is about 3 eV. [2]

c Use band theory to explain why:

i the blue light causes the resistance of the LDR to decrease [3]

ii the resistance of the LDR further decreases as the intensity of the blue light increases. [2]

d Calculate the maximum wavelength of radiation which this LDR could detect. [2]

14 Figure 30.33 shows the variation with temperature of the resistance of a semiconductor and of a metal.

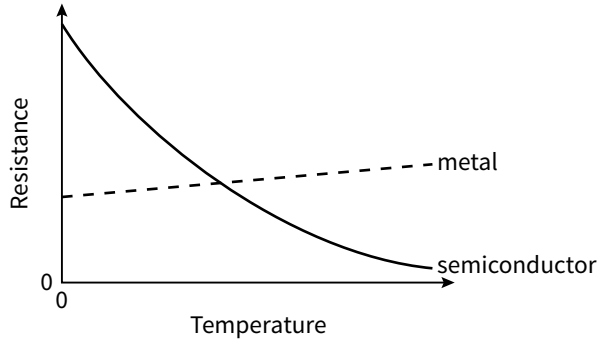
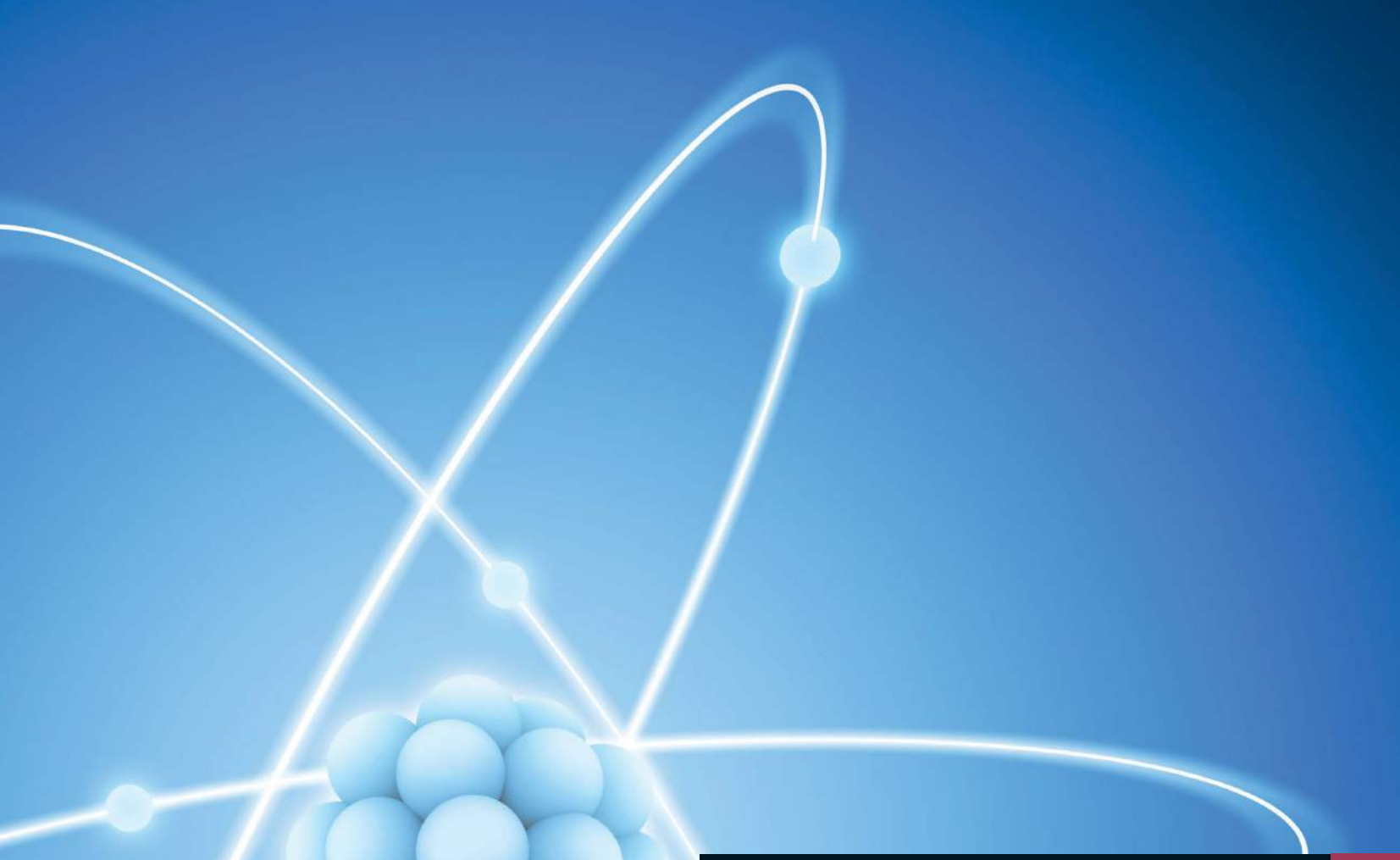


Figure 30.33 For End-of-chapter Question 14.

In terms of the band theory of solids and the behaviour of free electrons, explain the changes in resistance of:

a the metal [3]

b the semiconductor. [3]



Chapter 31: Nuclear physics

Learning outcomes

You should be able to:

- write simple nuclear decay equations and explain how they are balanced
- appreciate and use Einstein's mass–energy relationship $E = mc^2$
- relate nuclear binding energy to nuclear stability, fission and fusion
- describe the spontaneous and random nature of radioactive decay
- solve problems involving activity and decay constants
- solve problems involving exponential decay and half-life

Energy and the nucleus

Towards the end of the Second World War, nuclear weapons were dropped on the Japanese cities of Hiroshima and Nagasaki. Nuclear explosions release so much energy that their size is often given in 'megatonnes' – that is, their equivalent in millions of tonnes of high explosive (Figure 31.1). A more peaceful use of nuclear materials is in nuclear power generation. Because materials such as uranium are such concentrated stores of energy, a nuclear power station requires only a small van-load of fuel each week, whereas a coal-fired power station may require a train-load every hour.

In this chapter, we will look at energy in the atomic nucleus and how this relates to nuclear stability; we will also look at how we can write equations to represent radioactive decay.

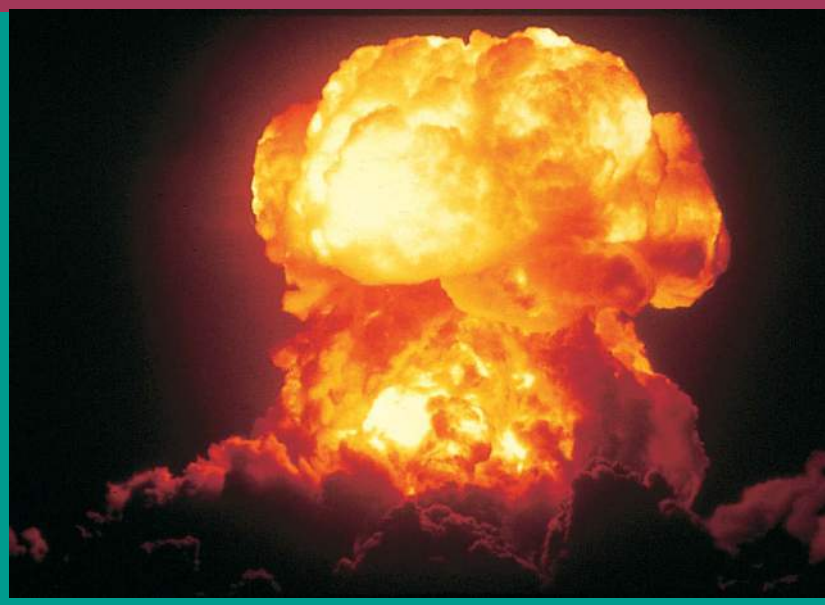


Figure 31.1 Our understanding of nuclear physics has proved to be a mixed blessing. Nuclear weapons dominated global politics for much of the 20th century.

Balanced equations

When an unstable nucleus undergoes radioactive decay, the nucleus before the decay is often referred to as the **parent** nucleus and the new nucleus after the decay of the α-particle is known as the **daughter** nucleus.

Radioactive decay processes can be represented by balanced equations. As with all equations representing nuclear processes, both nucleon number A and proton number Z are conserved.

- In α decay, the nucleon number decreases by 4 and the proton number decreases by 2.
- In β^- decay, the nucleon number is unchanged and the proton number increases by 1.
- In β^+ decay, the nucleon number is unchanged and the proton number decreases by 1.
- In γ emission there is no change in nucleon or proton number.

The emission of α- and β-particles can be shown on a graph of nucleon number plotted against proton number, as shown in Figure 31.2. The graph will appear different if neutron number is plotted against proton number.

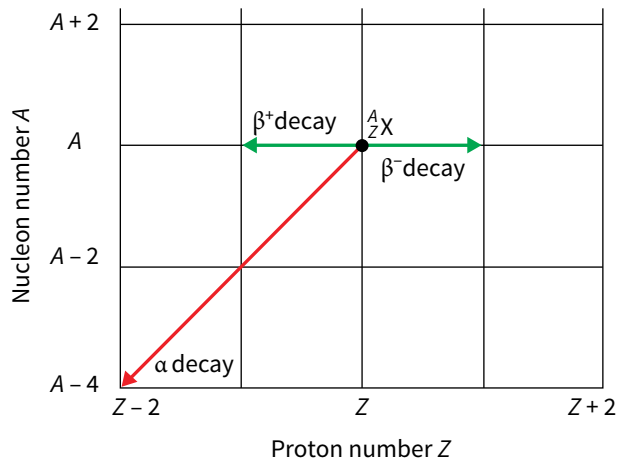


Figure 31.2 Emission of α- and β-particles.

QUESTIONS

1 Study the decay equations given in Worked examples 1 and 2, and write balanced equations for the following:

a A nucleus of radon-220 ($^{220}_{86}$Rn) decays by α emission to form an isotope of polonium, Po.

b A nucleus of a sodium isotope ($^{25}_{11}$Na) decays by β^- emission to form an isotope of magnesium, Mg.

2 Copy and complete this equation for the β^- decay of a nucleus of argon:

$^{41}_{18}\text{Ar} \rightarrow \text{K} + ?$

WORKED EXAMPLES

1 Radon is a radioactive gas that decays by α emission to become polonium. Here is the equation for the decay of one of its isotopes, radon-222:

$^{222}_{86}\text{Rn} \rightarrow {}^{218}_{84}\text{Po} + {}^{4}_{2}\text{He}$

Show that *A* and *Z* are conserved.

Compare the nucleon and proton numbers on both sides of the equation for the decay:

nucleon number *A*	222 = 218 + 4
proton number *Z*	86 = 84 + 2

Remember that in α-decay, A decreases by 4 and Z decreases by 2.
Don't confuse nucleon number A with neutron number N.

In this case, radon-222 is the parent nucleus and polonium-218 is the daughter nucleus.

2 A carbon-14 nucleus (parent) decays by β^--emission to become an isotope of nitrogen (daughter). Here is the equation that represents this decay:

$^{14}_{6}\text{C} \rightarrow {}^{14}_{7}\text{N} + {}^{0}_{-1}\text{e}$

Show that both nucleon number and proton number are conserved.

Compare the nucleon and proton numbers on both sides of the equation for the decay:

nucleon number *A*	14 = 14 + 0
proton number *Z*	6 = 7 − 1

Remember that in β^--decay, A remains the same and Z increases by 1.

Mass and energy

In Chapter 16, we saw that energy is released when the nucleus of an unstable atom decays. How can we calculate the amount of energy released by radioactive decay? To find the answer to this, we need to think first about the masses of the particles involved.

We will start by considering a stable nucleus, $^{12}_{6}\text{C}$. This consists of six protons and six neutrons. Fortunately for us (because we have a lot of this form of carbon in our bodies), this is a very stable nuclide. This means that the nucleons are bound tightly together by the strong nuclear force. It takes a lot of energy to pull them apart.

Figure 31.3 shows the results of an imaginary experiment in which we have done just that. On the left-hand side of the balance is a single $^{12}_{6}\text{C}$ nucleus. On the right-hand side are six protons and six neutrons, the result of dismantling the nucleus. The surprising thing is that the balance is tipped to the right. The separate nucleons have **more** mass than the nucleus itself. This means that the law of conservation of mass appears to have been broken. Have we violated what was thought to be a fundamental law of Nature, something that was held to be true for hundreds of years?

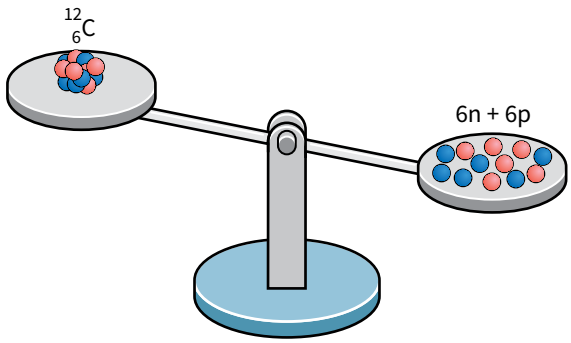


Figure 31.3 The mass of a nucleus is less than the total mass of its component protons and neutrons.

Notice that, in dismantling the $^{12}_{6}\text{C}$ nucleus, we have had to do work. The nucleons attract one another with nuclear forces and these are strong enough to make the nucleus very stable. So we have put energy into the nucleus to pull it apart, and this energy increases the potential energy of the individual nucleons. We can think of the nucleons within the nucleus as sitting in a deep potential well which results from the strong forces which hold the nucleus together. When we separate nucleons, we lift them out of this potential well, giving them more nuclear potential energy. This potential well is similar to that formed by the electric field around the nucleus; it is this well in which the atomic electrons sit, but it is much, much deeper. This explains why it is much easier to remove an electron from an atom than to remove a nucleon from the nucleus.

The problem of the appearing mass remains. To solve this problem, Einstein made the revolutionary hypothesis that energy has mass. This is not an easy idea. When bodies are in a higher energy state they have more mass than in a lower energy state. A bucket of water at the top of a hill will have more mass than when it is at the bottom because energy has been transferred to it in carrying it up the hill. A tennis ball travelling at 50 m s^{-1} will have more mass than the same tennis ball when stationary. In everyday life the amount of extra mass is so small that it cannot be measured, but the large changes in energy which occur in nuclear physics and high-energy physics

makes the changes in mass significant. Indeed, the increase in mass of particles, such as electrons, as they are accelerated to speeds near to the speed of light is a well-established experimental fact.

Another way to express this is to treat mass and energy as aspects of the same thing. Rather than having separate laws of conservation of mass and conservation of energy, we can combine these two. The total amount of mass and energy together in a system is constant. There may be conversions from one to the other, but the total amount of 'mass–energy' remains constant.

Einstein's mass–energy equation

Albert Einstein produced his famous mass–energy equation, which links energy E and mass m:

$$E = mc^2$$

where c is the speed of light in free space. The value of c is approximately $3.00 \times 10^8\ \text{m s}^{-1}$, but its precise value has been fixed as $c = 299\,792\,458\ \text{m s}^{-1}$.

Generally, we will be concerned with the changes in mass owing to changes in energy, when the equation becomes:

$$\Delta E = \Delta m c^2$$

According to Einstein's equation:

- the mass of a system **increases** when energy is supplied to it
- when energy is released from a system, its mass decreases.

Now, if we know the total mass of particles before a nuclear reaction and their total mass after the reaction, we can work out how much energy is released. Table 31.1 gives the mass in kilograms of each of the particles shown in Figure 31.3. Notice that this is described as the **rest mass** of the particle, that is, its mass when it is at rest (stationary); its mass is greater when it is moving because of its increase in energy. Nuclear masses are measured to a high degree of precision using mass spectrometers, often to seven or eight significant figures.

Particle	Rest mass / 10^{-27} kg
$^1_1$p	1.672 623
$^1_0$n	1.674 929
$^{12}_6$C nucleus	19.926 483

Table 31.1 Rest masses of some particles. It is worth noting that the mass of the neutron is slightly greater than that of the proton (roughly 0.1% greater).

We can use the mass values to calculate the mass that is released as energy when nucleons combine to form a nucleus. So for our particles in Figure 31.3, we have:

$$\text{mass before} = (6 \times 1.672\,623 + 6 \times 1.674\,929) \times 10^{-27}\ \text{kg}$$
$$= 20.085\,312 \times 10^{-27}\ \text{kg}$$

$$\text{mass after} = 19.926\,483 \times 10^{-27}\ \text{kg}$$

$$\text{mass difference } \Delta m = (20.085\,312 - 19.926\,483) \times 10^{-27}\ \text{kg}$$
$$= 0.158\,829 \times 10^{-27}\ \text{kg}$$

When six protons and six neutrons combine to form the nucleus of carbon-12, there is a very small loss of mass Δm, known as the **mass defect**.

> The mass defect of a nucleus is equal to the difference between the total mass of the individual, separate nucleons and the mass of the nucleus.

The loss in mass implies that energy is released in this process. The energy released E is given by Einstein's mass–energy equation. Therefore:

$$E = mc^2$$
$$= 0.158\,829 \times 10^{-27} \times (3.00 \times 10^8)^2$$
$$\approx 1.43 \times 10^{-11}\ \text{J}$$

This may seem like a very small amount of energy, but it is a lot on the scale of an atom. For comparison, the amount of energy released in a chemical reaction involving a single carbon atom would typically be of the order of 10^{-18} J, more than a million times smaller.

Now look at Worked example 3.

Mass–energy conservation

Einstein pointed out that his equation $\Delta E = \Delta m c^2$ applied to **all** energy changes, not just nuclear processes. So, for example, it applies to chemical changes, too. If we burn some carbon, we start off with carbon and oxygen. At the end, we have carbon dioxide and energy. If we measure the mass of the carbon dioxide, we find that it is very slightly less than the mass of the carbon and oxygen at the start of the experiment. The total potential energy of the system will be less than at the start of the experiment, hence the mass is less. In a chemical reaction such as this, the change in mass is very small, less than a microgram if we start with 1 kg of carbon and oxygen. Compare this with the change in mass that occurs during the fission of 1 kg of uranium, described later. The change in mass in a chemical reaction is a much, much smaller proportion of the original mass, which is why we don't notice it.

WORKED EXAMPLE

3 Use the data below to determine the minimum energy required to split a nucleus of oxygen-16 ($^{16}_{8}$O) into its separate nucleons. Give your answer in joules (J).

mass of proton = $1.672\,623 \times 10^{-27}$ kg

mass of neutron = $1.674\,929 \times 10^{-27}$ kg

mass of $^{16}_{8}$O nucleus = $26.551\,559 \times 10^{-27}$ kg

speed of light $c = 3.00 \times 10^{8}$ m s^{-1}

Step 1 Find the difference Δm in kg between the mass of the oxygen nucleus and the mass of the individual nucleons. The $^{16}_{8}$O nucleus has 8 protons and 8 neutrons.

Δm = final mass − initial mass

$\Delta m = [(8 \times 1.672\,623 + 8 \times 1.674\,929) - 26.551\,559] \times 10^{-27}$ kg

$\Delta m \approx 2.20 \times 10^{-28}$ kg

There is an increase in the mass of this system because external energy is supplied.

Step 2 Use Einstein's mass–energy equation to determine the energy supplied.

$\Delta E = \Delta m c^2$

$E = 2.20 \times 10^{-28} \times (3.00 \times 10^{8})^2 \approx 1.98 \times 10^{-11}$ J

QUESTIONS

3 The Sun releases vast amounts of energy. Its power output is 4.0×10^{26} W. Estimate by how much its mass decreases each second because of this energy loss.

4 **a** Calculate the energy released if a $^{4}_{2}$He nucleus is formed from separate protons and neutrons. The masses of the particles are given in Table **31.2**.

b Calculate also the energy released per nucleon.

Particle	Mass / 10^{-27} kg
$^{1}_{1}$p	1.672 623
$^{1}_{0}$n	1.674 929
$^{4}_{2}$He	6.644 661

Table 31.2 Masses of some particles.

5 A golf ball has a mass of 150 g. Calculate its increase in mass when it is travelling at 50 m s^{-1}. What is this as a percentage of its rest mass?

Another unit of mass

When calculating energy values using $E = mc^2$, it is essential to use values of mass in kg, the SI unit of mass. However, the mass of a nucleus is very small, perhaps 10^{-25} kg, and these numbers are awkward. As an alternative, atomic and nuclear masses are often given in a different unit, the **atomic mass unit** (symbol u).

1 u is defined as $\frac{1}{12}$ of the mass of a neutral atom of carbon-12.

It follows that the mass of an atom of $^{12}_{6}$C = 12.000 u exactly. (An alternative might have been to define the mass of a proton as 1.000 u exactly, but it is more practical to measure the mass of an atom of the carbon isotope $^{12}_{6}$C.)

Table 31.3 shows the masses of some nuclides in u.

Nuclide	Symbol	Mass / u
proton	$^{1}_{1}$p	1.007 825
neutron	$^{1}_{0}$n	1.008 665
helium-4	$^{4}_{2}$He	4.002 602
carbon-12	$^{12}_{6}$C	12.000 000
potassium-40	$^{40}_{19}$K	39.963 998
uranium-235	$^{235}_{92}$U	235.043 930

Table 31.3 Masses of some nuclides in atomic mass units. Some have been measured to several more decimal places than are shown here.

The atomic mass unit is related to the kilogram by:

$$1\text{ u} = 1.660\,538\,921(73) \times 10^{-27}\text{ kg}$$

This conversion factor is found by measuring the mass in kg of atoms of carbon-12. (The bracketed figures represent the experimental uncertainty.)

To convert a mass in u to kg, multiply by the conversion factor shown above (usually 1.6605×10^{-27} is sufficiently accurate).

From Table 31.3, you can see that the mass in u is close to the nucleon number A. For example, the mass of uranium-235 is slightly more than 235. The extra bit is known as the **mass excess**.

mass excess = mass (in u) − nucleon number

So the mass excess for U-235 is 235.043 930 − 235 = 0.043 930 u.

QUESTIONS

6 **a** The mass of an atom of $^{56}_{26}$Fe is 55.934 937 u. Calculate its mass in kg.

b The mass of an atom of $^{16}_{8}$O is $2.656\,015 \times 10^{-26}$ kg. Calculate its mass in u.

7 Table 31.3 gives the masses (in u) of several nuclides.

a Determine the mass excess of helium-4.

b One nuclide in the table has a negative mass excess. Name the nuclide and determine its mass excess.

Energy released in radioactive decay

Unstable nuclei may emit α- and β-particles with large amounts of kinetic energy. We can use Einstein's mass–energy equation $\Delta E = \Delta m c^2$ to explain the origin of this energy. Take, for example, the decay of a nucleus of uranium-238. It decays by emitting an α-particle and changes into an isotope of thorium:

$$^{238}_{92}\text{U} \rightarrow {}^{234}_{90}\text{Th} + {}^{4}_{2}\text{He}$$

The uranium nucleus is in a high-energy, relatively unstable state. It emits the α-particle and the remaining thorium nucleus is in a lower, more stable energy state. There is a decrease in the mass of the system. That is, the combined mass of the thorium nucleus and the α-particle is less than the mass of the uranium nucleus. According to Einstein's mass–energy equation, this difference in mass Δm is equivalent to the energy released as kinetic energy of the products. Using the most accurate values available:

mass of $^{238}_{92}$U nucleus $= 3.952\,83 \times 10^{-25}$ kg

total mass of $^{234}_{90}$Th nucleus and α-particle ($^{4}_{2}$He)
$= 3.952\,76 \times 10^{-25}$ kg

change in mass $\Delta m = (3.952\,76 - 3.952\,83) \times 10^{-25}$ kg

$\approx -7.0 \times 10^{-30}$ kg

The minus sign shows a decrease in mass, hence, according to the equation $\Delta E = \Delta m c^2$, energy is released in the decay process:

energy released $\approx 7.0 \times 10^{-30} \times (3.0 \times 10^{8})^2$

$\approx 6.3 \times 10^{-13}$ J

This is an enormous amount of energy for a single decay. One mole of uranium-238, which has 6.02×10^{23} nuclei, has the potential to emit total energy equal to about 10^{11} J.

We can calculate the energy released in all decay reactions, including β-decay, using the same ideas as above.

QUESTION

8 A nucleus of beryllium $^{10}_{4}$Be decays into an isotope of boron by β^--emission. The chemical symbol for boron is B.

a Write a nuclear decay equation for the nucleus of beryllium-10.

b Calculate the energy released in this decay and state its form.

mass of $^{10}_{4}$Be nucleus $= 1.662\,38 \times 10^{-26}$ kg

mass of boron isotope $= 1.662\,19 \times 10^{-26}$ kg

mass of electron $= 9.109\,56 \times 10^{-31}$ kg

Binding energy and stability

We can now begin to see why some nuclei are more stable than others. If a nucleus is formed from separate nucleons, energy is released. In order to pull the nucleus apart, energy must be put in; in other words, work must be done against the strong nuclear force which holds the nucleons together. The more energy involved in this, the more stable is the nucleus.

The minimum energy needed to pull a nucleus apart into its separate nucleons is known as the **binding energy** of the nucleus.

Take care: this is **not** energy stored in the nucleus; on the contrary, it is the energy that must be put in to the nucleus in order to pull it apart. In the example of $^{12}_{6}$C discussed above, we calculated the binding energy from the mass difference between the mass of the $^{12}_{6}$C nucleus and the masses of the separate protons and neutrons.

In order to compare the stability of different nuclides, we need to consider the binding energy per nucleon. We can determine the binding energy per nucleon for a nuclide as follows:

- Determine the mass defect for the nucleus.
- Use Einstein's mass–energy equation to determine the binding energy of the nucleus by multiplying the mass defect by c^2.
- Divide the binding energy of the nucleus by the number of nucleons to calculate the binding energy per nucleon.

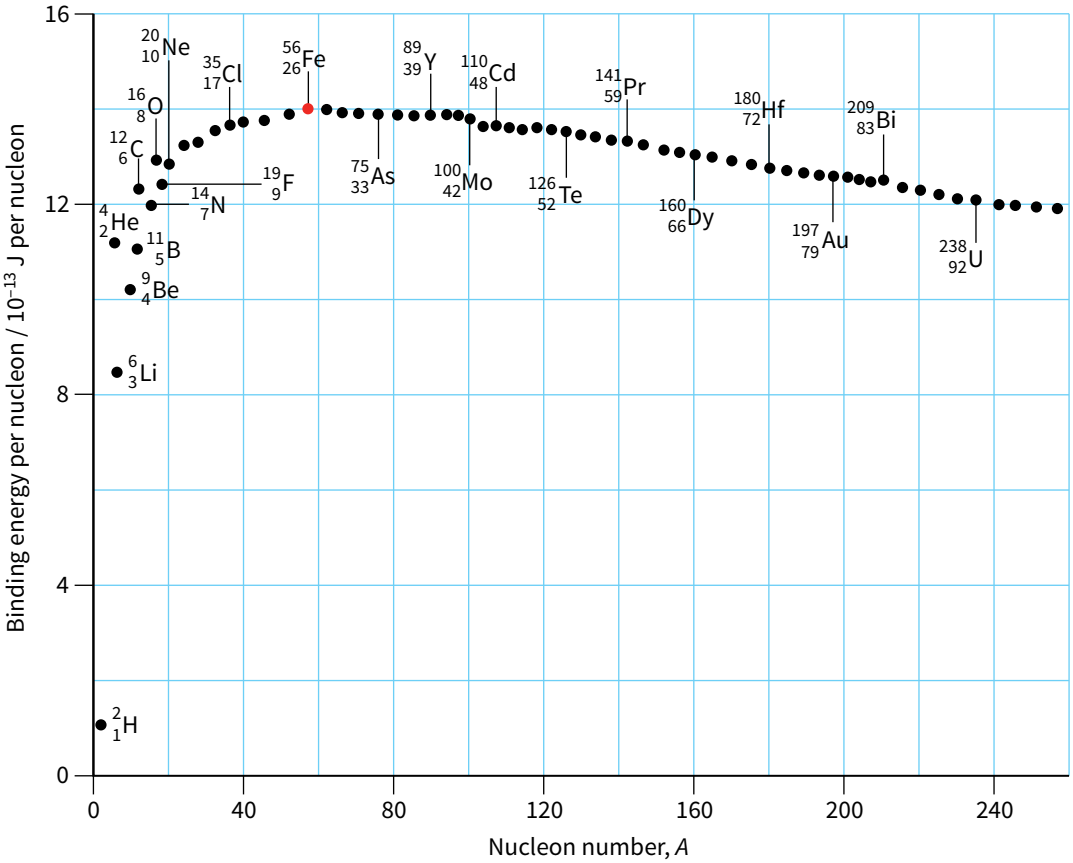


Figure 31.4 This graph shows the binding energy per nucleon for a number of nuclei. The nucleus becomes more stable as binding energy per nucleon increases.

WORKED EXAMPLE

4 Calculate the binding energy per nucleon for the nuclide $^{56}_{26}$Fe.

mass of neutron = 1.675×10^{-27} kg
mass of proton = 1.673×10^{-27} kg
mass of $^{56}_{26}$Fe nucleus = 9.288×10^{-26} kg

Step 1 Determine the mass defect.

number of neutrons = 56 − 26 = 30

mass defect
$= (30 \times 1.675 \times 10^{-27} + 26 \times 1.673 \times 10^{-27}) - 9.288 \times 10^{-26}$
$= 8.680 \times 10^{-28}$ kg

Step 2 Determine the binding energy of the nucleus.

binding energy = Δmc^2
$= 8.680 \times 10^{-28} \times (3.00 \times 10^{8})^2$
$= 7.812 \times 10^{-11}$ J

Step 3 Determine the binding energy per nucleon.

$$\text{binding energy per nucleon} = \frac{7.812 \times 10^{-11}}{56}$$
$$\approx 14 \times 10^{-13} \text{ J}$$

Figure 31.4 shows the binding energy per nucleon for stable nuclei, including the value for $^{56}_{26}$Fe (shown as a red dot) from Worked example 4. This is a graph plotted against the nucleon number A. The greater the value of the binding energy per nucleon, the more tightly bound are the nucleons that make up the nucleus.

If you examine this graph, you will see that the general trend is for light nuclei to have low binding energies per nucleon. Note, however, that helium has a much higher binding energy than its place in the Periodic Table might suggest. The high binding energy means that it is very stable. Other common stable nuclei include $^{12}_{6}$C and $^{16}_{8}$O, which can be thought of respectively as three and four α-particles bound together (Figure 31.5).

For nuclides with $A > 20$ approximately, there is not much variation in binding energy per nucleon. The greatest value of binding energy per nucleon is found

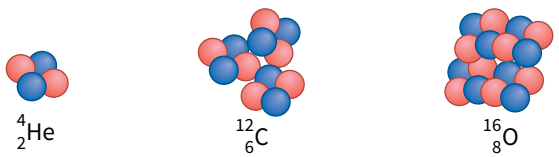


Figure 31.5 More stable nuclei are formed when 'α-particles' are bound together. In $^{12}_{6}$C and $^{16}_{8}$O, the 'α-particles' do not remain separate, as shown here; rather, the protons and neutrons are tightly packed together.

for $^{56}_{26}$Fe. This isotope of iron requires the most energy per nucleon to dismantle it into separate nucleons; hence iron-56 is the most stable isotope in nature.

QUESTIONS

9 Explain why hydrogen $^{1}_{1}$H does not appear on the graph shown in Figure 31.4.

10 The mass of a $^{8}_{4}$Be nucleus is 1.33×10^{-26} kg. A proton and a neutron have a mass of about 1.67×10^{-27} kg. For the nucleus of $^{8}_{4}$Be, determine:

a the mass defect in kg

b the binding energy

c the binding energy per nucleon.

Binding energy, fission and fusion

We can use the binding energy graph to help us decide which nuclear processes – fission, fusion, radioactive decay – are likely to occur (Figure 31.6).

Fission

Fission is the process in which a massive nucleus splits to form two smaller fragments (rather than simply emitting α- or β-radiation). For uranium, we have $A = 235$, and the typical fragments have $A = 140$ and 95. If we look at the binding energy curve, we see that these two fragments have greater binding energy (less potential energy) per nucleon than the original uranium nucleus. Hence, if the uranium nucleus splits in this way, energy will be released.

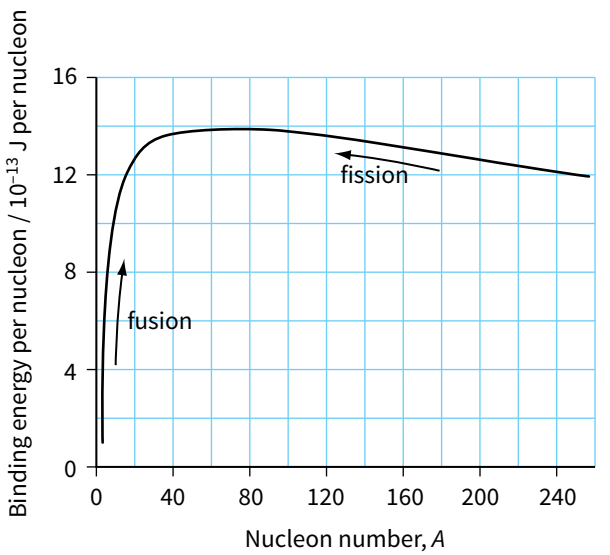


Figure 31.6 Both fusion and fission are processes that tend to increase the binding energy per nucleon of the particles involved.

Fusion

Fusion is the process by which two very light nuclei join together to form a heavier nucleus. (This is the process by which energy is released in the Sun, when hydrogen nuclei fuse to form helium nuclei.) If two light nuclei fuse, the final binding energy per nucleon will be greater than the original value. The high binding energy (low total energy) of the $^{4}_{2}$He nuclide means that it is rare for these nuclei to fuse.

QUESTION

11 Use the binding energy graph to suggest why fission is unlikely to occur with 'light nuclei' ($A < 20$), and why fusion is unlikely to occur for heavier nuclei ($A > 40$).

Randomness and decay

Listen to a counter connected to a Geiger–Müller (GM) tube that is detecting the radiation from a weak source, so that the count rate is about one count per second. Each count represents the detection of a single α-particle or a β-particle or a γ-ray photon. You will notice that the individual counts do not come regularly. The counter beeps or clicks in a random, irregular manner. If you try to predict when the next clicks will come, you are unlikely to be right.

You can see the same effect if you have a ratemeter, which can measure faster rates (Figure 31.7). The needle fluctuates up and down. Usually a ratemeter has a control for setting the 'time constant' – the time over which the meter averages out the fluctuations. Usually this can be set to 1 s or 5 s. The fluctuations are smoothed out more on the 5 s setting.



Figure 31.7 The time constant of this ratemeter can be adjusted to smooth out rapid fluctuations in the count rate.

Figure 31.8 shows a graph of count rate against time, with a smoothing of a few seconds. The count rate decreases with time as the number of radioactive nuclei that are left decreases. The fluctuations either side are caused by the randomness of the decay.

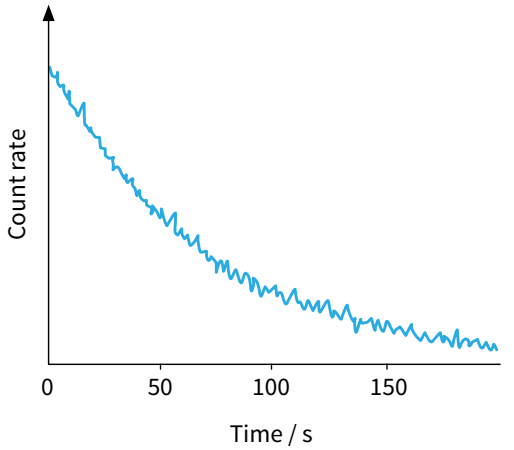


Figure 31.8 Count rate showing randomness of decay.

So it is apparent that radioactive decay is a random, irregular phenomenon. But is it completely unpredictable? Well, not really. We can measure the average rate of decay. We might measure the number of counts detected in 1000 s, and then calculate the average number per second. We cannot be sure about this average rate either, because the number of counts in 1000 s will fluctuate, too. All of our measurements of radioactive decay are inherently uncertain and imprecise but, by taking averages over a sufficiently long time period, we can reduce or smooth out the random fluctuations to reveal the underlying pattern.

Spontaneous decay

Radioactive decay occurs within the unstable nucleus of an atom. A nucleus emits radiation and becomes the nucleus of an atom of a different element. This is a spontaneous process, which means that we cannot predict, for a particular nucleus, when it will happen. If we sit and stare at an individual nucleus, we cannot see any change that will tell us that it is getting ready to decay. And if it doesn't decay in the first hour when we are watching it, we cannot say that it is any more likely to decay in the next hour. What is more, we cannot affect the probability of an individual nucleus decaying, for example by changing its temperature.

This is slightly odd, because it goes against our everyday experience of the way things around us change. We observe things changing. They gradually age, die, rot away. But this is not how things are on the scale of atoms and nuclei. Many of the atoms of which we are made have existed for billions of years, and will still exist long after we are gone. The nucleus of an atom does not age.

If we look at a very large number of atoms of a radioactive substance, we will see that the number of undecayed nuclei gradually decreases. However, we cannot predict when an **individual** nucleus will decay. Each nucleus 'makes up its own mind' when to decay, independently from its neighbours. This is because neighbouring nuclei do not interact with one another (unlike neighbouring atoms). The nucleus is a tiny fraction of the size of the atom, and the nuclear forces do not extend very far outside the nucleus. So one nucleus cannot affect a neighbouring nucleus by means of the nuclear force. Being inside a nucleus is a bit like living in a house in the middle of nowhere; you can just see out into the garden, but everything is darkness beyond, and the next house is 1000 km away.

The fact that individual nuclei decay spontaneously, and independently of their neighbours and of environmental factors, accounts for the random pattern of clicks that we hear from a Geiger counter and the fluctuations of the needle on the ratemeter.

To summarise, nuclear decay is **spontaneous** because:

- the decay of a particular nucleus is not affected by the presence of other nuclei
- the decay of nuclei cannot be affected by chemical reactions or external factors such as temperature and pressure.

Nuclear decay is **random** because:

- it is impossible to predict when a particular nucleus in a sample is going to decay
- each nucleus in a sample has the same chance of decaying per unit time.

The mathematics of radioactive decay

We have seen that radioactive decay is a random, spontaneous process. Because we cannot say when an individual nucleus will decay, we have to start thinking about very large numbers of nuclei. Even a tiny speck of radioactive material will contain more than 10^{15} nuclei. Then we can talk about the average number of nuclei that we expect to decay in a particular time interval; in other words, we can find out the **average** decay rate. Although we cannot make predictions for individual nuclei, we can say that certain types of nuclei are more likely to decay than others. For example, a nucleus of carbon-12 is stable; carbon-14 decays gradually over thousands of years; carbon-15 nuclei last, on average, a few seconds.

So, because of the spontaneous nature of radioactive decay, we have to make measurements on very large numbers of nuclei and then calculate averages. One quantity we can determine is the probability that an individual nucleus will decay in a particular time interval. For example, suppose we observe one million nuclei of a particular radioisotope. After one hour, 200 000 have decayed. Then the probability that an individual nucleus will decay in one hour is 0.2 or 20%, since 20% of the nuclei have decayed in this time. (Of course, this is only an approximate value, since we might repeat the experiment and find that only 199 000 decay because of the random nature of the decay. The more times we repeat the experiment, the more reliable our answer will be.)

We can now define the decay constant:

The probability that an individual nucleus will decay per unit time interval is called the **decay constant**, λ.

For the example above, we have:

decay constant $\lambda = 0.20\,\text{h}^{-1}$

Note that, because we are measuring the probability of decay per unit time interval, λ has units of h^{-1} (or s^{-1}, day^{-1}, year^{-1}, etc.).

The activity of a source is defined as follows:

The **activity** A of a radioactive sample is the rate at which nuclei decay or disintegrate.

Activity is measured in decays per second (or h^{-1}, day^{-1}). An activity of one decay per second is one becquerel (1 Bq):

$1\,\text{Bq} = 1\,\text{s}^{-1}$

Clearly, the activity of a sample depends on the decay constant λ of the isotope under consideration. The greater the decay constant (the probability that an individual nucleus decays per unit time interval), the greater is the activity of the sample. It also depends on the number of undecayed nuclei N present in the sample. For a sample of N undecayed nuclei, we have:

$$A = -\lambda N$$

We can also think of the activity as the number of α- or β-particles emitted from the source per unit time. Hence, we can also write the activity A as:

$$A = \frac{\Delta N}{\Delta t}$$

where ΔN is equal to the number of emissions (or decays) in a small time interval of Δt.

WORKED EXAMPLES

5 A radioactive source emits β-particles. It has an activity of 2.8×10^7 Bq. Estimate the number of β-particles emitted in a time interval of 2.0 minutes. State one assumption made.

Step 1 Write down the given quantities in SI units.

$A = 2.8 \times 10^7\,\text{Bq}$ $\quad \Delta t = 120\,\text{s}$

Step 2 Determine the number of β-particles emitted.

$$A = \frac{\Delta N}{\Delta t} \qquad \Delta N = A\,\Delta t$$

$\Delta N = 2.8 \times 10^7 \times 120 = 3.36 \times 10^9 \approx 3.4 \times 10^9$

We have assumed that the activity remains constant over a period of 2.0 minutes.

6 A sample consists of 1000 undecayed nuclei of a nuclide whose decay constant is $0.20\,\text{s}^{-1}$. Determine the initial activity of the sample. Estimate the activity of the sample after 1.0 s.

Step 1 Since activity $A = -\lambda N$, we have:

$A = 0.20 \times 1000 = 200\,\text{s}^{-1} = 200\,\text{Bq}$

Step 2 After 1.0 s, we might expect 800 nuclei to remain undecayed.

The activity of the sample would then be:

$A = 0.2 \times 800 = 160\,\text{s}^{-1} = 160\,\text{Bq}$

(In fact, it would be slightly higher than this. Since the rate of decay decreases with time all the time, less than 200 nuclei would decay during the first second.)

Count rate

Although we are often interested in finding the activity of a sample of radioactive material, we cannot usually measure this directly. This is because we cannot easily detect **all** of the radiation emitted. Some will escape past our detectors, and some may be absorbed within the sample itself. A GM tube placed in front of a radioactive source therefore only detects a fraction of the activity. The further it is from the source, the smaller the count rate. Therefore, our measurements give a received **count rate** R that is significantly lower than the activity A. If we know how efficient our detecting system is, we can deduce A from R. If the level of background radiation is significant, then it must be subtracted to give the **corrected** count rate.

QUESTIONS

12 A sample of carbon-15 initially contains 500 000 undecayed nuclei. The decay constant for this isotope of carbon is 0.30 s^{-1}. Determine the initial activity of the sample.

13 A small sample of radium gives a received count rate of 20 counts per minute in a detector. It is known that the counter detects only 10% of the decays from the sample. The sample contains 1.5×10^9 undecayed nuclei. Determine the decay constant of this form of radium.

14 A radioactive sample is known to emit α-, β- and γ-radiations. Suggest **four** reasons why the count rate measured by a Geiger counter placed next to this sample would be lower than the activity of the sample.

Decay graphs and equations

The activity of a radioactive substance gradually diminishes as time goes by. The atomic nuclei emit radiation and become different substances. The pattern of radioactive decay is an example of a very important pattern found in many different situations, a pattern called **exponential decay**. Figure 31.9 shows the decay graphs for three different isotopes, each with a different rate of decay.

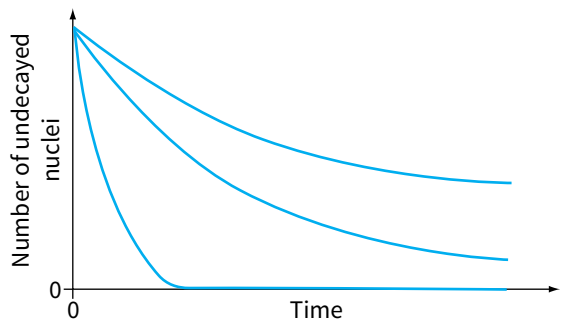


Figure 31.9 Some radioactive materials decay faster than others.

Although the three graphs look different, they all have something in common – their shape. They are curved lines having a special property. If you know what is meant by the **half-life** of a radioisotope, then you will understand what is special about the shape of these curves.

The half-life $t_{1/2}$ of a radioisotope is the mean time taken for half of the active nuclei in a sample to decay.

In a time equal to one half-life, the activity of the sample will also halve. This is because activity is proportional to the number of undecayed nuclei ($A \propto N$). It takes the same amount of time again for half of the remainder of the nuclei to decay, and a third half-life for half of the new remainder to decay (Figure 31.10).

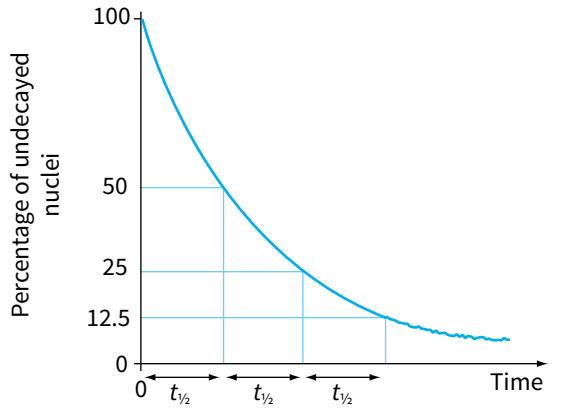


Figure 31.10 All radioactive decay graphs have the same characteristic shape.

In principle, the graph never reaches zero; it just gets closer and closer. In practice, when only a few undecayed nuclei remain the graph will cease to be a smooth curve (because of the random nature of the decay) and it will eventually reach zero. We use the idea of half-life because we cannot say when a sample will have completely decayed.

Mathematical decay

We can write an equation to represent the graph shown in Figure 31.10. If we start with N_0 undecayed nuclei, then the number N that remain undecayed after time t is given by:

$$N = N_0 e^{(-\lambda t)}$$

In this equation, λ is the decay constant, as before. (You may also see this written as $N = N_0 \exp(-\lambda t)$.) Note that you must take care with units. If λ is in s^{-1}, then the time t must be in s.

The symbol e represents the number e = 2.71828…, a special number in the same way that π is a special number. You will need to be able to use the e^x key on your calculator to solve problems involving e.

The activity A of a sample is proportional to the number of undecayed nuclei N. Hence the activity of the sample decreases exponentially:

$$A = A_0 e^{(-\lambda t)}$$

BOX 31.1: Determining half-life

If you are to determine the half-life of a radioactive substance in the laboratory, you need to choose something that will not decay too quickly or too slowly. In practice, the most suitable isotope is protactinium-234, which decays by emitting β^--radiation. This is available in a bottle containing a solution of a uranium compound (uranyl(VI) nitrate) (Figure 31.11). By shaking the bottle, you can separate the protactinium into the top layer of solvent in the bottle. The counter allows you to measure the decay of the protactinium.

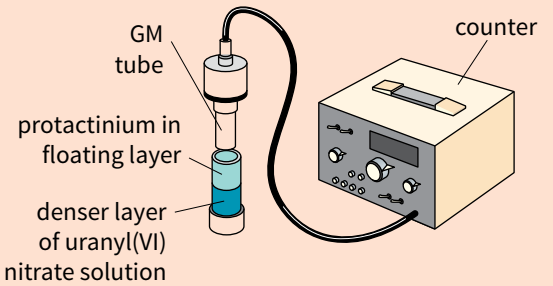


Figure 31.11 Practical arrangement for observing the decay of protactinium-234.

After recording the number of counts in consecutive 10-second intervals over a period of a few minutes, you can then draw a graph, and use it to find the half-life of protactinium-234.

Usually we measure the corrected count rate R in the laboratory rather than the activity or the number of undecayed nuclei. Since the count rate is a fraction of the activity, it too decreases exponentially with time:

$$R = R_0 e^{(-\lambda t)}$$

Now look at Worked examples 7 and 8.

WORKED EXAMPLES

7 Suppose we start an experiment with 1.0×10^{15} undecayed nuclei of an isotope for which λ is equal to $0.02\,s^{-1}$. Determine the number of undecayed nuclei after 20 s.

Step 1 In this case, we have $N_0 = 1.0 \times 10^{15}$, $\lambda = 0.02\,s^{-1}$ and $t = 20$ s. Substituting in the equation gives:

$N = 1.0 \times 10^{15} e^{(-0.02 \times 20)}$

Step 2 First calculate the expression in brackets; then use the e^x key and multiply by 1.0×10^{15}.

$N = 1.0 \times 10^{15} e^{(-0.40)}$

$= 6.7 \times 10^{14}$

8 A sample initially contains 1000 undecayed nuclei of an isotope whose decay constant $\lambda = 0.10\,min^{-1}$. Draw a graph to show how the sample will decay over a period of 10 minutes.

Step 1 We have $N_0 = 1000$ and $\lambda = 0.10\,min^{-1}$. Hence, we can write the equation for this decay:

$N = 1000 e^{(-0.10 \times t)}$

Step 2 Calculate values of the number N of undecayed nuclei at intervals of 1.0 min (60 s); this gives Table 31.4 and the graph shown in Figure 31.12.

t / min	0	1.0	2.0	3.0	4.0	5.0
N	1000	905	819	741	670	607
t / min	6.0	7.0	8.0	9.0	10.0	
N	549	497	449	407	368	

Table 31.4 For Worked example 8.

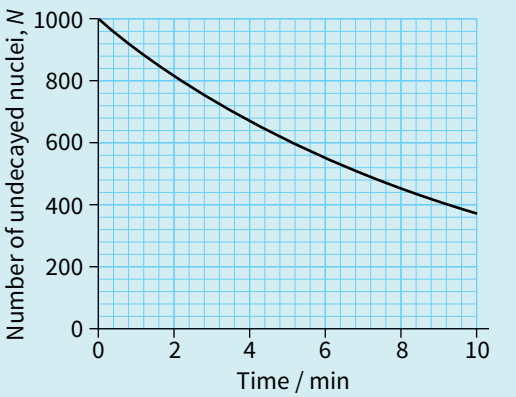


Figure 31.12 Radioactive decay graph.

QUESTIONS

15 The isotope nitrogen-13 has a half-life of 10 min. A sample initially contains 8.0×10^{10} undecayed nuclei.

a Write down an equation to show how the number undecayed, N, depends on time, t.

b Determine how many nuclei will remain after 10 min, and after 20 min.

c Determine how many nuclei will decay during the first 30 min.

16 A sample of an isotope for which $\lambda = 0.10\,s^{-1}$ contains 5.0×10^9 undecayed nuclei at the start of an experiment. Determine:

a the number of undecayed nuclei after 50 s

b its activity after 50 s.

17 The value of λ for protactinium-234 is $9.63 \times 10^{-3}\,s^{-1}$. Table 31.5 shows the number of undecayed nuclei, N, in a sample.

Copy and complete Table 31.5. Draw a graph of N against t, and use it to find the half-life $t_{1/2}$ of protactinium-234.

t/s	0	20	40	60	80	100	120	140
N	400	330						

Table 31.5 Data for Question 17.

Decay constant and half-life

A radioactive isotope that decays rapidly has a short half-life $t_{1/2}$. Its decay constant must be large, since the probability per unit time of an individual nucleus decaying must be high. What is the connection between the decay constant and the half-life?

In a time equal to one half-life $t_{1/2}$, the number of undecayed nuclei is halved. Hence the equation:

$$N = N_0 e^{(-\lambda t)}$$

becomes:

$$\frac{N}{N_0} = e^{(-\lambda t_{1/2})} = \tfrac{1}{2}$$

Therefore:

$$e^{(\lambda t_{1/2})} = 2$$

$$\lambda t_{1/2} = \ln 2 \approx 0.693$$

(remember if $e^x = y$, then $x = \ln y$).

The half-life of an isotope and the decay constant are inversely proportional to each other. That is:

$$\lambda = \frac{0.693}{t_{1/2}}$$

Thus if we know either $t_{1/2}$ or λ, we can calculate the other. For a nuclide with a very long half-life, we might not wish to sit around waiting to measure the half-life; it is easier to determine λ by measuring the activity (and using $A = \lambda N$), and use that to determine $t_{1/2}$.

Note that the units of λ and $t_{1/2}$ must be compatible; for example, λ in s^{-1} and $t_{1/2}$ in s.

QUESTIONS

18 Figure 31.13 shows the decay of a radioactive isotope of caesium, $^{134}_{55}Cs$. Use the graph to determine the half-life of this nuclide in years, and hence find the decay constant in $year^{-1}$.

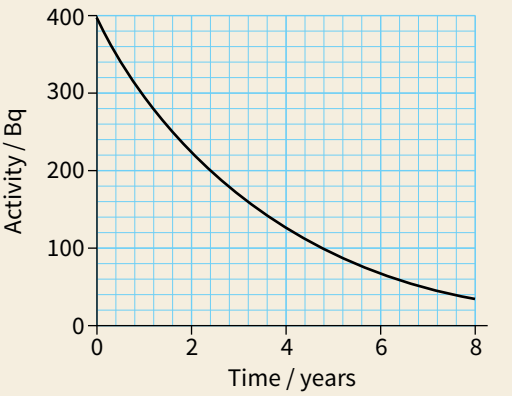


Figure 31.13 Decay graph for a radioactive isotope of caesium – see Question 18.

19 The decay constant of a particular isotope is known to be $3.0 \times 10^{-4}\,s^{-1}$. Determine how long it will take for the activity of a sample of this substance to decrease to one-eighth of its initial value.

20 The isotope $^{16}_{7}N$ decays with a half-life of 7.4 s.

a Calculate the decay constant for this nuclide.

b A sample of $^{16}_{7}N$ initially contains 5000 nuclei. Determine how many will remain after a time of:

i 14.8 s **ii** 20.0 s.

21 A sample contains an isotope of half-life $t_{1/2}$.

a Show that the fraction f of nuclei in the sample which remain undecayed after a time t is given by the equation:

$$f = (\tfrac{1}{2})^n \quad \text{when} \quad n = \frac{t}{t_{1/2}}$$

b Calculate the fraction f after each of the following times:

i $t_{1/2}$ **ii** $2t_{1/2}$

iii $2.5t_{1/2}$ **iv** $8.3t_{1/2}$

Summary

- Nuclear reactions can be represented by equations of the form

  $^{14}_{6}C \rightarrow {}^{14}_{7}N + {}^{0}_{-1}e$

- Einstein's mass–energy equation $\Delta E = \Delta mc^2$ relates mass changes to energy changes.
- The mass defect is equal to the difference between the mass of the separate nucleons and that of the nucleus.
- Atomic masses may be measured in atomic mass units:

  $1\,u = 1.660\,538\,921(73) \times 10^{-27}\,kg.$

- The mass excess of a nuclide is the difference between its mass (in u) and the nucleon number.
- The binding energy of a nucleus is the minimum energy required to break up the nucleus into separate nucleons.
- The binding energy per nucleon indicates the relative stability of different nuclides.
- The variation of binding energy per nucleon shows that energy is released when light nuclei undergo fusion and when heavier nuclei undergo fission, because these processes increase the binding energy per nucleon and hence result in more stable nuclides.
- Nuclear decay is a spontaneous and random process. This unpredictability means that count rates tend to fluctuate, and we have to measure average quantities.
- The half-life $t_{1/2}$ of a radioisotope is the mean time taken for half of the active nuclei in a sample to decay.
- The decay constant λ is the probability that an individual nucleus will decay per unit time interval.
- The activity A of a sample is related to the number of undecayed nuclei in the sample N by $A = -\lambda N$.
- The decay constant and half-life are related by the equation:

  $\lambda t_{1/2} = \ln 2 \quad \text{or} \quad \lambda t_{1/2} \approx 0.693$

- We can represent the exponential decrease of a quantity by an equation of the form:

  $x = x_0 e^{(-\lambda t)}$

  where x can be activity A, count rate R or number of undecayed nuclei N.

End-of-chapter questions

1 An antiproton is identical to a proton except that it has negative charge. If a proton and an antiproton collide they are annihilated and two photons are formed.

a Calculate the energy released in the reaction. [3]

b Calculate the energy released if 1 mole of protons and 1 mole of antiprotons were annihilated by this process. [3]

(Mass of a proton = mass of an antiproton = 1.67×10^{-27} kg.)

2 Calculate the mass that would be annihilated to release 1 J of energy. [2]

3 In a nuclear reactor the mass of uranium and the fission fragments falls at a rate of $70\,\mu g\,s^{-1}$. Calculate the maximum power output from the reactor assuming that it is 100% efficient. [3]

4 The equation shows the radioactive decay of radon-222.

$^{222}_{86}\text{Rn} \rightarrow {}^{218}_{84}\text{Po} + {}^{4}_{2}\alpha + \gamma$

Calculate the total energy output from this decay and state what forms of energy are produced. [6]

(Mass of $^{222}_{86}$Rn = 221.970 u, mass of $^{218}_{84}$Po = 217.963 u, mass of $^{4}_{2}\alpha$ = 4.002 u;
1 u is the unified atomic mass unit = 1.660×10^{-27} kg.)

(Hint: find the mass defect in u, then convert to kg.)

5 A carbon-12 atom consists of 6 protons, 6 neutrons and 6 electrons. The unified atomic mass unit (u) is defined as $\frac{1}{12}$ the mass of the carbon-12 atom.

Calculate:

a the mass defect in kilograms [2]

b the binding energy [2]

c the binding energy per nucleon. [2]

(Mass of a proton = 1.007 276 u, mass of a neutron = 1.008 665 u, mass of an electron = 0.000 548 u.)

6 The fusion reaction which holds most promise for the generation of electricity is the fusion of tritium $^{3}_{1}$H and deuterium $^{2}_{1}$H. The equation below shows the process:

$^{3}_{1}\text{H} + {}^{2}_{1}\text{H} \rightarrow {}^{4}_{2}\text{He} + {}^{1}_{1}\text{H}$

Calculate:

a the change in mass in the reaction [3]

b the energy released in the reaction [2]

c the energy released if one mole of deuterium were reacted with one mole of tritium. [2]

(Mass of $^{3}_{1}$H = 3.015 500 u, mass of $^{2}_{1}$H = 2.013 553 u, mass of $^{4}_{2}$He = 4.001 506 u, mass of $^{1}_{1}$H = 1.007 276 u.)

7 The initial activity a sample of 1 mole of radon-220 is $8.02 \times 10^{21}\,\text{s}^{-1}$. Calculate:

a the decay constant for this isotope [3]

b the half-life of the isotope. [2]

8 Figure 31.14 Shows the count rate recorded when a sample of the isotope indium-116 decays.

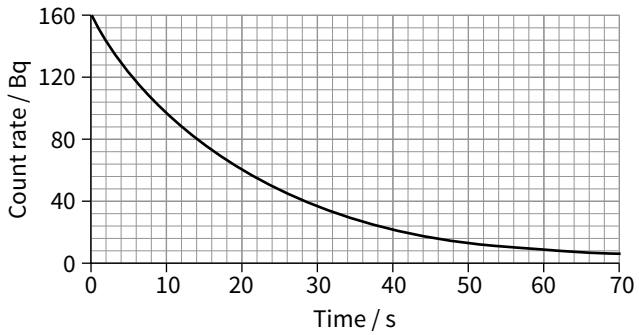


Figure 31.14 For End-of-chapter Question 8.

a Use the graph to find the half-life of the isotope. [2]

b Calculate the decay constant. [2]

9 The proportions of different isotopes in rocks can be used to date the rocks. The half-life of uranium-238 is 4.9×10^9 years. A sample has 99.2% of the proportion of this isotope compared with newly formed rock.

a Calculate the decay constant for this isotope of uranium. [2]

b Calculate the age of the rock. [3]

10 The table shows the received count rate when a sample of the isotope vanadium-52 decays.

Time / min	0	1	2	3	4	5	6	7	8
Count rate / Bq	187	159	134	110	85	70	60	56	40

a i Draw a graph of the count rate against the time. [5]

ii Comment on the scatter of the points. [1]

b From the graph, deduce the half-life of the isotope. [1]

c Describe the changes to the graph that you would expect if you were given a larger sample of the isotope. [2]

11 The graph in Figure 31.8 shows how randomness affects count rate. State and explain what happens if the experiment is performed using the same amount of radioactive material but at a higher temperature. [2]

12 This question is about the nucleus of uranium-235 ($^{235}_{92}$U), which has a mass of 3.89×10^{-25} kg.

a State the number of protons and neutrons in this nucleus. [1]

b The radius r of a nucleus is given by the equation:

$r = 1.41 \times 10^{-15} A^{1/3}$

where A is the nucleon number of the nucleus.

Determine the density of the $^{235}_{92}$U nucleus. [3]

c Explain why the total mass of the nucleons is different from the mass of the $^{235}_{92}$U nucleus. [2]

d Without calculations, explain how you can determine the binding energy per nucleon for the uranium-235 nucleus from its mass and the masses of a proton and a neutron. [4]

13 a Explain what is meant by **nuclear fusion** and explain why it only occurs at very high temperatures. [3]

b The main reactions which fuel the Sun are the fusion of hydrogen nuclides to form helium nuclides. However, other reactions do occur. In one such reaction, known as the triple alpha process, three helium nuclei collide and fuse to form a carbon-12 nucleus.

i Explain why temperatures higher than those required for the fusion of hydrogen are needed for the triple alpha process. [1]

ii Calculate the energy released in the triple alpha process. [3]

(Mass of a helium ($^4_2$He) nucleus = 4.001 506 u,
mass of a carbon ($^{12}_6$C) nucleus = 12.000 000 u,
1 u = 1.660×10^{-27} kg.)

14 The radioactive isotope of polonium, $^{218}_{81}$Po, decays by the emission of an α-particle with a half-life of 183 s.

a In an accident at a reprocessing plant some of this isotope, in the form of dust, is released into the atmosphere.
Explain why a spillage in the form of a dust is very much more dangerous to health than a liquid spillage. [2]

b It is calculated that 2.4 g of the isotope is released into the atmosphere.
Calculate the initial activity of the released polonium. [4]

c It is felt that it would safe to re-enter the laboratory when the activity falls to background, about 10 Bq.
Calculate how many hours must pass before it is safe to re-enter the laboratory. [3]

15 A nuclear reactor is fuelled by fission of uranium. The output from the reactor is 200 MW. The equation below describes a typical fission reaction.

$$^{239}_{52}\text{U} + ^{1}_{0}\text{n} \rightarrow ^{239}_{62}\text{U} \rightarrow ^{87}_{35}\text{Br} + ^{146}_{57}\text{La} + 3\,^{1}_{0}\text{n}$$

a Suggest and explain into what form the majority of the energy released in the reaction is converted. [2]

b **i** Calculate the energy released in the reaction. The kinetic energy of the captured neutron is negligible. [2]

ii Assume that the energy released in this fission is typical of all fissions of U-236. Calculate how many fissions occur each second. [1]

iii Calculate the mass of uranium-235 that is required to run the reactor for 1 year. [3]

(Mass of $^{239}_{52}$U = 3.90×10^{-25} kg, mass of $^{87}_{35}$Br = 1.44×10^{-25} kg, mass of $^{146}_{57}$La = 2.42×10^{-25} kg, mass of neutron = 1.67×10^{-27} kg, 1 year = 3.15×10^{7} s.)

16 The radioactive decay of nuclei is random and spontaneous. Explain what is meant by:

a **radioactive** decay [2]

b **random** decay [2]

c **spontaneous** decay. [2]

Chapter 32: Medical imaging

Learning outcomes

You should be able to:

- explain how X-ray beams are produced and controlled
- explain how ultrasound is produced and detected
- explain how ultrasound images are produced, revealing internal structures
- describe how conventional and CT scan X-ray images are produced
- explain the principles of magnetic resonance imaging

Applying physics

In this book, you have learned many important ideas from physics. You may have noticed that the same big ideas keep reappearing – for example, the idea of a field of force (magnetic, electric, gravitational), or the idea of energy transmitted as waves, or the idea that matter is made of particles with forces acting between them. This is an important characteristic of physics; ideas that are used in one area prove to be useful in another. Hopefully, you will see many of these connections now that you are approaching the end of your course.

Physics is also useful. It is applied in many areas of life. In this chapter, we look at one of these areas: medical imaging. This topic covers a range of techniques which doctors use to see inside our bodies. The best known is X-rays, good for showing up bones (Figure 32.1), and the subject of the first part of this chapter. The sections that follow will look at the physics behind two other medical diagnostic techniques: ultrasound scanning and magnetic resonance imaging. In this chapter, you will make use of several important aspects of physics that you have studied earlier in the course, including sound as a wave, electromagnetic radiation, the behaviour of charged particles, and magnetic fields.

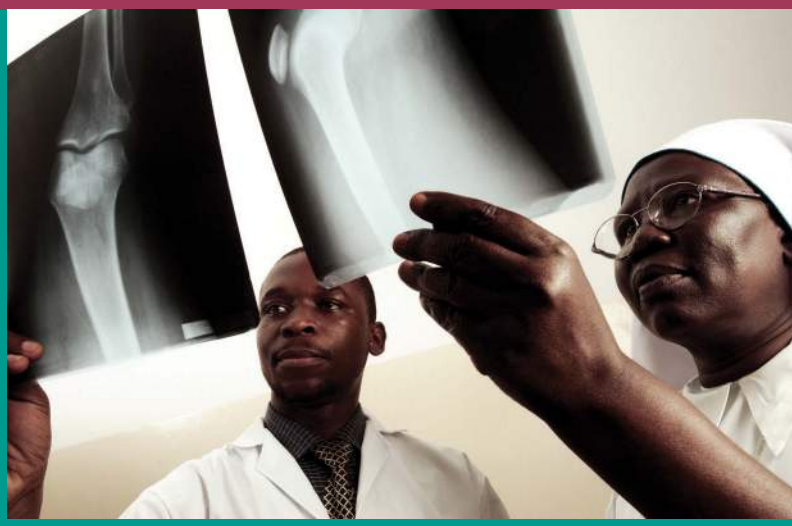


Figure 32.1 A radiographer and a doctor examine X-ray images of a patient's leg at a hospital in Uganda.

The nature and production of X-rays

X-rays are a form of electromagnetic radiation. They belong to the short-wavelength, high-frequency end of the electromagnetic spectrum, beyond ultraviolet radiation (Figure 32.2). They have wavelengths in the range 10^{-8} m to 10^{-13} m and are effectively the same as gamma-rays (γ-rays), the difference being in the way they are produced:

- X-rays are produced when fast-moving electrons are rapidly decelerated. As the electrons slow down, their kinetic energy is transformed to photons of electromagnetic radiation.
- γ-rays are produced by radioactive decay. Following alpha (α) or beta (β) emission, a gamma photon is often emitted by the decaying nucleus (see Chapter 16).

The X-rays used in medical applications are usually described as **soft X-rays**, because their energy is not very great, usually less than the energies of γ-rays produced by radioactive substances.

As with all electromagnetic radiation, we can think of X-rays either as waves or as photons (see Chapter 30). X-rays travel in straight lines through a uniform medium.

X-ray tube

Figure 32.3a shows a patient undergoing a pelvic X-ray to check for bone degeneration. The X-ray machine is above the patient; it contains the **X-ray tube** that produces the X-rays which pass downwards through the patient's body. Below the patient is the detection system. In this case an electronic detector is being used, but often photographic

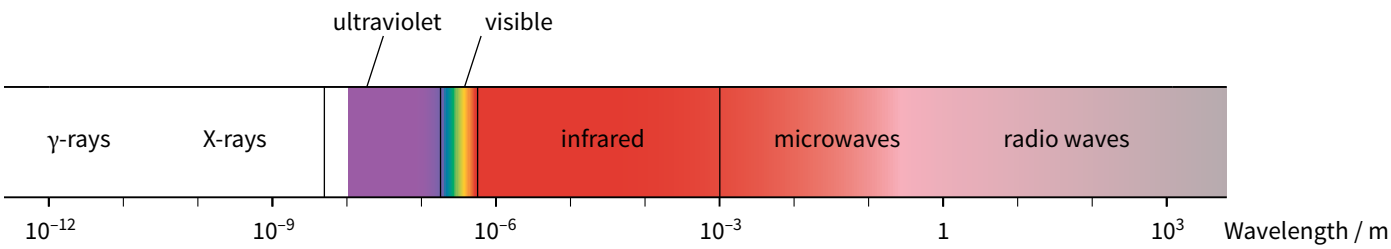


Figure 32.2 The electromagnetic spectrum; X-rays and γ-rays lie at the high-frequency, short-wavelength end of the spectrum.

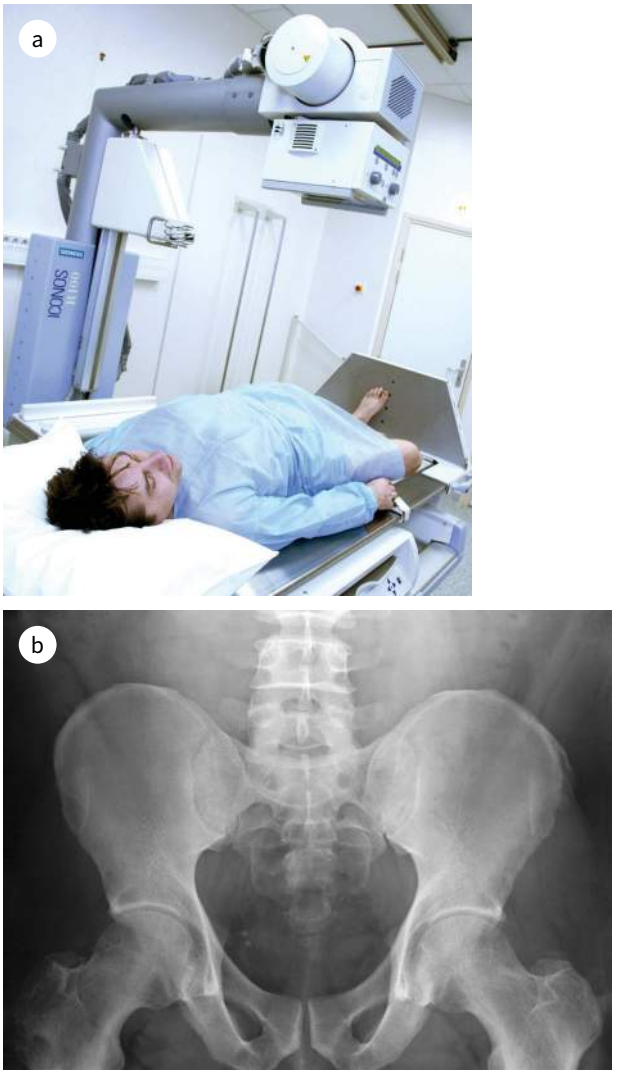


Figure 32.3 **a** A general-purpose X-ray system. **b** A typical X-ray image produced by such a machine, showing the region around the pelvis.

film is used in the detection system. Figure 32.3b shows the resulting image.

Figure 32.4 shows the principles of the modern X-ray tube. The tube itself is evacuated, and contains two electrodes:

- Cathode – the heated filament acts as the cathode (negative) from which electrons are emitted.
- Anode – the rotating anode (positive) is made of a hard metal such as tungsten. (The anode metal is often referred to as the 'target metal'.)

An external power supply produces a voltage of up to 200 kV between the two electrodes. This accelerates a beam of electrons across the gap between the cathode and the anode. The kinetic energy of an electron arriving at the anode is 200 keV. When the electrons strike the anode at high speed, they lose some of their kinetic energy in the

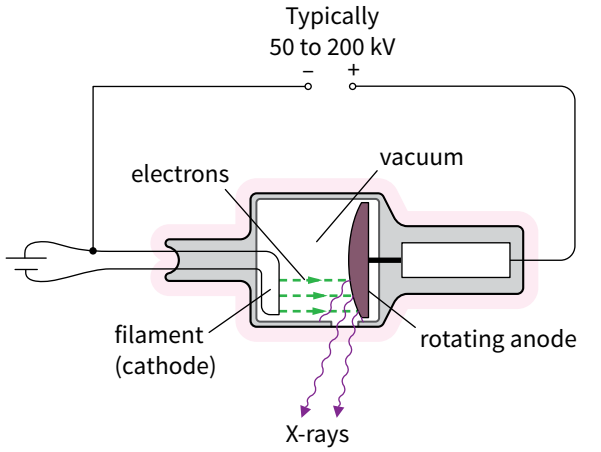


Figure 32.4 A simplified diagram of an X-ray tube.

form of X-ray photons, which emerge in all directions. Part of the outer casing, the window, is thinner than the rest and allows X-rays to emerge into the space outside the tube. The width of the X-ray beam can be controlled using metal tubes beyond the window to absorb X-rays. This produces a parallel-sided beam called a **collimated beam**.

Only a small fraction, about 1%, of the kinetic energy of the electrons is converted to X-rays. Most of the incident energy is transferred to the anode, which becomes hot. This explains why the anode rotates; the region that is heated turns out of the beam so that it can cool down by radiating heat to its surroundings. Some X-ray tubes have water circulating through the anode to remove this excess heat.

X-ray spectrum

The X-rays that emerge from an X-ray tube have a range of energies, as represented in the X-ray spectra shown in Figure 32.5. Each spectrum has two components, the

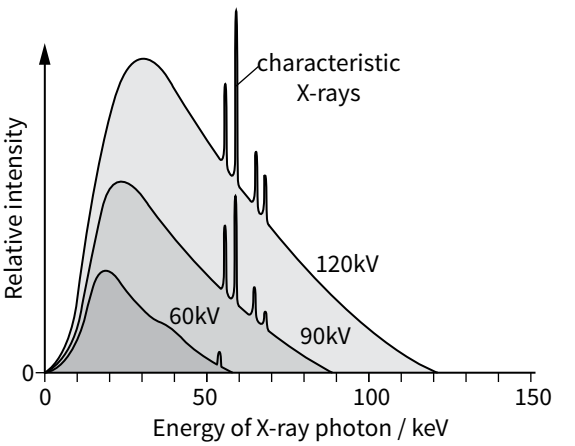


Figure 32.5 X-ray spectra for a tungsten target with accelerating voltages of 60 kV, 90 kV and 120 kV. The continuous curve shows the braking radiation while the sharp spikes are the characteristic X-rays.

broad background 'hump' of **braking radiation** (also known as Bremsstrahlung radiation) and a few sharp 'lines' of **characteristic radiation**. These arise from the different ways in which an individual electron loses its energy when it crashes into the anode.

When an electron strikes the anode, it will be attracted towards the nucleus of an atom in the anode. This will cause it to change both speed and direction – in other words, it decelerates. A decelerating electron (or any other charged particle) loses energy by emitting electromagnetic radiation. The result is a single X-ray photon or, more usually, several photons. The electron interacts with more nuclei until it has lost all its energy and it comes to a halt. The X-rays emitted in this way all contribute to the background braking radiation.

The energy E gained by the electron when it is accelerated through a potential difference of V between the cathode and the anode is given by $E = eV$. This is the maximum energy that an X-ray photon can have, and so the maximum X-ray frequency f_{max} that can be produced can be calculated using the formula $E = hf$. Hence:

$$f_{max} = \frac{eV}{h}$$

An electron may cause a rearrangement of the electrons in an anode atom, with an electron dropping from a high energy level to a lower energy level. As it does so, it emits a single photon whose energy is equal to the difference in energy levels. You should recall from Chapter 30 that this is how a line spectrum arises and the photon energies are characteristic of the atom involved. So the characteristic spectral lines of X-rays from a tungsten anode have different energies from those of a molybdenum or copper target. In practice, these characteristic X-rays are relatively unimportant in medical applications.

QUESTIONS

1 **a** Summarise the energy changes that take place in an X-ray tube.

b An X-ray tube is operated with a potential difference of 80 kV between the cathode and the tungsten anode. Calculate the kinetic energy (in electronvolts and joules) of an electron arriving at the anode. Estimate the impact speed of such an electron (assume that the electron is non-relativistic).

2 Determine the minimum wavelength of X-rays emitted from an X-ray tube operated at a voltage of 120 kV.

We can also see from Figure 32.5 that X-rays of a whole range of energies are produced. The lowest energy X-rays will not have sufficient energy to penetrate through the body, so will have no effect on the resulting image. However, they will contribute to the overall X-ray dose that the patient receives. These X-rays must be filtered out; this is done using aluminium absorbers across the window of the tube.

Controlling intensity and hardness

The intensity of an X-ray beam is a measure of the energy passing through unit area (see the next section). To increase the intensity of a beam, the current in the X-ray tube must be increased. Since each electron that collides with the anode produces X-rays, a greater current (more electrons per second) will produce a beam of greater intensity (more X-ray photons per second). A more intense beam means that the X-ray image will be formed in a shorter time.

Another consideration is the **hardness** of the X-rays. An X-ray may be thought of as 'hard' or 'soft'. Soft X-rays have lower energies and hence longer wavelengths than hard X-rays. Soft X-rays are less penetrating (they are more easily absorbed) and so they contribute more to the patient's exposure to hazardous radiation. It is often better to use hard X-rays, which pass through the body more easily.

The hardness of an X-ray beam can be increased by increasing the voltage across the X-ray tube, thereby producing X-rays of higher energies (see Figure 32.5). Another method is to use a filter which absorbs the lower energy soft X-rays so that the average energy of the X-rays is higher.

X-ray attenuation

As you can see if you look back to Figure 32.1, bones look white in an X-ray photograph. This is because they are good absorbers of X-rays, so that little radiation arrives at the photographic film to cause blackening. Flesh and other soft tissues are less absorbing, so the film is blackened. Modern X-ray systems use digital detectors instead of photographic films. The digital images are easier to process, store and transmit using computers.

X-rays are a form of ionising radiation; that is, they ionise the atoms and molecules of the materials they pass through. In the process, the X-rays transfer some or all of their energy to the material, and so a beam of X-rays is gradually absorbed as it passes through a material.

The gradual decrease in the intensity of a beam of X-rays as it passes through matter is called **attenuation**. We will now look at the pattern of attenuation of X-rays as they travel through matter.

Decreasing intensity

You should recall from Chapter 13 that the intensity of a beam of radiation indicates the rate at which energy is transferred across unit cross-sectional area. Intensity is defined thus:

> Intensity is the power per unit cross-sectional area.

We can determine the intensity I using the equation:

$$I = \frac{P}{A}$$

where P is power and A is the cross-sectional area normal to the radiation. The unit of intensity is $W\,m^{-2}$.

The intensity of a collimated beam of X-rays (i.e. a beam with parallel sides, so that it does not spread out) decreases as it passes through matter. Picture a beam entering a block of material. Suppose that, after it has passed through 1 cm of material, its intensity has decreased to half its original value. Then, after it has passed through 2 cm, the intensity will have decreased to one quarter of its original value (half of a half), and then after 3 cm it will be reduced to one eighth. You should recognise this pattern ($1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \ldots$) as a form of exponential decay.

We can write an equation to represent the attenuation of X-rays as they pass through a uniform material as follows:

$$I = I_0 e^{-\mu x}$$

where I_0 is the initial intensity (before absorption), x is the thickness of the material, I is the transmitted intensity and μ is the **attenuation** (or **absorption**) **coefficient** of the material. Figure 32.6 shows this pattern of absorption. It also shows that bone is a better absorber of X-rays than flesh; it has a higher attenuation coefficient. (The attenuation coefficient also depends on the energy of the X-ray photons.)

The unit of the attenuation coefficient μ is m^{-1} (or cm^{-1} etc.).

Now look at Worked example 1.

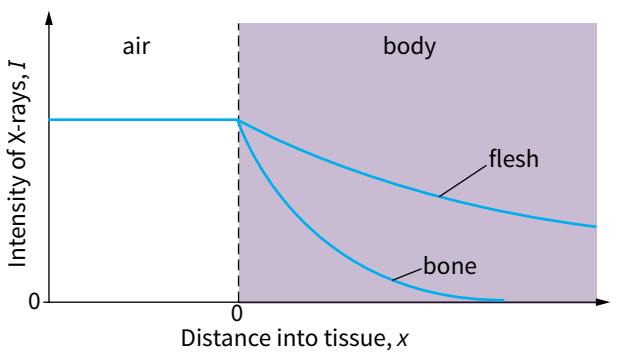


Figure 32.6 The absorption of X-rays follows an exponential pattern.

WORKED EXAMPLE

1 The attenuation (absorption) coefficient of bone is $600\,m^{-1}$ for X-rays of energy 20 keV. A beam of such X-rays has an intensity of $20\,W\,m^{-2}$. Calculate the intensity of the beam after passing through a 4.0 mm thickness of bone.

Step 1 Write down the quantities that you are given; make sure that the units are consistent.

$I_0 = 20\,W\,m^{-2}$

$x = 4.0\,mm = 0.004\,m$

$\mu = 600\,m^{-1}$

Step 2 Substitute in the equation for intensity and solve.

Hint: *Calculate the exponent (the value of $-\mu x$) first.*

$I = I_0 e^{-\mu x}$

$= 20 \times e^{-(600 \times 0.04)} = 20 \times e^{-2.4}$

$= 1.8\,W\,m^{-2}$

So the intensity of the X-ray beam will have been reduced to about 10% of its initial value after passing through just 4.0 mm of bone.

Half thickness

If we compare the graphs (or equations) for the attenuation of X-rays as they pass through a material with the decay of a radioactive nuclide we see that they are both exponential decays. From Chapter 31, you should become familiar with the concept of the half-life of a radioactive isotope (the time taken for half the nuclei in any sample of the isotope to decay). In a similar manner we refer to the half-thickness of an absorbing material. This is the thickness of material which will reduce the transmitted intensity of an X-ray beam of a particular frequency to half its original value.

QUESTIONS

3 Use the equation $I = I_0 e^{-\mu x}$ to show that the half-thickness $x_{1/2}$ is related to the attenuation coefficient μ by:

$$x_{1/2} = \frac{\ln 2}{\mu}$$

4 An X-ray beam transfers 400 J of energy through an area of $5.0\,cm^2$ each second. Calculate its intensity in $W\,m^{-2}$.

5 An X-ray beam of initial intensity $50\,W\,m^{-2}$ is incident on soft tissue of attenuation coefficient $1.2\,cm^{-1}$. Calculate its intensity after it has passed through a 5.0 cm thickness of tissue.

Improving X-ray images

The X-ray systems in use in hospitals and clinics today are highly developed pieces of technology. They do not simply show bones against a background of soft tissue. They can also show very fine detail in the soft tissue, including the arrangement of blood vessels.

Radiographers (the people in charge of X-ray systems) have three main aims:

- to reduce as much as possible the patient's exposure to harmful X-rays
- to improve the **sharpness** of the images, so that finer details can be resolved
- to improve the **contrast** of the image, so that the different tissues under investigation show up clearly in the image.

Reducing dosage

X-rays, like all ionising radiation, can damage living tissue, causing mutations which can lead to the growth of cancerous tissue. It is therefore important that the dosage is kept to a minimum.

A radiographer may choose to record the X-ray image on film or digitally. X-rays are only weakly absorbed by photographic film, so, historically, patients had to be exposed to long and intense doses of X-rays. Today, **intensifier screens** are used. These are sheets of a material that contains a phosphor, a substance that emits visible light when it absorbs X-ray photons. The film is sandwiched between two intensifier screens. Each X-ray photon absorbed results in several thousand light photons, which then blacken the film. This reduces the patient's exposure by a factor of 100–500.

In digital systems, **image intensifiers** are used (Figure 32.7). The incoming X-rays strike a phosphor screen, producing visible light photons. These then release electrons (by the photoelectric effect) from the photocathode. The electrons are accelerated and focused by the positively charged anode so that they strike a screen, which then gives out visible light. The image on this screen can be viewed via a television camera. At the same time, the image can be stored electronically. Digital systems have the advantage that images can be easily stored, shared and viewed.

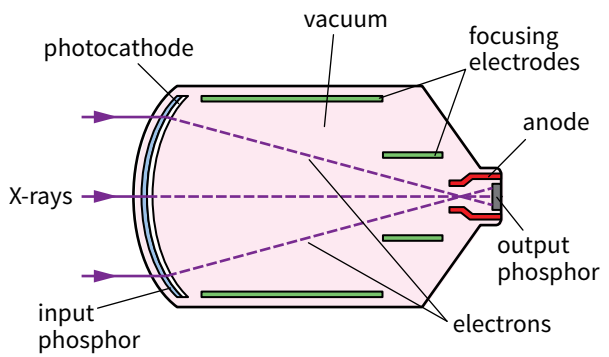


Figure 32.7 An X-ray image intensifier.

Image intensifiers are particularly useful in a technique called **fluoroscopy**. A continuous X-ray beam is passed through the patient onto a fluorescent screen where a real-time image is formed. Using an image intensifier ensures that the patient is not exposed to dangerous levels of X-rays over a long period.

Improving sharpness

Figure 32.8 shows a remarkably sharp X-ray image of blood vessels in the human abdomen. The sharpness of the image is determined by the width of the X-ray beam. You will remember that the shadow of an object is much sharper if it is illuminated by a small lamp, rather than a large lamp (Figure 32.9). So a good X-ray source must produce a **narrow** beam of **parallel** X-rays, as if they were coming from a distant point source.

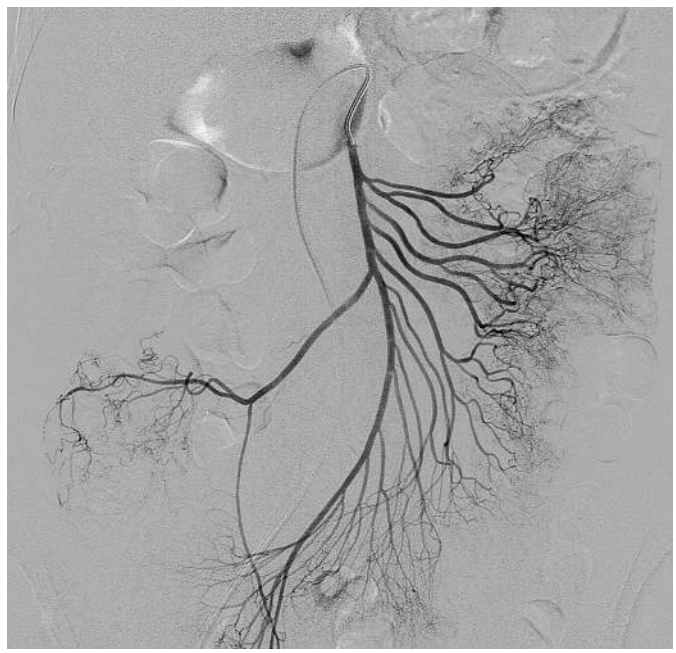


Figure 32.8 An X-ray image of blood vessels branching out from an artery carrying oxygenated blood to the intestines.

Three factors determine the width of the X-ray beam:

- the width of the electron beam and the target it strikes – as shown in Figure 32.10, the wider the electron beam, the wider the X-ray beam
- the size of the aperture at the exit window – this can be reduced using adjustable lead plates (Figure 32.11)
- collimation of the beam – the beam is passed through lead slits (Figure 32.12), ensuring that it is an approximately parallel-sided beam and does not fan out.

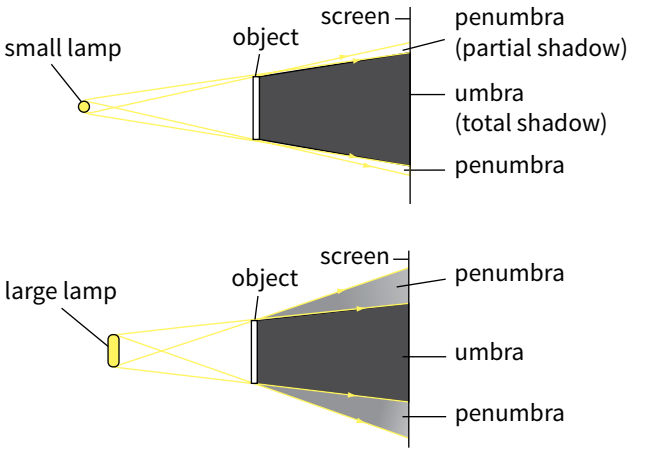


Figure 32.9 The small lamp casts a smaller penumbra and this improves the sharpness of the shadow.

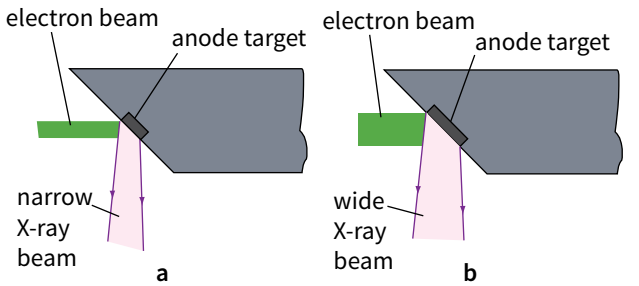


Figure 32.10 A wide anode target results in a wide X-ray beam, giving fuzzy edges to the shadow image.

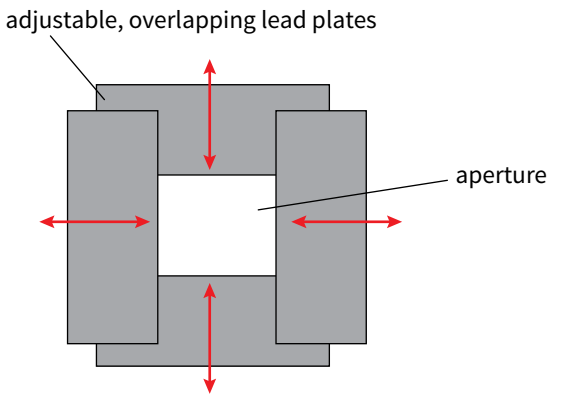


Figure 32.11 The smaller the aperture, the narrower the X-ray beam.

Inevitably some X-rays are scattered as they pass through the body. If these reach the detector they cause fogging and this reduces the sharpness of the image. Scattered X-rays approach the detector screen at an angle, and so an anti-scatter screen (Figure 32.13) can be used to absorb them. This consists of a series of plates made of a material (such as lead) which is opaque to X-rays, separated by plates made of a material (such as aluminium) which is transparent to X-rays. The plate is placed just above the screen, and the lead absorbs the scattered X-rays.

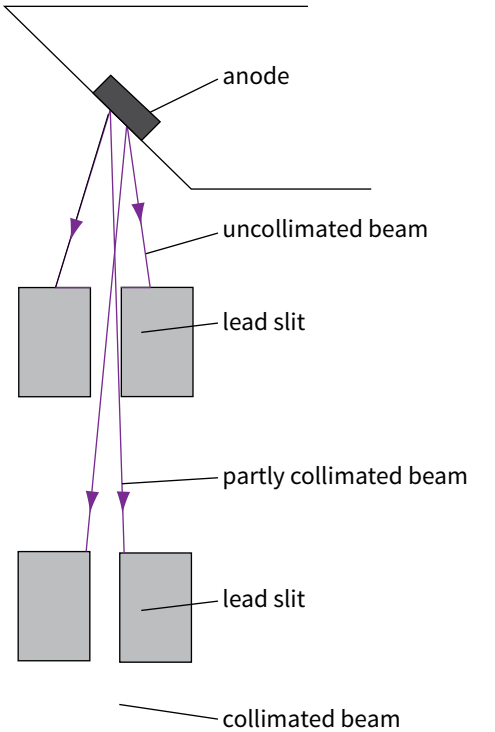


Figure 32.12 Collimating an X-ray beam. The first set of slits produces a partly collimated beam but, due to the finite size of the anode target, there is still some spreading of the beam. The second set of slits reduces this spread further, making the final beam almost parallel-sided.

Improving contrast

Good contrast is said to be achieved if there is a clear difference in the blackening of the photographic film as the X-ray passes through different types of tissue. The contrast is largely determined by the hardness of the X-rays. Bone is a good absorber of the radiation. If the doctor is diagnosing a break in a bone, he or she will use

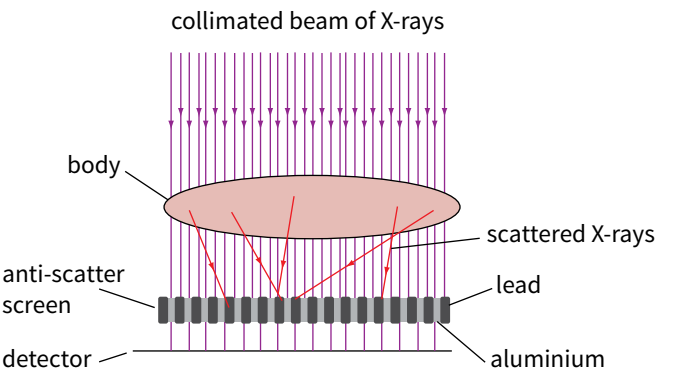


Figure 32.13 An anti-scatter screen absorbs X-rays which arrive at an angle to the main beam.

hard X-rays. In contrast, investigation of the tissue of the breast, where the tissue is a poor absorber, will require a longer exposure, using much softer (long-wavelength, low-frequency) X-rays.

As we have seen, different tissues show up differently in X-ray images. In particular, bone can readily be distinguished from soft tissue such as muscle because it is a good absorber of X-rays. However, it is often desirable to show up different soft tissues that absorb X-rays equally. In order to do this, **contrast media** are used.

A contrast medium is a substance such as iodine or barium which is a good absorber of X-rays. The patient may swallow a barium-containing liquid (a 'barium meal'), or have a similar liquid injected into the tissue of interest. This tissue is then a better absorber of X-rays and its edges show up more clearly on the final image.

Figure 32.14 shows an X-ray image of the intestine of a patient who has been given a barium meal. The large pale areas show where the barium has accumulated. Other parts of the intestine have become smeared with barium, and this means that the outline of the tissue shows up clearly.



Figure 32.14 X-ray image of a patient's intestine after taking a barium meal. Barium shows up as pale in this image, which has also been artificially coloured to highlight features of interest.

Contrast media are elements with high values of atomic number Z. This means that their atoms have many electrons with which the X-rays interact, so they are more absorbing. Soft tissues mostly consist of compounds of hydrogen, carbon and oxygen (low Z values), while bone has the heavier elements calcium and phosphorus, and contrast media have even higher Z values – see Table 32.1.

Substance	Elements (Z values)	Average Z
soft tissue	H (1), C (6), O (8)	7
bone	H (1), C (6), O (8), P (15), Ca (20)	14
contrast media	I (53), Ba (56)	55

Table 32.1 Proton (atomic) numbers of the constituents of different tissues, and of contrast media.

QUESTIONS

6 Outline the advantages of using an anti-scatter grid when taking an X-ray image.

7 The data in Table 32.2 shows how the attenuation coefficient μ depends on the energy of the X-rays in bone and muscle. When making a diagnostic X-ray image, it is desirable that bone should be clearly distinguished from muscle. Use the data in Table 32.2 to explain why it would be best to use lower energy (50 keV) X-rays for this purpose.

Maximum X-ray energy	Bone: μ / cm^{-1}	Muscle: μ / cm^{-1}
4.0 MeV	0.087	0.049
250 keV	0.32	0.16
100 keV	0.60	0.21
50 keV	3.32	0.54

Table 32.2 Data for Questions 7 and 8.

8 When low-energy X-rays are used, the attenuation coefficient μ is (roughly) proportional to the cube of the proton number Z of the absorbing material. Use the data in Table 32.2 to show that bone absorbs X-rays eight times as strongly as muscle.

Computerised axial tomography

A conventional X-ray image has an important limitation. Because an X-ray is essentially a two-dimensional shadow image, it shows the bones, organs, etc. at different depths within the body superimposed on each other. For example,

in Figure 32.15, it is difficult to distinguish the bones of the front and back of the ribcage. This can be overcome by taking several images at different angles. An experienced radiographer can then study these images and deduce what is going on inside the patient.



Figure 32.15 Computer-generated X-ray image of a person in a yoga position. This shows the difficulty of distinguishing one bone from another when they overlap.

An ingenious technique for extending this approach was invented by Geoffrey Hounsfield and his colleagues at EMI in the UK in 1971. They developed the **computerised axial tomography** scanner (CAT scanner or CT scanner). Figure 32.16 illustrates the principle of a modern scanner.

- The patient lies in a vertical ring of X-ray detectors.
- The X-ray tube rotates around the ring, exposing the patient to a fan-shaped beam of X-rays from all directions.
- Detectors opposite the tube send electronic records to a computer.
- The computer software builds up a three-dimensional image of the patient.
- The radiographer can view images of 'slices' through the patient on the computer screen.

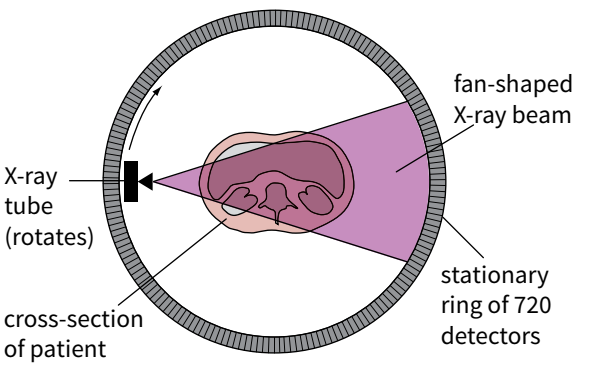


Figure 32.16 Operation of a modern CT scanner. The X-ray tube rotates around the patient while the detectors are stationary.

CT scanners have undergone many developments since they were first invented. In a fifth-generation scanner, the patient's bed slides slowly through the ring of detectors as the X-ray tube rotates. The tube thus traces out a spiral path around the patient, allowing information to be gathered about the whole body.

Figure 32.17 shows a child undergoing a CT scan. On the monitor you can see a cross-section of the patient's head.



Figure 32.17 A boy undergoes a CT scan in an investigation of an eye condition.

This technique is called **computerised axial tomography** because it relies on a computer to control the scanning motion and to gather and manipulate the data to produce images; because the X-ray tube rotates around an axis; and because it produces images of slices through the patient – the Greek word **tomos** means **slice**.

Building up the image

As the X-ray tube is rotated around the body, hundreds of pieces of information are gathered and an image is built up. As shown in Figure 32.18, we imagine the body as being divided up into a large number of tiny cubes called **voxels**. (This is the same as dividing a two-dimensional picture into a 2-D array of pixels, but a three-dimensional body must be divided into a 3-D array of voxels.)

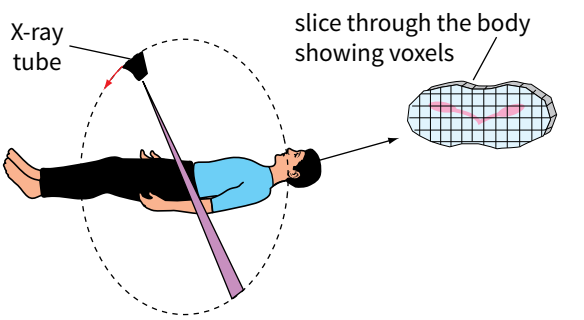


Figure 32.18 In CT scanning, we picture the body divided into an array of tiny cubic volumes called voxels.

To understand how the image is constructed from the data, we will simplify the procedure by considering a section made up of four voxels and imagine exposing this 2 × 2 grid to a beam of X-rays from four different directions. Different parts of the body have different 'densities'; that is, some are stronger absorbers of X-rays. We will represent this by labelling our four voxels with densities 5, 6, 2 and 8, as shown in Figure 32.19. This diagram shows how the detectors read different values when the array is exposed to X-rays from different angles, and how we can then work back to the original densities.

Step 1

The beam passes through the array from the side. The top part of the beam has passed through voxels with values 5 and 6, which makes 11. The value for the bottom part is 8 + 2 = 10. This is recorded in a 2 × 2 memory grid as 11, 11 in the top row and 10, 10 below.

Step 2

The beam is rotated through 45° so that it passes through the array diagonally. The top of the beam passes though just one voxel, of value 5. The central part passes through two voxels giving 6 + 2 = 8. The lower part passes through one voxel of value 8. This gives a detected grid as shown, and this is added to the memory grid.

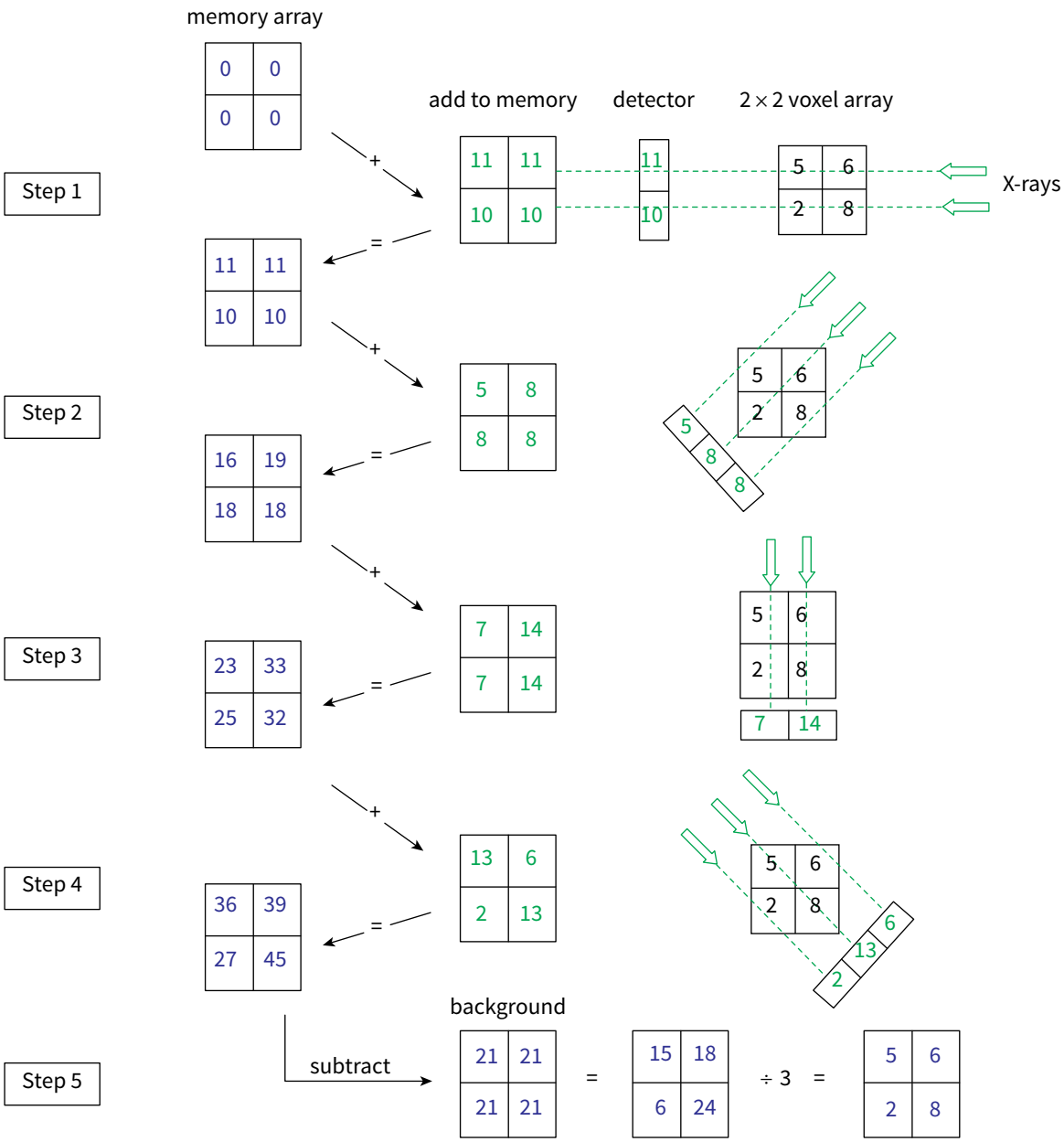


Figure 32.19 Data is built up from a CT scan of a 2 × 2 voxel array, and then processed to deduce the original array.

Steps 3 and 4

The beam is rotated twice more through 45° and each time the detected values are added to the memory grid.

Step 5

Now each voxel in the array has been exposed to X-rays from four different directions. How can we extract the original values from the final memory grid? Note that, in each step, the total density detected had a value of 21 (10 + 11 in Step 1, 5 + 8 + 8 in Step 2, and so on). We subtract this background value from each square in the memory grid, and then divide the remainder by three. The final values in the memory grid are the same as in the original 2 × 2 array.

Our 2 × 2 array is an example in two dimensions. In three dimensions, we would need to consider a 2 × 2 × 2 array of cubes rather than squares. This is known as an 8-voxel cube. A real object would be made up of a very large number of tiny voxels in three dimensions, requiring very powerful computers to analyse all the data collected.

For a well-defined image in a CT scan, we need the voxels to be small. Two things are needed to achieve this:

- The X-ray beam must be well collimated so that it consists of parallel rays – the rays must not spread outwards.
- The detector must consist of a regular array of tiny detecting elements – the smaller each individual detector is, the better will be the resolution in the final image.

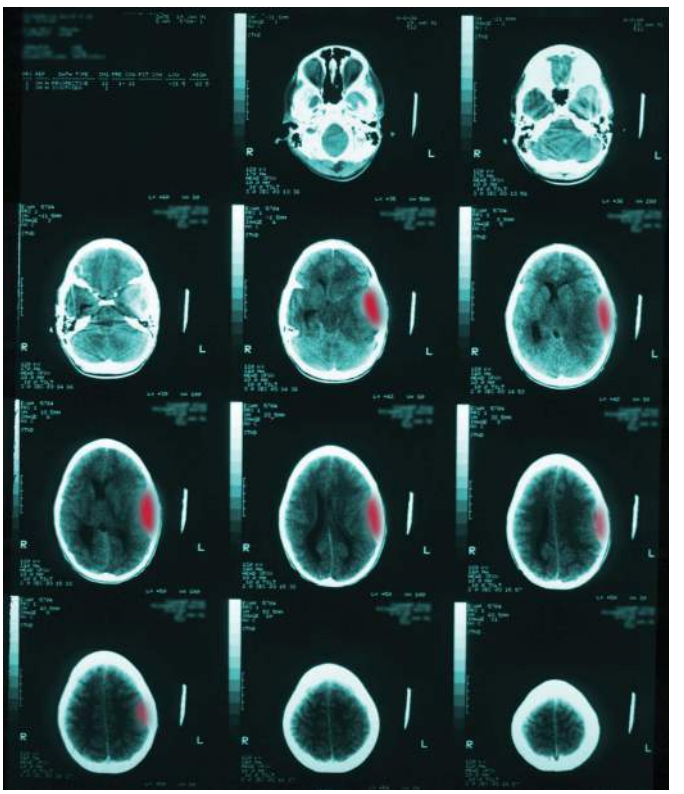


Figure 32.20 Sections through the head of a 10-year-old boy. You can see the haematoma (bruising) arising from being struck on the side of the head; this causes pressure on his brain.

In practice, the body is exposed to X-rays from many directions, giving a large number of values from which a complex computer program can deduce the variation of X-ray absorption throughout the body. From this, an image of any section through the body can be constructed on a monitor screen. Figure 32.20 shows a sequence of sections through a child's head.

Advantages of a CT scan

Although single X-ray images still have many uses (and they can be made very quickly), CT scans have a number of advantages:

- They produce images that show three-dimensional relationships between different tissues.
- They can distinguish tissues with quite similar densities (attenuation coefficients).

So, for example, a CT scan can show up the precise position, shape and size of a tumour. This allows it to be precisely targeted in treatment with high-energy X-rays or γ-rays.

However, it is worth noting that a CT scan involves using X-rays and any exposure to ionising radiation carries a risk for the patient. These risks are fairly small; it is estimated, with modern scanning equipment, that the radiation dose received is about one-third the dose received from background radiation in a year, or is equivalent to the dose received on four long-haul flights. Nevertheless, it is important to be aware of the dangers, particularly if there are other underlying health problems or if a woman is pregnant.

QUESTIONS

9 Suggest why a patient may be asked to hold his or her breath during a CT scan.

10 A patient with an injury to the skull, perhaps as a result of a road accident, is likely to undergo a CT scan. Explain why a CT scan is preferable to a conventional X-ray in a case like this.

Using ultrasound in medicine

Ultrasound scanning is routinely used to check the condition of a baby in the womb (Figure 32.21). There do not seem to be any harmful side-effects associated with this procedure, and it can provide useful information on the baby's development. Indeed, for many children, their

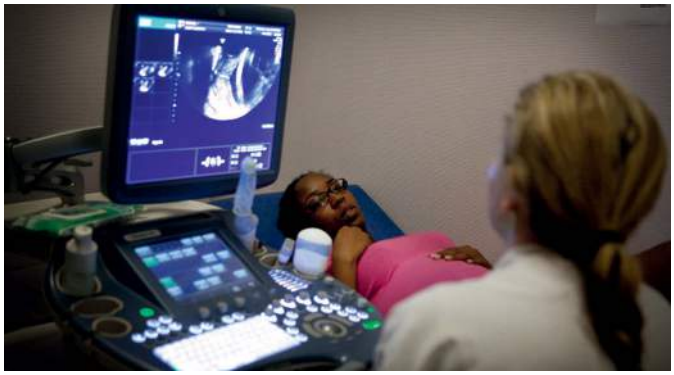


Figure 32.21 An expectant mother undergoes an ultrasound scan. The image of her baby is built up by computer and appears on the monitor.

first appearance in the family photo album is in the form of an ante-natal (before birth) scan!

This technique has many other uses in medicine. It can be used to detect gallstones or kidney stones (two very painful complaints), so men as well as women may undergo this type of scan.

The technique of ultrasound scanning is rather similar to the way in which sailors use echo sounding and echo location to detect the seabed and shoals of fish. Ultrasound waves are directed into the patient's body. These waves are partially reflected at the boundaries between different tissues and the reflected waves are detected and used to construct the image.

In this chapter, we will look at the principles of ultrasound scanning, and we will also look at another technique in which ultrasound is used to measure the rate of blood flow in the body.

Working with ultrasound

Ultrasound is any sound wave that has a frequency above the upper limit of human hearing. This is usually taken to mean frequencies above 20 kHz (20 000 Hz), although the limit of hearing decreases with age to well below this figure. In medical applications, the typical frequencies used are in the megahertz range.

Sound waves are longitudinal waves. They can only pass through a material medium; they cannot pass through a vacuum. The speed of sound (and hence of ultrasound) depends on the material. In air, it is approximately 330 m s^{-1}; it is higher in solid materials. A typical value for body tissue is 1500 m s^{-1}. Using the wave equation $v = f\lambda$, we can calculate the wavelength of 2.0 MHz ultrasound waves in tissue:

$$\lambda = \frac{v}{f} = \frac{1500}{2.0} \times 10^{6}$$

$$= 7.5 \times 10^{-4}\,\text{m} \approx 1\,\text{mm}$$

This means that 2.0 MHz ultrasound waves will be able to distinguish detailed features whose dimensions are of the order of 1 mm. Higher-frequency waves have shorter wavelengths and these are used to detect smaller features inside the body. Unfortunately, higher frequency waves are absorbed more strongly and so a more intense beam must be used.

Producing ultrasound

Like audible sound, ultrasound is produced by a vibrating source. The frequency of the source is the same as the frequency of the waves it produces. In ultrasound scanning, ultrasonic waves are produced by a varying electrical voltage in a transducer. The same device also acts as a detector. (You should recall from Chapter 25 that a transducer is any device that changes energy from one form to another.)

At the heart of the transducer is a **piezo-electric crystal**, such as quartz. This type of crystal has a useful property: when a voltage is applied across it in one direction, it shrinks slightly – see Figure 32.22a. When the voltage is reversed, it expands slightly. So an alternating voltage with frequency f causes the crystal to contract and expand at the same frequency f. We say that the voltage induces a strain in the crystal. In the best piezo-electric substances, the maximum value of strain is about 0.1%; in other words, the crystal's width changes by about one part in a thousand.

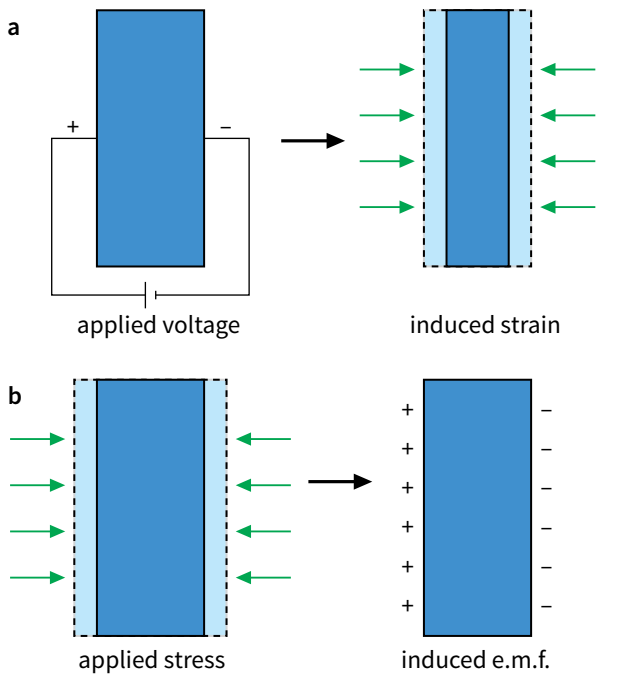


Figure 32.22 The piezo-electric effect. **a** An applied voltage causes a piezo-electric crystal to contract or expand. **b** An applied stress causes an induced e.m.f. across the crystal.

In a piezo-electric transducer, an alternating voltage is applied across the crystal, which then acts as the vibrating source of ultrasound waves. A brief pulse of ultrasound waves is sent into the patient's body; the transducer then receives an extended pulse of reflected ultrasound waves.

Detecting ultrasound

The transducer also acts as the detector of reflected ultrasound waves. It can do this because the piezo-electric effect works in reverse: a varying stress applied to the crystal produces a varying e.m.f. across the crystal – see Figure 32.22b. To maximise the effect, the frequency of the waves must match the resonant frequency of the crystal. The optimum size of the crystal is half the wavelength $\left(\frac{\lambda}{2}\right)$ of the ultrasound waves.

Figure 32.23 shows the construction of a piezo-electric ultrasound transducer. Note the following features:

- The crystal is now usually made of polyvinylidene difluoride. Previously, quartz and lead zirconate titanate were used.
- The outer case supports and protects the crystal.
- At the base is the acoustic window, made from a material that is a good transmitter of ultrasound.
- Behind the crystal is a large block of damping material (usually epoxy resin). This helps to stop the crystal vibrating when a pulse of ultrasound has been generated. This is necessary so that the crystal is not vibrating when the incoming, reflected ultrasound waves reach the transducer.

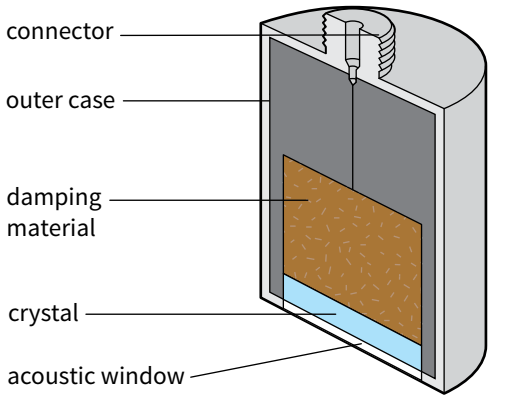


Figure 32.23 A section through an ultrasound transducer.

QUESTIONS

11 Quartz is an example of a piezo-electric material. The speed of sound in quartz is 5700 m s^{-1}.

a Calculate the wavelength of ultrasound waves of frequency 2.1 MHz in a quartz crystal.

b If the crystal is to be used in an ultrasound transducer, its thickness must be half a wavelength. Calculate the thickness of the transducer.

12 Piezo-electric crystals have many applications other than in ultrasound scanning. For example, they are used:

a in gas lighters (to produce a spark)

b in inkjet printers (to break up the stream of ink into droplets)

c in guitar pickups (to connect the guitar to an amplifier)

d in the auto-focus mechanism of some cameras (to move the lens back and forth).

For each of these examples, state whether the piezo-electric effect is being used to convert energy in the vibrations of the crystal to electrical energy or the other way round.

Echo sounding

The principle of an ultrasound scan is to direct ultrasound waves into the body. These pass through various tissues and are partially reflected at each boundary where the wave speed changes. The reflected waves are then detected and used to construct an internal image of the body.

Figure 32.24 shows what happens when a beam of ultrasound reaches a boundary between two different media. The beam is partially refracted (that is, the transmitted beam has changed direction) and partially reflected. This diagram should remind you of the way in which a ray of light is refracted and reflected when it strikes the boundary between two media. It is the change in speed which causes the refraction of a wave.

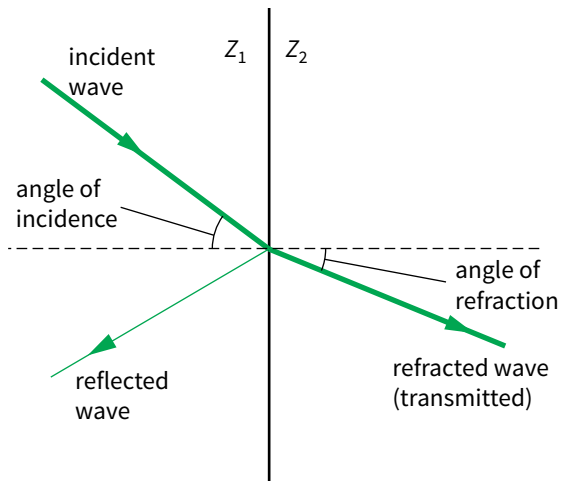


Figure 32.24 An ultrasound wave is both refracted and reflected when it strikes the boundary between two different materials.

For ultrasound, we are interested in the fraction of the incident intensity of ultrasound that is reflected at the boundary. This depends on the **acoustic impedance** Z of each material. This quantity depends on the density ρ and the speed of sound c in the material. Acoustic impedance is defined as follows:

acoustic impedance = density × speed of sound

$Z = \rho c$

Since the unit of density is $kg\,m^{-3}$ and the unit of speed is $m\,s^{-1}$, the unit of acoustic impedance Z is $kg\,m^{-2}\,s^{-1}$.

Table 32.3 shows values of ρ, c and Z for some materials that are important in medical ultrasonography.

Material	Density / $kg\,m^{-3}$	Speed of sound / $m\,s^{-1}$	Acoustic impedance / $10^6\,kg\,m^{-2}\,s^{-1}$
air	1.3	330	0.0004
water	1000	1500	1.50
Biological			
blood	1060	1570	1.66
fat	925	1450	1.34
soft tissue (average)	1060	1540	1.63
muscle	1075	1590	1.71
bone (average; adult)	1600	4000	6.40
Transducers			
barium titanate	5600	5500	30.8
lead zirconate titanate	7650	3790	29.0
quartz	2650	5700	15.1
polyvinylidene difluoride	1780	2360	4.20

Table 32.3 The density (ρ), speed of sound in air (c) and acoustic impedance (Z) of some materials important in medical scanning.

Calculating reflected intensities

When an ultrasound beam reaches the boundary between two materials, the greater the **difference** in acoustic impedances, the greater the reflected fraction of the ultrasound waves. For normal incidence (i.e. angle of incidence = 0°) the ratio of the reflected intensity I_r to the incident intensity I_0 is given by:

$$\frac{I_r}{I_0} = \frac{(Z_2 - Z_1)^2}{(Z_2 + Z_1)^2}$$

or

$$\frac{I_r}{I_0} = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1}\right)^2$$

where Z_1 and Z_2 are the acoustic impedances of the two materials (see Figure 32.24). The ratio $\frac{I_r}{I_0}$ indicates the fraction of the intensity of the beam that is reflected.

WORKED EXAMPLE

2 A beam of ultrasound is normally incident on the boundary between muscle and bone. Use Table 32.3 to determine the fraction of its intensity which is reflected.

Step 1 Write down the values of Z_1 (for muscle) and Z_2 (for bone).

$Z_1 = 1.71 \times 10^6\,kg\,m^{-2}\,s^{-1}$

$Z_2 = 6.40 \times 10^6\,kg\,m^{-2}\,s^{-1}$

Step 2 Substitute these values in the equation for $\frac{I_r}{I_0}$.

$$\frac{I_r}{I_0} = \frac{(Z_2 - Z_1)^2}{(Z_2 + Z_1)^2}$$

Hint: *We can use this equation because we know that the angle of incidence = 0°.*

$$\frac{I_r}{I_0} = \frac{(6.40 - 1.71)^2}{(6.40 + 1.71)^2}$$

$= 0.33$

Hint: *We can ignore the factor of 10^6 in the Z values because this is a factor common to all the values, so they cancel out.*

So 33% of the intensity of ultrasound will be reflected at the muscle–bone boundary.

Comparing acoustic impedances

A big change in acoustic impedance gives a large fraction of reflected intensity. Inspection of Table 32.3 shows that:

- a very large fraction ($\frac{I_r}{I_0} \approx 99.95\%$) of the incident ultrasound will be reflected at an air–tissue boundary
- a large fraction will be reflected at a tissue–bone boundary (as shown in Worked example 2)
- very little will be reflected at a boundary between soft tissues including fat and muscle.

This means that bone shows up well in an ultrasound scan, but it is difficult to see different soft tissues (Figure 32.25). Another problem is that the patient's skin is in contact with air, and 99.95% of the ultrasound will be reflected before it has entered the body. To overcome this, the transducer must be 'coupled' to the skin using a gel whose impedance matches that of the skin. This process of **impedance matching** explains why the patient's skin is smeared with gel before a scan.



Figure 32.25 Ultrasound scan of a fetus at 20 weeks; the baby's skin is clearly visible, as are its bony skull and ribs.

The acoustic impedance of the gel is typically $1.65 \times 10^6\,\text{kg m}^{-2}\,\text{s}^{-1}$ and that of skin is $1.71 \times 10^6\,\text{kg m}^{-2}\,\text{s}^{-1}$. With gel between the skin and the transducer, the percentage of the intensity reflected is 0.03%.

The poor match of impedance between air and tissue means that ultrasound cannot penetrate the lungs. The operator must take care to avoid any bubbles of gas in the intestines. Bones are also difficult to see through. For an ultrasound scan of the heart, the probe must be directed through the gap between two ribs.

As ultrasound waves pass through the body, they are gradually absorbed. Their absorption follows the same exponential pattern as we saw earlier for X-rays. The intensity I decreases with distance x according to the equation

$$I = I_0 e^{-\alpha x}$$

Here, α is the absorption coefficient, equivalent to the quantity μ in the absorption equation for X-rays; its value varies with the nature of the tissue through which the ultrasound is passing, and with the frequency of the ultrasound. In practice, absorption is not a serious problem in an ultrasound scan as scanning relies on the reflection of ultrasound at the boundaries between different tissues.

Ultrasound scanning

There are several different types of ultrasound scan which are used in practice. To illustrate the basic principles, we will concentrate on the A-scan and the B-scan.

A-scan

This is the simplest type of scan. A pulse of ultrasound is sent into the body and the reflected 'echoes' are detected and displayed on an oscilloscope or computer screen as a voltage–time graph.

A pulse generator controls the ultrasound transducer. It is also connected to the time base of the oscilloscope. Simultaneously, the pulse generator triggers a pulse of ultrasound which travels into the patient and starts a trace on the screen. Each partial reflection of the ultrasound is detected and appears as a spike on the screen (Figure 32.26).

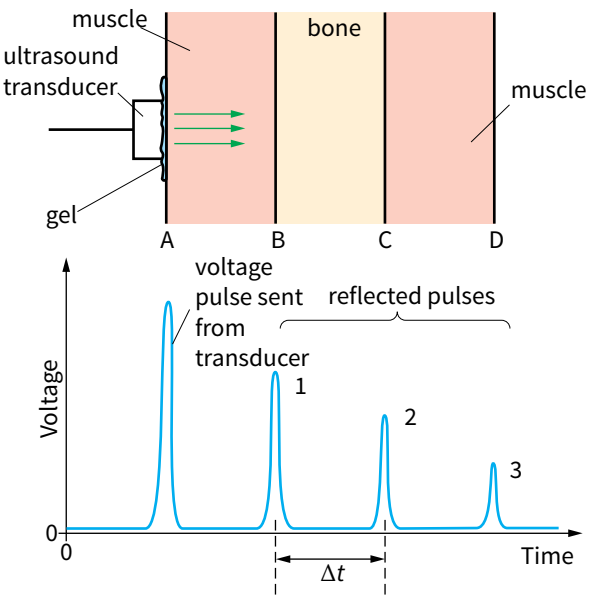


Figure 32.26 An A-scan. Information about the depth of reflecting tissues can be obtained from the positions of the spikes along the time axis; their relative amplitudes can indicate the nature of the reflecting surfaces.

QUESTIONS

13 Calculate the acoustic impedance of muscle tissue. (Density = 1075 kg m^{-3}, speed of sound = 1590 m s^{-1}.)

14 Determine the fraction of the intensity of an ultrasound beam that is reflected when a beam is incident normally on a boundary between water and fat. (Use values from Table 32.3.)

15 The ultrasound image shown in Figure 32.25 clearly shows the baby's skin and some bones. Explain why these show up clearly while softer organs inside its body do not.

16 Explain why ultrasound cannot readily be used to examine the brain. Suggest one or more alternative scanning techniques that can be used for this.

In Figure 32.26, pulses 1, 2 and 3 are reflected at the various boundaries. Pulse 1 is the reflection at the muscle–bone boundary at B. Pulse 2 is the reflection at the bone–muscle boundary at C. The time Δt is the time taken for the ultrasound to travel **twice** the thickness of the bone. Finally, pulse 3 is the reflection at the muscle–air boundary at D. The thickness of the bone can be determined from this A-scan.

time interval between pulses 1 and 2 = Δt

$$\text{thickness of bone} = \frac{\text{distance travelled by ultrasound}}{2}$$

$$= \frac{c\Delta t}{2}$$

where c is the speed of the ultrasound in the bone (see Worked example 3).

Because ultrasound waves are gradually attenuated as they pass through the body (their energy is absorbed so that their amplitude and intensity decrease), the echoes from tissues deeper in the body are weaker and must be amplified.

A-scans are used for some straightforward procedures such as measuring the thickness of the eye lens.

WORKED EXAMPLE

3 In a particular A-scan, similar to Figure 32.26, the time interval between pulses 1 and 2 is 12 μs. The speed of ultrasound in bone is about 4000 m s^{-1}. Determine the thickness of the bone.

Step 1 Determine the distance travelled by the ultrasound in the time interval of 12 μs.

distance = speed × time

distance = $4000 \times 12 \times 10^{-6} = 4.8 \times 10^{-2}$ m

Step 2 Calculate the thickness of the bone.

Hint: *The distance you have just calculated must be halved because the ultrasound has to travel through the bone twice.*

$$\text{thickness of bone} = \frac{4.8 \times 10^{-2}}{2}$$

$$= 2.4 \times 10^{-2} \text{ m (2.4 cm)}$$

B-scan

In a B-scan, a detailed image of a cross-section through the patient is built up from many A-scans. The ultrasound transducer is moved across the patient's body in the area of interest. Its position and orientation are determined by small sensors attached to it.

Each reflected pulse is analysed to determine the depth of the reflecting surface (from the time of echo) and the nature of the surface (from the amplitude of the reflected wave). A two-dimensional image is then built up on a screen by positioning dots to represent the position of the reflecting surfaces and with brightness determined by the intensity of the reflection, brighter dots indicating more reflected ultrasound (see Figure 32.27).

Figure 32.28 shows the result of a typical B-scan. Because it takes several seconds for the scanner to move across the body, problems can arise if the organs of interest are moving – this gives a blurred image.

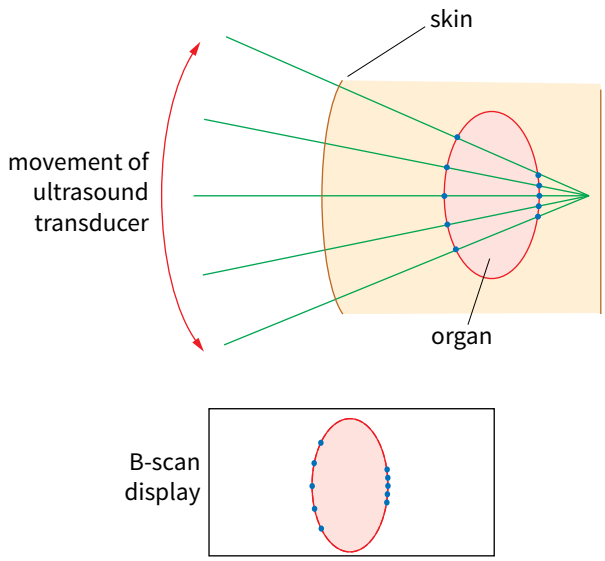


Figure 32.27 In a B-scan, dots are produced on the screen rather than pulses as in the A-scan. By moving the transducer, a series of dots on the screen traces out the shape of the organ being examined.

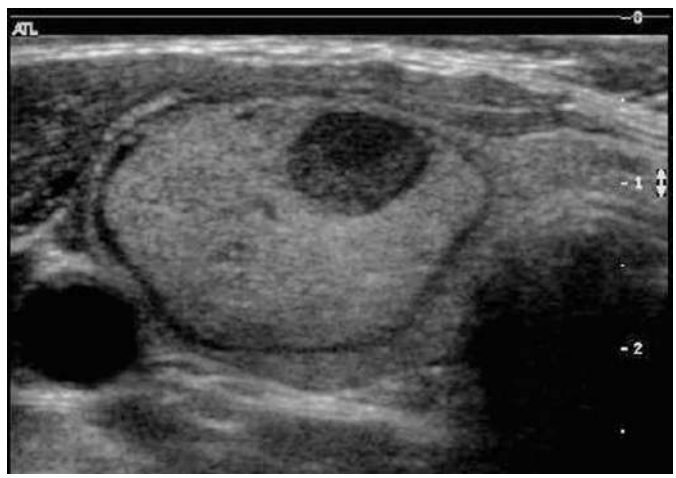


Figure 32.28 An ultrasonic B-scan of an abnormal thyroid gland.

QUESTIONS

17 Two consecutive peaks in an ultrasound A-scan are separated by a time interval of 0.034 ms. Calculate the distance between the two reflecting surfaces. (Assume that the speed of sound in the tissue between the two surfaces is 1540 m s^{-1}.)

18 Explain why an ultrasound B-scan, rather than X-rays, is used to examine a fetus.

Magnetic resonance imaging

Magnetic resonance imaging, or MRI, is a diagnostic technique used in medicine. It can provide images (including moving images) of the insides of a patient. It does not rely on exposing patients to ionising radiation such as X-rays; rather, it relies on the fact that some atomic nuclei behave like tiny magnets in an external magnetic field.

(MRI was originally known as nuclear magnetic resonance imaging, but the word 'nuclear' was dropped because it was associated in patients' minds with bombs and power stations. To emphasise: MRI does **not** involve radioactive decay, fission or fusion.)

As in CT scanning, MRI scanning involves electromagnetic radiation, in this case radio frequency (RF) electromagnetic waves. The patient lies on a bed in a strong magnetic field (Figure 32.29), RF waves are sent into their body, and the RF waves that emerge are detected. From this, a picture of the patient's insides can be built up by computer. As we will see, MRI gives rather different information from that obtained by the other non-invasive techniques such as X-rays or ultrasound.

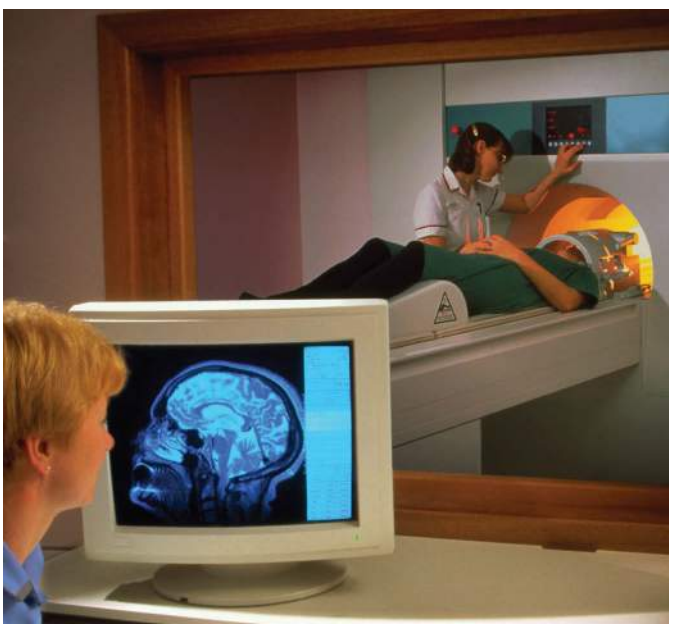


Figure 32.29 A patient undergoing an MRI scan of the brain. This is a form of tomography; the display shows different 'slices' through the patient's brain.

Principles of nuclear magnetic resonance

The nuclei of certain atoms have a property called **spin**, and this causes them to behave as tiny magnets in a magnetic field. In MRI, it is usually the nuclei of hydrogen atoms that are studied, since hydrogen atoms are present in all tissues. A hydrogen nucleus is a proton, so we will consider protons from now on.

A proton has positive charge. Because it spins, it behaves like a tiny magnet with N and S poles. Figure 32.30a shows a number of protons aligned randomly.

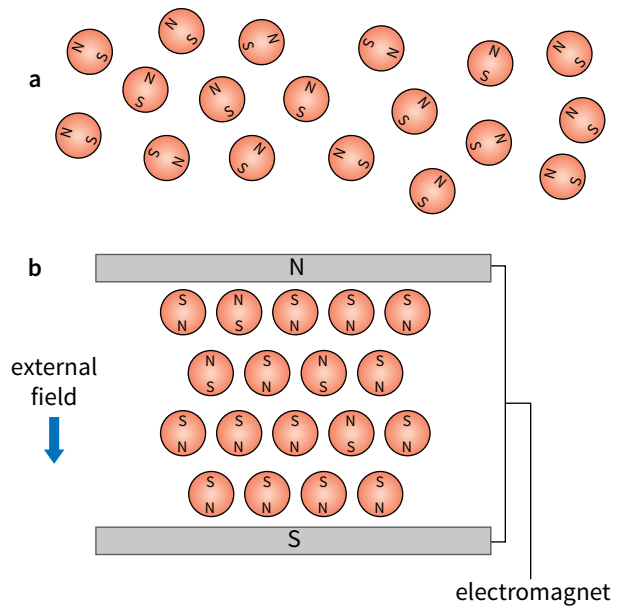


Figure 32.30 How protons behave in a strong magnetic field. **a** Protons are randomly directed when there is no external magnetic field. **b** Because protons are magnetic, a strong external magnetic field causes most of them to align themselves with the field.

When a very strong external magnetic field is applied, the protons respond by lining up in the field (just as plotting compasses line up to show the direction of a magnetic field). Most line up with their N poles facing the S pole of the external field, a low energy state; a few line up the other way round, which is an unstable, higher energy state (Figure 32.30b).

A proton does not align itself directly along the external field. In practice, its magnetic axis rotates around the direction of the external field (Figure 32.31), just like the axis of a spinning top. This rotation or gyration action is known as **precession**.

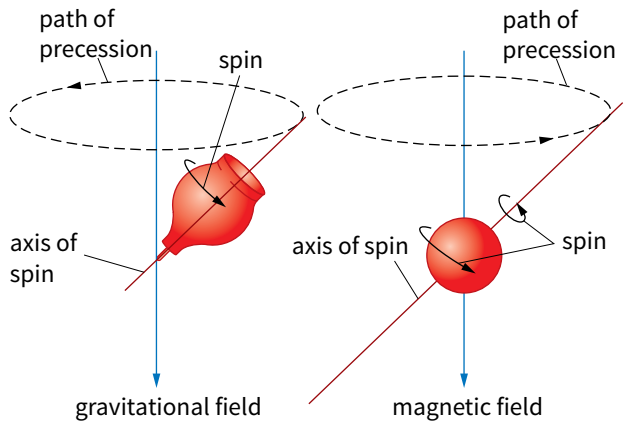


Figure 32.31 A spinning top (left) rotates about its axis; at the same time, its axis precesses about the vertical, which is the direction of the gravitational field. In a similar way, a proton (right) spins and its axis of rotation precesses about the direction of the external magnetic field.

The angular frequency of precession is called the **Larmor frequency** ω_0, and depends on the individual nucleus and the magnetic flux density B_0 of the magnetic field:

$$\omega_0 = \gamma B_0$$

So, the stronger the external field, the faster the protons precess about it. The quantity γ is called the **gyromagnetic ratio** for the nucleus in question and is a measure of its magnetism. (Note that the Larmor frequency is measured in radians per second. This means that, strictly speaking, it is an angular velocity, not a frequency.)

For protons, γ has the approximate value $2.68 \times 10^8\,\text{rad}\,\text{s}^{-1}\,\text{T}^{-1}$. To determine the frequency f_0 of the precessing nuclei, we can use the equation:

$$\omega_0 = 2\pi f_0$$

Therefore:

$$f_0 = \frac{\lambda B_0}{2\pi}$$

In an MRI scanner, the external magnetic field is very strong, of the order of 1.5 T (thousands of times the strength of the Earth's field). The precession frequency f_0 is:

$$f_0 = \frac{2.68 \times 10^8 \times 1.5}{2\pi} = 6.4 \times 10^7\,\text{Hz} = 64\,\text{MHz}$$

This frequency lies in the radio frequency (RF) region of the electromagnetic spectrum. You should recall that **resonance** requires a system with a natural frequency of vibration; when it is stimulated with energy of the same frequency, it absorbs energy. In MRI, protons precessing about the strong external field are exposed to a burst or pulse of RF waves whose frequency equals the frequency of precession. Each proton absorbs a photon of RF energy and flips up into the higher energy state; this is **nuclear magnetic resonance** (Figure 32.32).

Now we come to the useful bit. The RF waves are switched off and the protons gradually relax into their lower energy state. As they do so, they release their excess energy in the form of RF waves. These can be detected, and the rate of **relaxation** tells us something about the environment of the protons.

In Figure 32.32, you can see that the relaxation of the protons follows an exponential decay pattern. Curves like this are characterised by two **relaxation times**:

- T_1, the spin–lattice relaxation time, where the energy of the spinning nuclei is transferred to the surrounding 'lattice' of nearby atoms
- T_2, the spin–spin relaxation time, where the energy is transferred to other spinning nuclei.

These relaxation times depend on the environment of the nuclei. For biological materials, it depends on their water content:

- Water and watery tissues (e.g. cerebrospinal fluid) have relaxation times of several seconds.
- Fatty tissues (e.g. white matter in the brain) have shorter relaxation times, several hundred milliseconds.
- Cancerous tissues have intermediate relaxation times.

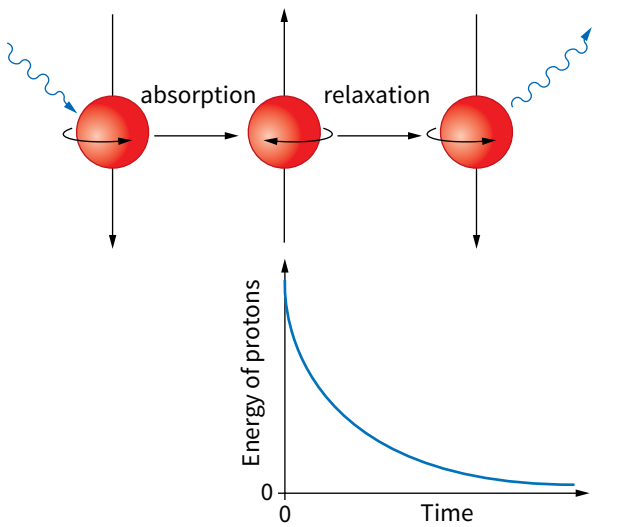


Figure 32.32 In nuclear magnetic resonance, a spinning nucleus is flipped into a higher energy state when it absorbs a photon of RF energy; then it relaxes back to its lower energy state.

This means that different tissues can be distinguished by the different rates at which they release energy after they have been forced to resonate. That is the basis of medical applications of nuclear magnetic resonance.

QUESTIONS

19 Protons precess at a frequency of 42.6 MHz in an external field of magnetic flux density 1.0 T.

a Determine the frequency at which will they precess in a field of magnetic flux density 2.5 T.

b State the frequency of RF radiation that will cause the protons to resonate in this stronger magnetic field.

20 Figure 32.33 shows how the amplitude of RF waves coming from watery tissue varies after resonance. Copy the graph and add lines and labels to show the graphs you would expect to see for cancerous and fatty tissues.

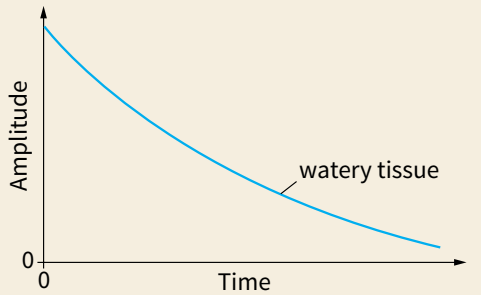


Figure 32.33 See Question 20.

MRI scanner

Figure 32.34 shows the main components of an MRI scanner. The main features are:

- A large superconducting magnet which produces the external magnetic field (up to 2.0 T) needed to align the protons. Superconducting magnets are cooled to 4.2 K (−269 °C) using liquid helium.
- An RF coil that transmits RF pulses into the body.
- An RF coil that detects the signal emitted by the relaxing protons.
- A set of gradient coils. (For clarity, only one pair of gradient coils is shown in Figure 32.34.) These produce an additional external magnetic field that varies across the patient's body. These coils are arranged such that they alter the magnitude of the magnetic flux density across the length, depth and width of the patient. This ensures that the Larmor frequency of the nuclei within the patient will be slightly different for each part of the body. This means that only a small volume of the body is at exactly the right field value for resonance and so the computer can precisely locate the source of the RF signal within the patient's body and construct an image.
- A computer that controls the gradient coils and RF pulses, and which stores and analyses the received data, producing and displaying images.

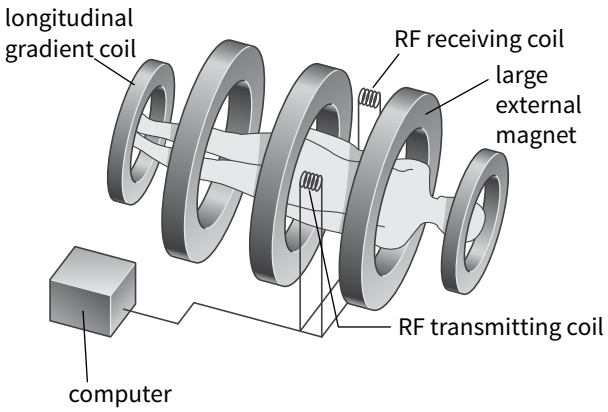


Figure 32.34 The main components of an MRI scanner.

Procedure

The patient lies on a bed which is moved into the centre of the electromagnet. The central imaging section is about 0.9 m long and 0.6 m in diameter. The magnetic field is very uniform, with variations smaller than 50 parts per million in its strength. The gradient field is superimposed on this fixed field. An RF pulse is then transmitted into the body, causing protons to flip (resonate). Then the receiving coils pick up the relaxation signal and pass it to the computer.

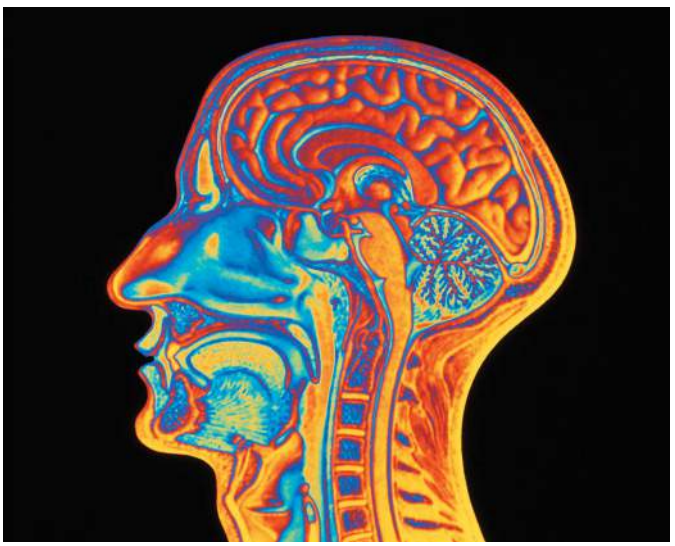


Figure 32.35 MRI scan through a healthy human head. Different tissues, identified by their different relaxation times, are coloured differently.

The result is an image like the one shown in Figure 32.35. This image has been coloured to show up the different tissues, which are identified by their different relaxation times.

Advantages and disadvantages of MRI

MRI has several advantages compared to other scanning techniques:

- It does not use ionising radiation which causes a hazard to patients and staff.
- There are no moving mechanisms, just changing currents and magnetic fields.
- The patient feels nothing during a scan (although the gradient coils are noisy as they are switched), and there are no after-effects.
- MRI gives better soft-tissue contrast than a CT scan, although it does not show bone as clearly.
- Computer images can be generated showing any section through the volume scanned, or as a three-dimensional image.

One disadvantage of MRI is that any metallic objects in the patient, such as surgical pins, can become heated. Also, heart pacemakers can be affected, so patients with such items cannot undergo MRI scans. Loose steel objects must not be left in the room as these will be attracted to the magnet, and the room must be shielded from external radio fields.

Figure 32.36 shows how an MRI scan can be combined with a CT scan to show detail of both bone and soft tissue, allowing medical staff to see how the two are related. Compare this with Figure 32.35.



Figure 32.36 A combined CT scan and MRI scan, showing how the tissues revealed by MRI relate to the bone structure shown by X-rays.

QUESTION

21 An MRI scan might be considered a safer procedure than a CT scan.

a Explain why it might be considered to be safer.

b Why might a CT scan be chosen in preference to an MRI scan?

c Explain why MRI is described as **non-invasive.**

Summary

- X-rays are short-wavelength, high-frequency electromagnetic radiation, produced when electrons are decelerated.
- The intensity of an X-ray beam is the power transmitted per unit cross-sectional area.
- The intensity of a collimated X-ray beam decreases exponentially according to the equation $I = I_0 e^{-\mu x}$, where μ is the attenuation coefficient of the medium. μ has units of m^{-1} (or cm^{-1} or mm^{-1}).
- X-ray images can be improved using image intensifiers and contrast media (such as barium or iodine).
- Ultrasound is a longitudinal wave with a frequency greater than 20 kHz.
- Ultrasound transducers use the piezo-electric effect to generate and detect ultrasound waves.
- The acoustic impedance Z of a material depends on its density ρ and the speed c of sound:

  $Z = \rho c$

- The fraction of the intensity of an ultrasound wave reflected at a boundary is given by:

  $\frac{I_r}{I_0} = \frac{(Z_2 - Z_1)^2}{(Z_2 + Z_1)^2}$ or $\frac{I_r}{I_0} = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1}\right)^2$

- To transfer a high proportion of the intensity of an ultrasound pulse into the patient's body, an impedance-matching gel must be used with acoustic impedance almost the same as that of the skin.
- In MRI scanning, spinning, precessing protons are forced to resonate using radio frequency pulses. RF radiation from relaxing protons is used to obtain diagnostic information about internal organs, particularly soft tissues.
- The main components of an MRI scanner are: superconducting magnet, RF transmitter coil, RF receiver coil, set of gradient coils and computer.

End-of-chapter questions

1 **a** Explain what is meant by **ionising radiation** and explain why it can be harmful to humans. [2]

b Which of the following scans use ionising radiation? [2]

X-ray shadow imaging
ultrasound A-scan
ultrasound B-scan
magnetic resonance imaging
CT scan

2 Calculate the minimum wavelength (in air) of X-rays produced when the accelerating potential across the source is 20 kV. [2]

3 Explain why a gel is used between the skin and the transducer when an ultrasound scan of a fetus is taken. [2]

4 For ultrasound of frequency 3.5 MHz, the acoustic impedance of muscle is $1.78 \times 10^6\,kg\,m^{-2}\,s^{-1}$, and that of soft tissue is $1.63 \times 10^6\,kg\,m^{-2}\,s^{-1}$.

Calculate the percentage of the incident ultrasound reflected at a muscle–soft tissue boundary. [3]

5 A transducer produces ultrasonic waves of frequency 800 kHz. The speed of sound in the crystal is 5200 m s^{-1}.

Calculate the optimum thickness for the crystal. [2]

6 Explain what is meant by the **Larmor frequency** and state its units. [2]

7 State and explain **two** reasons why full-body CT scans are not offered for regular checking of healthy patients. [2]

8 **a** Explain with the aid of a simple, labelled diagram how X-rays are produced. [5]

b Discuss the energy changes in the production of X-rays. [3]

9 Figure 32.37 shows the spectrum of X-rays produced from an X-ray source.

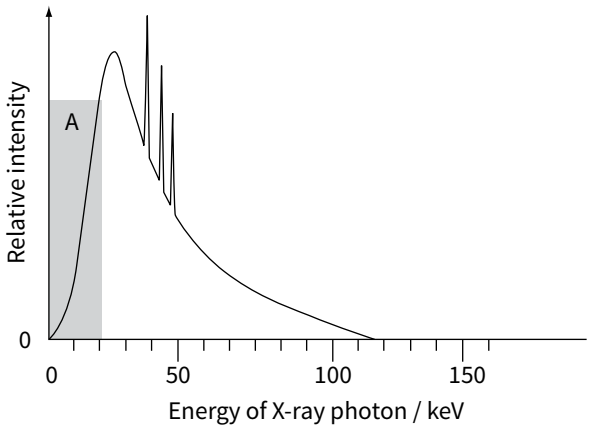


Figure 32.37 For End-of-chapter Question 9.

a Describe the process by which:

i the three sharp peaks of high-intensity X-rays are produced [2]

ii the broad band of X-rays is produced. [2]

b The X-rays in the shaded region, labelled A, are filtered out using an aluminium filter. Explain:

i why it is advantageous to filter these X-rays out [2]

ii why aluminium is a suitable material to filter them out. [2]

c Calculate the maximum frequency of X-rays produced by this tube. [3]

10 **a** Describe the main principles of how a CT scan image is produced and compare this with the formation of an X-ray shadow image. [7]

b A CT image is sometimes superimposed on an image from an MRI scan. Explain the advantages of using this dual approach. [2]

11 **a** An X-ray beam, containing X-rays with a variety of frequencies and which has an intensity of 4.0×10^5 W, is incident on an aluminium plate of thickness 5.0 cm. The average linear attenuation coefficient is 250 m^{-1}.

i Calculate the intensity of the transmitted beam. [3]

ii Explain the advantages of passing the X-rays through this aluminium plate prior to their being incident on a patient. [3]

b Outline the use of an anti-scatter grid and explain its role in improving an X-ray shadow image. [5]

12 **a** Explain what is meant by **acoustic impedance** and outline its role in the use of ultrasound scans. [3]

b Brain tissue has a density of $1.04 \times 10^3\,\text{kg m}^{-3}$ and ultrasound travels at $1.58 \times 10^3\,\text{m s}^{-1}$ through it. Calculate the acoustic impedance of brain tissue. [2]

c Figure 32.38 shows the trace formed on the screen of an oscilloscope when ultrasound is reflected from the front and rear surfaces of the head of a fetus. The time-base of the oscilloscope is set at $10\,\mu\text{s div}^{-1}$.

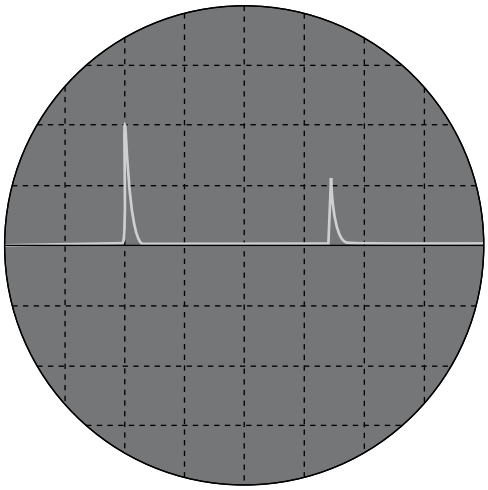


Figure 32.38 For End-of-chapter Question 12.

i Explain why the second peak is lower than the first. [1]

ii Calculate the diameter of the head of the fetus. [3]

13 **a** With reference to nuclear magnetic resonance, explain what is meant by the **gyromagnetic ratio**. [1]

b A doctor is conducting an MRI scan on a patient. There is a steady magnetic field of 1.8 T and superimposed on this is a much weaker field which oscillates at an RF frequency of 48 MHz.

i Assuming the RF frequency causes resonance of the protons, calculate their gyromagnetic ratio. [3]

ii Explain what **resonance** means in this application. [2]

iii Describe what happens when the RF field is removed and how this leads to information about the internal organs of the patient. [3]

P2:
Planning, analysis and evaluation

Learning outcomes

You should be able to:

- develop a systematic approach to carrying out experiments, including planning, setting up apparatus, investigating and recording results, analysing data and writing conclusions.
- plan an investigation to test a relationship or investigate a problem, identifying the dependent, independent and control variables
- use logarithms and logarithmic graphs
- combine uncertainties
- plot error bars on graphs and find uncertainties in gradients and intercepts

The practical work in the second year of your A level course builds on what you have covered in the first year. Tests and examinations you may take during your studies will ask you to demonstrate your abilities in two key areas:

- planning experiments
- analysis and evaluation of your results, including any conclusions you can draw.

In this chapter we will look at the different skills that you need to demonstrate your practical abilities.

Planning

As you progress through your A level physics studies, you should think about and continually develop your approach to planning experiments. The experiments you will be asked to plan by your teacher will usually provide you with a scenario and sometimes a relationship or an equation that you are to use and test. Often particular items of apparatus are mentioned and you should use these items, even if you think there is a better method. Sometimes the experiment will seem familiar to you and sometimes it will be completely new. Before you start, it is important to read the scenario carefully. It is also important to read, understand and re-read any questions you need to answer, before starting on your plan.

In producing your plan, you should draw a diagram showing the actual apparatus to be used, and pay particular attention to:

- the procedure to be followed
- the measurements to be taken
- the control of variables
- the analysis of the data
- the safety precautions to be taken.

Defining the problem – identifying the variables

It may seem obvious, but the first thing is to identify the problem. To do that you must identify:

- the **independent variable** in the experiment
- the **dependent variable** in the experiment
- the quantities that are to be **controlled** or kept constant.

It is usually a good idea to start with a clear statement about the variables as the first part of your plan.

Here is an example of the sort of problem you might face in planning an experiment.

The deflection of a balloon by a jet of air is shown in Figure P2.1. You need to plan an investigation to show that $\tan\theta$ is inversely proportional to v^2, where θ is the angle between the ground and the string of the balloon and v is the speed of the air hitting the balloon. You are unlikely to have seen this experiment before, but this should not concern you.

In this example, the speed v of the air is the variable that you will need to alter and so this is the independent variable; the angle θ is the variable that changes as a result, and so this is the dependent variable.

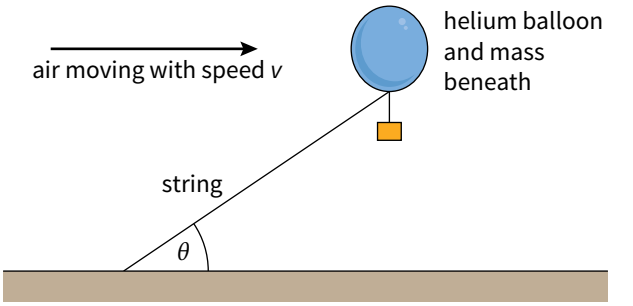


Figure P2.1 A balloon is deflected as the air moves at different speeds.

But what quantities are kept constant? These are the quantities that are controlled. You may be able to think of many, such as the total mass of the balloon and the mass placed underneath it. This total mass is certainly one quantity that should be kept constant, but it is not something that is likely to change during the course of the experiment. In terms of planning the experiment, you need to think about quantities that may easily change during the experiment if care is not taken. For example you might realise that:

- the balloon may be deflected downwards if the air blows more strongly; then the air will hit the top, rather than the middle, of the balloon
- the balloon may warm up and expand in size; then more air will hit the balloon.

In either case the experiment will then not just be testing the effect of the air speed. Your plan should clearly state what you need to keep constant. In our example:

- make sure that the jet of air is always horizontal and hits the middle of the balloon
- keep the temperature of the air inside the balloon constant.

As you can see, you need to think carefully about the experiment. Avoid giving wrong suggestions, for example keeping the length of the string constant. If the string is longer the balloon may be out of the jet of air, and so it is not entirely a wrong suggestion, but it is not a primary quantity to be kept constant.

QUESTION

1 An experiment is being planned to measure how the resistance of a wire depends on the cross-sectional area of the wire. What are the independent and dependent variables? Suggest **three** quantities that might be controlled.

Methods of data collection

The next task is to think about how you are going to carry the experiment out. Once you have a method in mind, you should be able to:

- describe the method to be used to vary the independent variable
- describe how the independent variable is to be measured
- describe how the dependent variable is to be measured
- describe how other variables are to be controlled
- describe, with the aid of a clear, labelled diagram, the arrangement of apparatus for the experiment and the procedures to be followed.

It may be worthwhile jotting down your thoughts about the experiment on a rough piece of paper before you start, but do make sure that you write up all your points. It is particularly important to say what you actually measure and how the measurement is made. It may seem obvious to you that a particular quantity is measured, but unless you write it down it is not part of your plan.

Always check that in your account you have clearly said:

- what you will measure and how you will measure it
- what you will change and how you will change it
- what you will keep constant and how this is achieved.

Let's use the example of the balloon deflected by a current of air to show how you could approach this part of the plan.

Describing the experiment

First, describe how to change the independent variable and state what instrument is used to measure it. The apparatus shown in Figure P2.1 does not help very much and you must use your general knowledge and suggest, for example, that a wind fan be used. To change v will mean either changing the distance from the fan to the balloon or adjusting the power supply voltage to the fan.

You will also a need a wind speed indicator, sometimes called an anemometer, to measure the independent variable. Perhaps you have never seen or used a wind speed indicator, but clearly there must be an instrument to actually measure v. You may have to think very carefully to find a sensible instrument when the quantity is unusual.

The instrument to measure the dependent variable is much simpler – a protractor – although it may have to be a large protractor. Alternatively, you could use a ruler to measure the height h from the bench to the top of the string and the length l of the string, and then use $\sin\theta = \frac{h}{l}$ to find θ.

At this stage, try to suggest how to keep at least one of your 'controlled quantities' constant. For example, for the suggestions made earlier:

- compensate for deflection of the balloon downwards by a faster wind by lowering the fan, so that air from the fan stays horizontal and is always aimed at the centre of the balloon.
- keep the temperature of the air inside the balloon constant by leaving the balloon in a room with constant temperature for many hours before the experiment starts, and ensuring that the fan used blows air from the room.

As you can see, you have to think carefully about what happens during the experiment.

As you now have a clear idea of the experiment in your mind, draw a labelled diagram showing everything that you have mentioned. In this example you could draw the fan, possibly its supply, a protractor, and even an anemometer.

Now describe your planned experiment, making sure that you describe a logical sequence of steps to follow. If you find this difficult, a labelled diagram of each step can sometimes be useful. For example, you might draw a diagram where you remove the balloon and put your wind speed measurement device in place of the balloon to measure v. Did you realise that the reading for v must be made exactly where the balloon was placed?

Additional details

It is also helpful to give additional details. In particular, make sure you suggest anything that needs to be done to ensure there is a large change in the dependent variable.

In the experiment with the balloon, you need a large change in θ as v changes. The readings would not be useful if θ was always very close to one value, for example 90°. How can this be achieved?

Obviously the largest air speed must be strong enough to cause a significant deflection. If the deflection is too small, then the mass under the balloon can be decreased; if it is too large, then the mass can be increased. It might be sensible to have the air speed as large as possible and adjust the mass under the balloon until θ is about 70°, and then check that θ varies from 70° to 90° as the fan is slowly moved further away. Of course, the mass under the balloon is then kept constant.

You might also think of any difficulties in carrying out the experiment. For example, draughts must be avoided and you must wait until the balloon has stopped swinging before taking a reading of θ.

Safety

It may seem strange, but you should always comment on safety when asked to carry out any experiment. In some situations the risks may be unimportant, and it may be sufficient to mention simple ideas such as wearing goggles to protect the eyes when heating liquids or handling stretched wires, using a safety screen, ensuring that the apparatus is stable and not easily knocked over, using a sand tray under heavy weights and making sure that weights do not fall on your foot, and switching off currents when not in use so that wires do not overheat.

In our example with the balloon, keeping away from the rotating blades in the fan and wearing goggles to avoid air blowing into your eye should be sufficient. Do give some detail in your suggestions and do not just say 'use goggles'.

WORKED EXAMPLE

1 Plan an experiment to measure the resistivity ρ of glass, which is about $10^{10}\,\Omega\,m$. You have available a number of sheets of glass of the same size but different thicknesses. Resistivity ρ is defined as $\rho = \frac{RA}{l}$.

Step 1 Identify the variables.

- The independent variable is the thickness l of the glass.
- The dependent variable is the resistance R of the glass. Finding R involves measuring the p.d. across the glass and the current in the glass.
- The control variable is the area of the glass. Since this is mentioned in the question, suggest also that the temperature must be constant.

Step 2 Describe the method of data collection in logical steps.

To alter the independent variable, use glass sheets of different thickness but the same area. The thickness of each piece of glass is measured with a micrometer at several places and averaged.

The area A is required. This can be found by measuring the length and breadth of each sheet with a rule and multiplying the values together.

Draw a circuit diagram of an ammeter in series with the glass sheet and a power supply, with a voltmeter across the glass. Connections are made to the large surfaces of the glass. This can be done using aluminium foil, or metal plates as in a capacitor, which closely touch each large face of the glass sheet. Use a diagram to show how this is done.

The logical steps are then to record ammeter and voltmeter readings with one thickness of glass. Then repeat the readings with different thicknesses, suggesting sensible thicknesses of glass, perhaps every mm from 1 mm to 10 mm. If you are going to perform the experiment these thicknesses may be available, but if you are merely planning the experiment then you must suggest sensible values.

Step 3 Add any additional details. How can you obtain reasonable values? Think about the size and thickness of the glass to be used and whether you can detect a reasonable change in the dependent variable, the resistance. You might, for example, suggest using a sheet of glass 1 m^2 in area and 1 mm thick. Its resistance is then:

$R = \frac{\rho l}{A} = 10^{10} \times 0.001 = 10^{7}\,\Omega$

Can this be measured with ordinary laboratory apparatus? What voltages and what meters are suitable? A voltage of 10 V produces a current of 1 μA, which is measurable, but 100 V gives a current of 10 μA, which may be easier to measure but more dangerous. With glass of thickness between 1 and 10 mm the current will be 1–10 μA and so the ammeter should measure from 1 to 10 μA or up to 10 μA.

As you can see, this means that you need some idea of the size of quantities that can be measured. In this example you need to know what currents and voltages can be measured with ordinary laboratory equipment.

You may also give additional detail by describing how to attach the metal foil as contacts onto the large faces of the glass sheet with weights on top, or suggest that the glass be cleaned and dried.

WORKED EXAMPLE (continued)

Step 4 State any safety points. Glass can cut a person's skin and so gloves should be worn. If voltages above about 50 V are to be used, then use rubber gloves to avoid an electric shock or cover all exposed metal parts with insulation.

Step 5 Give your method of analysis. Remember, every derived quantity must be explained, so do not forget to state that for each thickness the voltage and current readings are used to find the resistance with the formula $R = \frac{V}{I}$.

Since $R = \frac{\rho l}{A}$, choose to plot a graph with R on the y-axis and l on the x-axis. The graph should be a straight line through the origin – a diagram may help here.

The gradient of the graph is $\frac{\rho}{A}$, so ρ = gradient × A.

QUESTIONS

2 What other graph can be plotted in the example above on resistivity and how is the gradient used to find ρ?

3 A nail is placed with its sharp end just touching a piece of wood. When a mass falls with a velocity v and hits the nail, it drives the nail into the wood. It is suggested that the depth d that the nail moves into the wood is related to v by the equation $d = kv^2$, where k is a constant.

a Suggest:

i the independent, dependent and control variables

ii how the velocity of the falling mass can be measured as it hits the nail

iii sensible values for d and how they may be achieved and measured

iv the graph to be plotted and what it shows if the relationship is true.

b Write a logical step-by-step method to test the relationship.

Analysis of the data

Whether you are dealing with data you have collected in an experiment, or data provided to you, you will need to analyse it. You need to describe how the data is used in order to reach a conclusion, and give details of any derived quantities that are calculated.

First look carefully at the quantities in the relationship you have suggested (or at the formula that may be suggested when you are given an experiment to carry out). In our example with the balloon, $\tan\theta$ is inversely proportional to v^2, which means that the formula is $\tan\theta = \frac{k}{v^2}$, where k is a constant.

If possible you should suggest plotting a graph which you know is a straight line if the equation is correct. In our example, since the equation for a straight line is $y = mx + c$, the y-axis of the graph should be $\tan\theta$ and the x-axis should be $\frac{1}{v^2}$.

You must clearly state:

- what is plotted on each axis of your graph
- that the relationship is valid if the graph gives a straight line through the origin.

You may prefer to draw a sketch graph to show what you mean, but always state clearly what type of graph you are going to use.

More complicated analysis of data

In Chapter P1, we saw how to interpret equations of the form $y = mx + c$ and how to use a straight-line graph to find the constants m and c. However, you also need to be able to deal with quantities related by equations of the form $y = ax^n$ and $y = ae^{kx}$. For these, you need to be able to use logarithms (logs).

There are two common types of logarithm (see Chapter 20). The first type is sometimes called a natural logarithm, or a logarithm to base e, and is written as ln. The second type is a logarithm to base 10 and is written as lg. The ln type is more useful when dealing with an exponential formula such as e^{kx} but, otherwise, either type may be used. Look closely at any question to see which type is used. Do not mix the different types together in the answer to one question.

The unit of a quantity involving logarithms is specified in an unusual way. For example, the natural logarithm of a quantity s measured in metres is written as ln (s / m) and not as ln (s) / m or ln (s) / ln (m). You can see that the unit is written inside the bracket with the quantity.

You need to be able to take logarithms of equations of the form $y = ax^n$ and $y = ae^{kx}$. (Recall that an equation remains balanced if the same operation is performed on each side.)

Consider the equation: $y = ax^n$
Taking logarithms of both sides gives:

$$\lg y = \lg a + n \lg x$$

$$\ln y = \ln a + n \ln x$$

Now consider the equation: $y = ae^{kx}$
Taking logarithms of both sides gives:

$$\ln y = \ln a + kx$$

(To obtain these results we have used the rules for logarithms set out on page 318.)

QUESTIONS

4 Calculate:

- **a** lg 10
- **b** ln 10
- **c** lg 100
- **d** lg 5
- **e** the antilogarithm to base 10 of 1 (i.e. find x where $\lg x = 1$)
- **f** the antilogarithm to base e of 0.5 (i.e. find x where $\ln x = 0.5$)

5 The number $48 = 3 \times 2^4$. Calculate lg 48 and lg 3 + 4 lg 2. Why are they the same?

Which graph to plot?

In handling data, our aim is usually to process the data to obtain a straight line graph. Then we can deduce quantities from the gradient and the intercepts. Table P2.1 shows graphs which can be plotted for different relationships, and the quantities which can be deduced from the graphs.

A relationship of the form $y = ax^n$

A ball falls under gravity in the absence of air resistance. It falls a distance s in time t. The results are given in the first two columns of Table P2.2.

A graph of the distance fallen against time gives the curve shown in Figure P2.2. Because this is a curve, it tells us little about the relationship between the variables. If,

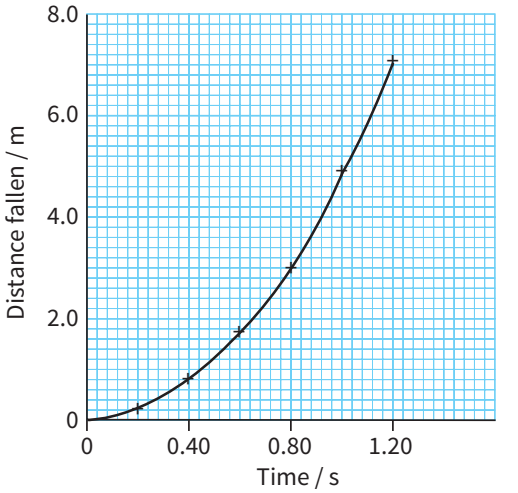


Figure P2.2 A distance–time graph plotted using the data in Table P2.2.

however, we suspect that the relationship is of the form $y = ax^n$, we can test this idea by plotting a graph of ln s against ln t (a 'log–log plot'). Table P2.2 shows the values for ln s and ln t, and the resulting graph is shown in Figure P2.3. (Notice that here we are using natural logs, but we could equally well use logs to base 10.)

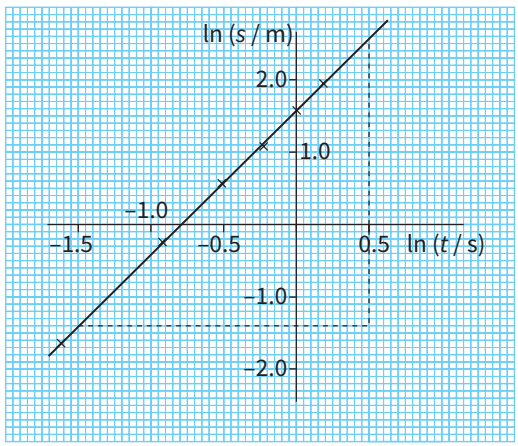


Figure P2.3 A log–log plot for the data shown in Table P2.2.

Relationship	Graph	Gradient	Intercept on y-axis	because ...
$y = mx + c$	y against x	m	c	
$y = ax^n$	ln y against ln x lg y against lg x	n	ln a lg a	ln y = n ln x + ln a lg y = n lg x + lg a
$y = ae^{kx}$	ln y against x	k	ln a	ln y = kx + ln a

Table P2.1 Choice of axes for straight-line graphs.

Time t / s	Distance fallen s / m	ln (t / s)	ln (s / m)
0.20	0.20	−1.61	−1.61
0.40	0.78	−0.92	−0.25
0.60	1.76	−0.51	0.57
0.80	3.14	−0.22	1.14
1.00	4.90	0.00	1.59
1.20	7.05	0.18	1.95

Table P2.2 Results for a ball falling under gravity.

From this graph the gradient is equal to the value of n, the power of t:

$$n = \text{gradient} = \frac{(2.55 - (-1.4))}{(0.5 - (-1.5))}$$
$$= \frac{3.95}{2.0}$$
$$= 1.98 \approx 2.0$$

So the equation is of the form $s = at^2$. The intercept on the y-axis is equal to $\ln a$, so:

$$\ln a = 1.6$$

By taking the antilogarithm we get:

$$a = 4.95\,\text{m s}^{-2} \approx 5.0\,\text{m s}^{-2}$$

If we think of the equation for free fall $s = \frac{1}{2}gt^2$, the constant $a = \frac{1}{2}g$. But $g = 9.8\,\text{m s}^{-2}$, which is consistent with the value we get for our constant.

A relationship of the form $y = ae^{kx}$

A current flows from a charged capacitor when it is connected in a circuit with a resistor. The current decreases exponentially with time (the same pattern we see in radioactive decay).

Figure P2.4 shows the circuit and Table P2.3 shows typical values of current I and time t from such an experiment.

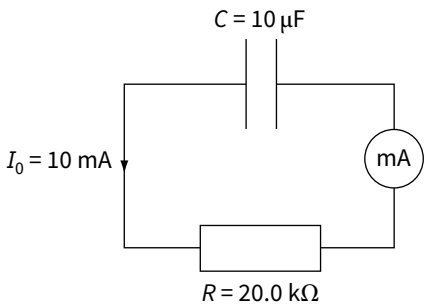


Figure P2.4 A circuit for investigating the discharge of a capacitor.

Current I / mA	Time t / s	ln (I / mA)
10.00	0.00	2.303
6.70	0.20	1.902
4.49	0.40	1.502
3.01	0.60	1.102
2.02	0.80	0.703
1.35	1.00	0.300

Table P2.3 Results from a capacitor discharge experiment.

The graph obtained from these results (Figure P2.5) shows a typical decay curve, but we cannot be sure that it is exponential. To show that the curve is of the form $I = I_0e^{kt}$ we plot $\ln I$ against t (a 'log-linear plot'). Values of $\ln I$ are included in Table P2.3. (Here, we must use logs to base e rather than to base 10.)

The graph of $\ln I$ against t is a straight line (Figure P2.6), confirming that the decrease in current follows an exponential pattern. The negative gradient shows exponential decay, rather than growth.

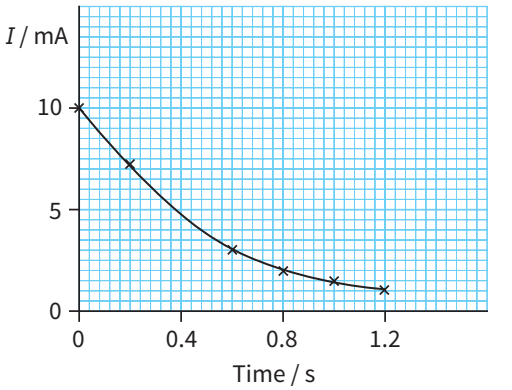


Figure P2.5

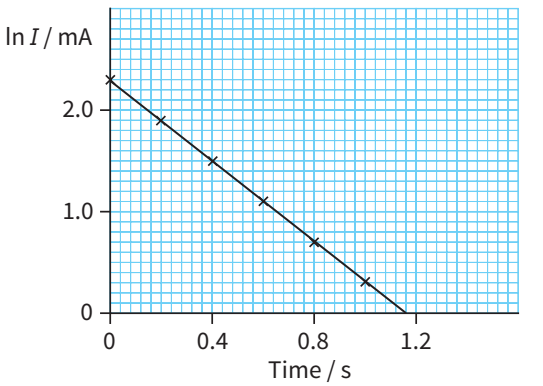


Figure P2.6

The gradient of the graph gives us the value of the constant k:

$$k = \text{gradient} = \frac{(0 - 2.30)}{(1.16 - 0)} = -1.98\,\text{s}^{-1} \approx -2.0\,\text{s}^{-1}$$

From the graph, we can also see that the intercept on the y-axis has the value 2.30 and hence (taking the antilogarithm) we have $I_0 = 9.97 \approx 10\,\text{mA}$. Hence we can write an equation to represent the decreasing current as follows:

$$I = 10\,e^{-2.0t}$$

We could use this equation to calculate the current at any time t.

QUESTIONS

6 In the expressions that follow, x and y are variables in an experiment. All the other quantities in the expressions are constants.

In each case, state the graph you would plot to produce a straight line. Give the gradient of each line in terms of the constants in the expression.

a $y = kx^{3/2}$

b $y = cx^{q}$

c $m = \frac{8x}{By^2}$

d $y = y_0 e^{kx}$

e $R = \frac{(y - y_0)}{x^2}$

7 The period of oscillation T of a small spherical mass supported by a length l of thread is given by the expression:

$$T = 2\pi\sqrt{\frac{l}{g}}$$

where g is the acceleration due to gravity.

Design a laboratory experiment using this expression to determine the acceleration due to gravity. You should draw a diagram showing the arrangement of your equipment. In your account, you should pay particular attention to:

a the procedure to be followed

b the measurements to be taken

c analysis of the data to determine g

d any safety precautions that you would take.

Treatment of uncertainties

All results should include an estimate of the absolute uncertainty. For example, when measuring the time for a runner to complete the 100 m you may express this as 12.1 ± 0.2 s. This can also be expressed as a percentage uncertainty (see Chapter P1); the percentage uncertainty is equal to $\frac{0.2}{12.1} \times 100\% = 1.65\%$, so we write the value as 12.1 s ± 1.7%, or even 12.1 s ± 2%.

Combining uncertainties

When quantities are combined, what is the uncertainty in the result?

Suppose that quantity $A = 1.0 \pm 0.1$ and that $B = 2.0 \pm 0.2$, so that the value of $A + B$ is 3.0. The maximum likely value of $A + B$, taking into account the uncertainties, is 3.3, and the minimum likely value is 2.7. You can see that the combined uncertainty is ±0.3, so $A + B = 3.0 \pm 0.3$. Similarly $B - A = 1.0 \pm 0.3$.

When quantities are added or subtracted, their absolute uncertainties are added. A simple example is measuring the length of a stick using a millimetre scale. There is likely to be an uncertainty of 0.5 mm at both ends, giving a total uncertainty of 1.0 mm.

When quantities are multiplied or divided, combining uncertainties is a little more complex. To find the combined uncertainty in this case, we add the percentage uncertainties of the two quantities to find the total percentage uncertainty.

Remember – you always add uncertainties; never subtract. Where quantities are:

- added or subtracted, then add **absolute** uncertainties
- multiplied or divided, then add **percentage** uncertainties.

WORKED EXAMPLE

2 The potential difference across a resistor is measured as 6.0 ± 0.2 V, while the current is measured as 2.4 ± 0.1 A.

Calculate the resistance of the resistor and the absolute uncertainty in its measurement.

Step 1 Find the percentage uncertainty in each of the quantities:

$$\text{percentage uncertainty in p.d.} = \frac{0.2}{6.0} \times 100\% = 3.3\%$$

$$\text{percentage uncertainty in current} = \frac{0.1}{2.4} \times 100\% = 4.2\%$$

Step 2 Add the percentage uncertainties.

sum of uncertainties = (3.3 + 4.2)% = 7.5%

Step 3 Calculate the resistance value and find the absolute uncertainty

$$R = \frac{V}{I} = \frac{6.0}{2.4} = 2.5\,\Omega$$

7.5% of 2.5 = 0.1875 ≈ 0.2 Ω

The resistance of the resistor is 2.5 ± 0.2 Ω.

When you calculate the uncertainty in the square of a quantity, since this is an example of multiplication, you should double the percentage uncertainty. For example, if $A = 2.0 \pm 0.2$ cm, then A has a percentage uncertainty of 10% so $A^2 = 4.0\,\text{m}^2 \pm 20\%$; or giving the absolute uncertainty, $A^2 = 4.0 \pm 0.8\,\text{cm}^2$.

QUESTIONS

8 You measure the following quantities:

$A = 1.0 \pm 0.4\,\text{m}$ $B = 2.0 \pm 0.2\,\text{m}$

$C = 2.0 \pm 0.5\,\text{m s}^{-1}$ $D = 0.20 \pm 0.01\,\text{s}$

Calculate the result and its uncertainty for each of the following expressions. You may express your uncertainty either as an absolute value or as a percentage.

a $A + B$

b $B - A$

c $C \times D$

d $\frac{B}{D}$

e A^2

f $2 \times A$

g the square root of $(A \times B)$ (For part **g**, you should recall that the square root of x can be written as $x^{\frac{1}{2}}$.)

9 A rifle bullet is photographed in flight using two flashes of light separated by a time interval of 1.00 ± 0.02 ms. The first image of the bullet on the photograph appears to be at a position of 22.5 ± 0.5 cm on a scale underneath the flight path. The position of the second image is 37.5 ± 0.7 cm on the same scale. Find the speed of the bullet and its absolute uncertainty.

Uncertainties and graphs

We can use error bars to show uncertainties on graphs. Table P2.4 shows results for an experiment on stretching a spring.

Load / N	Length of spring / cm	Extension / cm
0	12.4 ± 0.2	0.0
1.00	14.0 ± 0.2	1.6 ± 0.4
2.00	15.8 ± 0.2	3.4 ± 0.4
3.00	17.6 ± 0.2	5.2 ± 0.4
4.00	18.8 ± 0.2	6.4 ± 0.4
5.00	20.4 ± 0.2	8.0 ± 0.4

Table P2.4 Results from an experiment on stretching a spring.

When plotting the graph, the points are plotted as usual, and then they are extended to show the maximum and minimum likely values, as shown in Figure P2.7. Then the best fit line is drawn.

To estimate the error in the gradient we draw not only the best fit line but also a 'worst acceptable' line, passing through the extremes in the error bars as shown in Figure P2.8.

The gradients for both best fit and worst fit lines are calculated and the error is the difference in their gradients:

error = (gradient of best fit line) – (gradient of worst acceptable line)

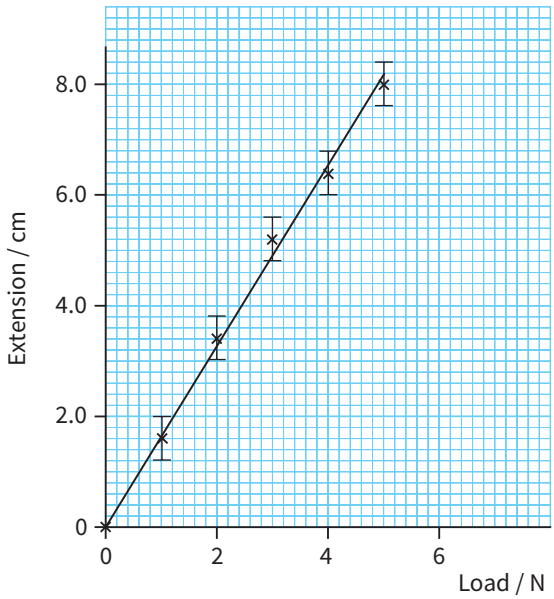


Figure P2.7 A graph representing the data in Table P2.4, with error bars and a line of best fit drawn.

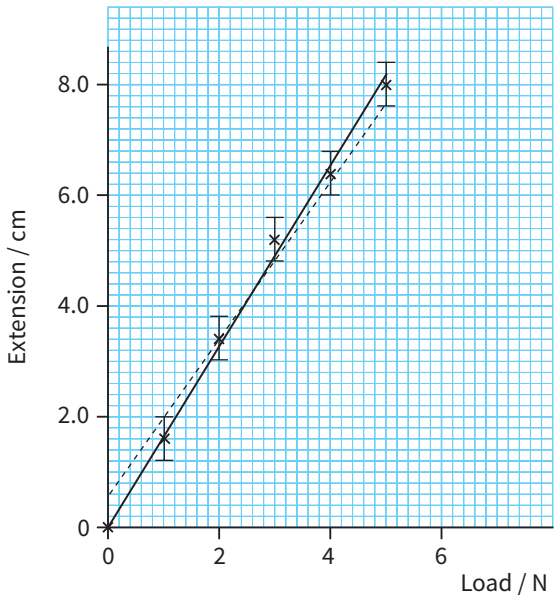


Figure P2.8 The same graph as in Figure P2.7, with a 'worst acceptable' line drawn (dashed).

In our experiment, the gradients are:

$$\text{line of best fit: gradient} = \left(\frac{8.2 - 0}{5.0 - 0}\right)\text{cm N}^{-1}$$

$$= 1.64\,\text{cm N}^{-1} \approx 1.6\,\text{cm N}^{-1}$$

$$\text{line of worst fit: gradient} = \left(\frac{7.6 - 0.6}{5.0 - 0}\right)\text{cm N}^{-1}$$

$$= 1.4\,\text{cm N}^{-1}$$

So the uncertainty in the gradient = $1.6 - 1.4 = \pm 0.2\,\text{cm N}^{-1}$

The gradient is therefore: $1.6 \pm 0.2\,\text{cm N}^{-1}$.

Uncertainties and logarithms

When a log graph is used and we need to include error bars, we must find the logarithm of the measured value and the logarithm of either the largest or the smallest likely value. The uncertainty in the logarithm will be the difference between the two.

QUESTIONS

10 The values of load shown in Table P2.4 are given without any indication of their uncertainties. Suggest a reason for this.

11 A student measures the radius r and the resistance R of several equal lengths of wire. The results are shown in Table P2.5. It is suggested that R and r are related by the equation:

$R = ar^b$ where a and b are constants.

a A graph is plotted with $\ln R$ on the y-axis and $\ln r$ on the x-axis. Express the gradient and y-intercept in terms of a and b.

b Values of r and R measured in an experiment are given in Table P2.5.

r / mm	R / Ω	$\ln r$ / mm	$\ln R$ / Ω
2.0 ± 0.1	175.0		
3.0 ± 0.1	77.8		
4.0 ± 0.1	43.8		
5.0 ± 0.1	28.0		
6.0 ± 0.1	19.4		

Table P2.5 Measurements for Question 11.

Copy and complete the table by calculating and recording values of $\ln(r/\text{mm})$ and $\ln(R/\Omega)$ and include the absolute uncertainties in $\ln(r/\text{mm})$.

c Plot a graph of $\ln(r/\text{mm})$ against $\ln(R/\Omega)$. Include error bars for $\ln(r/\text{mm})$.

d Draw the line of best fit and a worst acceptable straight line on your graph.

e Determine the gradient of the line of best fit. Include the uncertainty in your answer.

f Using your answer to **e**, determine the value of b.

g Determine the value of a and its uncertainty.

WORKED EXAMPLE

3 The resistance of a resistor is given as 47 ± 5 Ω. The value of $\ln(R/\Omega)$ is to be plotted on a graph. Calculate the value and uncertainty in $\ln(R/\Omega)$.

Step 1 Calculate the logarithm of the given value:

$\ln(R/\Omega) = \ln 47 = 3.85$

Step 2 Calculate the logarithm of the maximum likely value:

maximum likely value = 47 + 5 = 52 Ω

$\ln 52 = 3.95$

Step 3 The uncertainty is the difference between the two logarithms:

uncertainty in $\ln R = 3.95 - 3.85 = 0.10$

Thus $\ln(R/\Omega) = 3.85 \pm 0.10$

Conclusions and evaluation of results

In the previous experiment we can conclude that the extension/load for the spring in this example is $1.6 \pm 0.2\ \text{cm N}^{-1}$. If a hypothesis is made that the extension is proportional to the load then there is enough evidence here for the conclusion to be supported, as a straight line can be drawn from the origin through all the error bars. If this is not possible then the hypothesis is not validated.

Now, suppose that the hypothesis is that the spring obeys Hooke's law and stretches by 5.0 cm when a load of 2.5 N is applied. The first part is validated for the reasons given above. However, an extension of 5.0 cm for a load of 2.5 N gives a value of $2.0\ \text{cm N}^{-1}$ for the gradient. This is clearly outside the range allowed for by the uncertainty in our measurements, and therefore the hypothesis is not supported.

End-of-chapter questions

1 The volume of air inside a bottle affects its resonant frequency.

- **a** What are the dependent and independent variables? [1]
- **b** Suggest **one** quantity to be controlled. [1]
- **c** How would you produce sounds of different frequency to show that the bottle resonates? [1]
- **d** How would you find the frequency of the sound which makes the bottle resonate? [1]
- **e** How would you find the volume of air inside the bottle? [1]
- **f** How would you change the volume of air inside the bottle while keeping all other factors constant? [1]
- **g** Suggest a safety precaution involving sound. [1]

2 The terminal velocity of an air bubble that rises in water is affected by the size of the bubble.

- **a** What are the dependent and independent variables? [1]
- **b** Suggest a quantity to be controlled. [1]
- **c** How would you measure the terminal velocity of an air bubble that rises in water? [1]
- **d** How would you generate bubbles of air of different sizes in water? [1]

3 The count rate from a radioactive source emitting γ-radiation is inversely proportional to the square of the distance from the source. Sources emitting γ-radiation also emit α- and β-radiation and are roughly spherical with a diameter of 2 cm.

- **a** What are the dependent and independent variables? [1]
- **b** Suggest a quantity to be controlled. [1]
- **c** How could you make sure that only γ-radiation is detected? [2]
- **d** How would you measure the count rate? Draw a diagram of the apparatus and explain how it is used. [2]
- **e** How would you make the uncertainty in the count rate as small as possible? [1]
- **f** Suggest **one** difficulty in measuring the distance and how this difficulty may be reduced. [2]
- **g** Suggest a safety precaution. [1]

4 The size of a small toy balloon depends on atmospheric pressure.

- **a** What are the dependent and independent variables? [1]
- **b** Suggest a quantity to be controlled. [1]
- **c** Draw a diagram of an apparatus to investigate the change in size of the balloon as atmospheric pressure changes. [2]
- **d** State how the pressure is changed in your apparatus and how it is measured? [2]
- **e** Suggest a safety precaution. [1]

5 Quantities A and B have the following values: $A = 3.0 \pm 0.2$ cm, $B = 2.0 \pm 0.1$ cm. Find the value of the following expressions and their absolute uncertainties.

- **a** AB [1]
- **b** $\frac{A}{B}$ [1]
- **c** A^2 [1]
- **d** $A - B$ [1]
- **e** $A^2 - B^2$ [1]
- **f** $\sqrt{A}$ [1]

6 Explain how you draw the best fit line and the worst fit line on a graph and how you find the uncertainty in the intercept on the *y*-axis. [2]

Questions 7–11 ask you to design an experiment based on the information given. All these questions have the same marking structure, with marks allocated to the different aspects as shown below.

a *the procedure to be followed* *[2]*

b *the measurements to be taken* *[4]*

c *the analysis of data* *[3]*

d *the safety precautions to be taken* *[1]*

e *additional detail.* *[5]*

7 The resistance R of a light-dependent resistor (LDR) varies with the distance d from a very bright source of light. It is suggested that R and d are related by the formula $R = kd^n$, where k and n are constants. Design a laboratory experiment to test this relationship. The LDR has a resistance of 50 Ω in bright light and 200 kΩ in the dark.

8 A ruler with a small mass at one end is clamped at the other end, as shown in Figure P2.9, and oscillates up and down when plucked by hand.

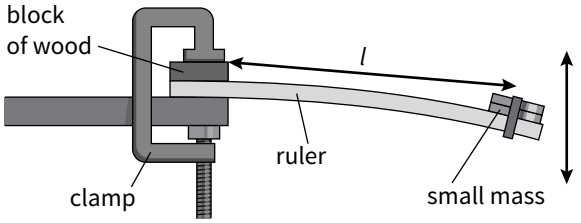


Figure P2.9 For End-of-chapter Question 8.

It is suggested that the period of oscillation T of the ruler is related to the length l by the relationship $T = kl^n$, where k and n are constants. Design a laboratory experiment to test this relationship and to find the value of n.

9 A current-carrying coil produces a magnetic field. It is suggested that the magnetic field strength B at the centre of the coil is proportional to the current I in the coil. Design a laboratory experiment that uses a Hall probe to test this relationship.

10 A bar magnet dropped into a coil induces an e.m.f. in the coil. It is suggested that E, the maximum induced e.m.f., is proportional to v, the speed of the magnet. Design a laboratory experiment to test this relationship. You might like to look at Figure 28.24 in Chapter 28.

11 A student has a number of different transformers of varying numbers of turns. An alternating input current to the transformer induces an output e.m.f. It is suggested that the output e.m.f. V_s is directly proportional to the frequency f of the applied current. Design a laboratory experiment to test this relationship.

12 The period T of a simple pendulum is related to its length l by the equation:

$$T = 2\pi\sqrt{\frac{l}{g}}$$

where g is the acceleration of free fall.

a A graph is plotted with T^2 on the y-axis and l on the x-axis. Express the gradient in terms of g. [1]

b A student measures the time t for 10 oscillations for different lengths l. Table P2.6 shows her data.

l / m	t / s	T	T^2
0.300	11.1 ± 0.1		
0.400	12.8 ± 0.1		
0.500	14.2 ± 0.1		
0.600	15.8 ± 0.1		
0.700	16.9 ± 0.1		
0.800	18.1 ± 0.1		

Table P2.6

i Calculate and record values of T and T^2, including the absolute uncertainties in T and T^2. [2]

ii Plot a graph of T^2/s^2 against l/cm including error bars for T^2. [2]

iii Draw a straight line of best fit and a worst acceptable line on your graph. [2]

iv Determine the gradient of your line and include the uncertainty in your answer. [2]

v Use your value of the gradient to determine g and include the absolute uncertainty in your value. [2]

vi Using your value of g and its uncertainty, calculate the value of t when the length l is 0.900 m. Include the absolute uncertainty in your answer. [2]

13 Readings are taken of the resistance R of a thermistor at different temperatures T. It is suggested that the relationship between R and T is $R = kT^n$, where k and n are constants.

a A graph is plotted with lg$R$ on the y-axis and lg$T$ on the x-axis. State the value of the gradient and the y-intercept in terms of k and n. [2]

b Values for T and R are shown in Table P2.7.

T / K	R / Ω	lg (T / K)	lg (R / Ω)
273	550 ± 10		
283	480 ± 10		
293	422 ± 10		
303	370 ± 10		
313	330 ± 10		

Table P2.7 For End-of-chapter Question 13.

Complete the table and include absolute uncertainties in lg (R / Ω).

c **i** Plot a graph of lg (R / Ω) against lg (T / K). Include error bars. [2]

ii Draw the line of best fit and a worst acceptable line on your graph. [2]

iii Determine the gradient of your line of best fit and the uncertainty in your value. [2]

iv Determine the y-intercept of your graph (this is where the x-value, in this case lg (T / K), is zero). Give the uncertainty in your value. [2]

v Determine values for n and k and the uncertainties in your answers. [3]

Appendix 1: Physical quantities and units

Physical quantities have a numerical value and a unit. In physics, it is essential to give the units of physical quantities. For example, mass can be measured in kilograms. Hence you might write the mass of the trolley as:

mass of trolley = 0.76 kg

It would be a serious error to omit the unit kg at the end of the numerical value.

The scientific system of units is called the Système Internationale d'Unités (or SI system). The seven **base units** of this system are listed in Table 1. Each of the units is carefully defined, but the definitions need not concern us here.

All other units can be derived from the seven base units. For example:

- volume is measured in cubic metres (m^3)
- velocity is measured in metres per second ($m\,s^{-1}$)
- density is measured in kilograms per cubic metre ($kg\,m^{-3}$).

Physical quantity	Unit
mass	kilogram, kg
length	metre, m
time	second, s
temperature	kelvin, K
electric current	ampere, A
amount of substance	mole, mol
luminous intensity	candela, cd

Table 1 The seven base units of the SI system. (Note that you are not required to use the candela in this book.)

Prefixes

In physics, you will have to cope with very small and very large numbers. Numbers are written using powers of 10 to make them less awkward. This is known as **scientific notation**. Prefixes are used as an abbreviation for some of the powers of 10. For example, the height of a 5400 m high mountain may be written as either 5.4×10^3 m or 5.4 km. The prefixes you will need most often are shown in Table 2.

Prefix	Symbol	Value
pico	p	10^{-12}
nano	n	10^{-9}
micro	μ	10^{-6}
milli	m	10^{-3}
centi	c	10^{-2}
deci	d	10^{-1}
kilo	k	10^{3}
mega	M	10^{6}
giga	G	10^{9}
tera	T	10^{12}

Table 2 Some of the prefixes used in the SI system.

Estimation

When you carry out an experiment or a calculation, it is sensible to look at the answer that you get (and the results of intermediate calculations) to see if they seem reasonable. The only way you can know if an answer is absurd is if you are aware of some benchmarks. Some suggestions are given below. Try to add to this list as you go through your physics course.

mass of a person	70 kg
height of a person	1.5 m
walking speed	$1\,m\,s^{-1}$
speed of a car on the motorway	$30\,m\,s^{-1}$
volume of a can of drink	$300\,cm^3$
density of water	$1000\,kg\,m^{-3}$
weight of an apple	1 N
typical current in domestic appliance	13 A
e.m.f. of a car battery	12 V

Appendix 2:
Data, formulae and relationships

Data

Values are given here to three significant figures, except where more significant figures are useful.

speed of light in free space	c	$3.00 \times 10^{8}\,\text{m s}^{-1}$
permeability of free space	μ_0	$4\pi \times 10^{-7}\,\text{H m}^{-1}$
permittivity of free space	ε_0	$8.85 \times 10^{-12}\,\text{C}^2\,\text{N}^{-1}\,\text{m}^{-2}\,(\text{F m}^{-1})$
elementary charge	e	$1.60 \times 10^{-19}\,\text{C}$
Planck constant	h	$6.63 \times 10^{-34}\,\text{J s}$
rest mass of electron	m_e	$9.11 \times 10^{-31}\,\text{kg}$
rest mass of proton	m_p	$1.67 \times 10^{-27}\,\text{kg}$
rest mass of neutron	m_n	$1.675 \times 10^{-27}\,\text{kg}$
rest mass of α-particle	m_α	$6.646 \times 10^{-27}\,\text{kg}$
molar gas constant	R	$8.31\,\text{J K}^{-1}\,\text{mol}^{-1}$
Avogadro constant	N_A	$6.02 \times 10^{23}\,\text{mol}^{-1}$
Boltzmann constant	k	$1.38 \times 10^{-23}\,\text{J K}^{-1}$
gravitational constant	G	$6.67 \times 10^{-11}\,\text{N m}^2\,\text{kg}^{-2}$
acceleration of free fall*	g	$9.81\,\text{m s}^{-2}$

*Note that this is the value of g which you should use in answering questions; g varies significantly over the Earth's surface, with values ranging from $9.78\,\text{m s}^{-2}$ at the equator to $9.83\,\text{m s}^{-2}$ at the poles.

Conversion factors

unified atomic mass unit	$1\,\text{u} = 1.661 \times 10^{-27}\,\text{kg}$
electronvolt	$1\,\text{eV} = 1.60 \times 10^{-19}\,\text{J}$
	$1\text{ day} = 8.64 \times 10^{4}\,\text{s}$
	$1\text{ year} \approx 3.16 \times 10^{7}\,\text{s}$
	$1\text{ light year} \approx 9.5 \times 10^{15}\,\text{m}$

Mathematical equations

arc length = $r\theta$

circumference of circle = $2\pi r$

area of circle = πr^2

curved surface area of cylinder = $2\pi rh$

volume of cylinder = $\pi r^2 h$

surface area of a sphere = $4\pi r^2$

volume of sphere = $\frac{4}{3}\pi r^3$

Pythagoras' theorem: $a^2 = b^2 + c^2$

cosine rule: $a^2 = b^2 + c^2 - 2bc\cos A$

sine rule: $\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}$

for small angle θ: $\sin\theta \approx \tan\theta \approx \theta$ and $\cos\theta \approx 1$

$\lg(AB) = \lg(A) + \lg(B)$

$\lg\left(\frac{A}{B}\right) = \lg(A) - \lg(B)$

$\ln(x^n) = n\ln(x)$

$\ln(e^{kx}) = kx$

Formulae and relationships

uniformly accelerated motion	$v = u + at$
	$s = \frac{1}{2}(u+v)t$
	$s = ut + \frac{1}{2}at^2$
	$v^2 = u^2 + 2as$
work done on or by a gas	$W = p\Delta V$
gravitational potential	$\varphi = \frac{-Gm}{r}$
hydrostatic pressure	$p = \rho gh$
pressure of an ideal gas	$p = \frac{1}{3}\frac{Nm}{V}\langle c^2\rangle$
simple harmonic motion	$a = -\omega^2 x$
velocity of a particle in s.h.m.	$v = v_0\cos\omega t$
	$v = \pm\omega\sqrt{(x_0^2 - x^2)}$
Doppler effect	$f_0 = \frac{f_s v}{v \pm v_s}$
electric potential	$V = \frac{Q}{4\pi\varepsilon_0 r}$
capacitors in series	$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \ldots$
capacitors in parallel	$C = C_1 + C_2 + \ldots$
energy of charged capacitor	$W = \frac{1}{2}QV$
electric current	$I = nAvq$
resistors in series	$R = R_1 + R_2 + \ldots$
resistors in parallel	$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \ldots$
Hall voltage	$V_H = \frac{BI}{ntq}$
alternating current or voltage	$x = x_0\sin\omega t$
radioactive decay	$x = x_0\exp(-\lambda t)$ or $x = x_0 e^{-\lambda t}$
decay constant	$\lambda = \frac{0.693}{t_{\frac{1}{2}}}$

Appendix 3
The Periodic Table

Group																	
1	2											13	14	15	16	17	18
1 **H** hydrogen 1.0																	2 **He** helium 4.0
3 **Li** lithium 6.9	4 **Be** beryllium 9.0											5 **B** boron 10.8	6 **C** carbon 12.0	7 **N** nitrogen 14.0	8 **O** oxygen 16.0	9 **F** fluorine 19.0	10 **Ne** neon 20.2
11 **Na** sodium 23.0	12 **Mg** magnesium 24.3	3	4	5	6	7	8	9	10	11	12	13 **Al** aluminium 27.0	14 **Si** silicon 28.1	15 **P** phosphorus 31.0	16 **S** sulphur 32.1	17 **Cl** chlorine 35.5	18 **Ar** argon 39.9
19 **K** potassium 39.1	20 **Ca** calcium 40.1	21 **Sc** scandium 45.0	22 **Ti** titanium 47.5	23 **V** vanadium 50.9	24 **Cr** chromium 52.0	25 **Mn** manganese 54.9	26 **Fe** iron 55.8	27 **Co** cobalt 58.9	28 **Ni** nickel 58.7	29 **Cu** copper 63.5	30 **Zn** zinc 65.4	31 **Ga** gallium 69.7	32 **Ge** germanium 72.6	33 **As** arsenic 74.9	34 **Se** selenium 79.0	35 **Br** bromine 79.9	36 **Kr** krypton 83.8
37 **Rb** rubidium 85.5	38 **Sr** strontium 87.6	39 **Y** yttrium 88.9	40 **Zr** zirconium 91.2	41 **Nb** niobium 92.9	42 **Mo** molybdenum 95.9	43 **Tc** technetium –	44 **Ru** ruthenium 101.1	45 **Rh** rhodium 102.9	46 **Pd** palladium 106.4	47 **Ag** silver 107.9	48 **Cd** cadmium 112.4	49 **In** indium 114.8	50 **Sn** tin 118.7	51 **Sb** antimony 121.8	52 **Te** tellurium 127.6	53 **I** iodine 126.9	54 **Xe** xenon 131.3
55 **Cs** caesium 132.9	56 **Ba** barium 137.3	57–71 **La** lanthanum	72 **Hf** hafnium 178.5	73 **Ta** tantalum 180.9	74 **W** tungsten 183.8	75 **Re** rhenium 186.2	76 **Os** osmium 190.2	77 **Ir** iridium 192.2	78 **Pt** platinum 195.1	79 **Au** gold 197.0	80 **Hg** mercury 200.6	81 **Tl** thallium 204.4	82 **Pb** lead 207.2	83 **Bi** bismuth 209.0	84 **Po** polonium –	85 **At** astatine –	86 **Rn** radon –
87 **Fr** francium –	88 **Ra** radium –	89–103 actinoids	104 **Rf** rutherfordium –	105 **Db** dubnium –	106 **Sg** Seaborgium –	107 **Bh** bohrium –	108 **Hs** hassium –	109 **Mt** meitnerium –	110 **Ds** darmstadtium –	111 **Rg** roentgenium –	112 **Cn** copernicium –		114 **Fl** flerovium –		116 **Lv** livermorium –		

Key

atomic number
atomic symbol
name
relative atomic mass

lanthanoids	57 **La** lanthanum 140	58 **Ce** cerium 140.1	59 **Pr** praseodymium 140.9	60 **Nd** neodymium 144.4	61 **Pm** promethium –	62 **Sm** samarium 150.4	63 **Eu** europium 152.0	64 **Gd** gadolinium 157.3	65 **Tb** terbium 158.9	66 **Dy** dysprosium 162.5	67 **Ho** holmium 164.9	68 **Er** erbium 167.3	69 **Tm** thulium 168.9	70 **Yb** ytterbium 173.1	71 **Lu** lutetium 175.0
actinoids	89 **Ac** actinium –	90 **Th** thorium 232.0	91 **Pa** protactinium 231.0	92 **U** uranium 238.0	93 **Np** neptunium –	94 **Pu** plutonium –	95 **Am** americium –	96 **Cm** curium –	97 **Bk** berkelium –	98 **Cf** californium –	99 **Es** einsteinium –	100 **Fm** fermium –	101 **Md** mendelevium –	102 **No** nobelium –	103 **Lr** lawrencium –

Glossary

absolute scale of temperature; *see* thermodynamic scale.

absolute zero The temperature at which a system has minimum internal energy; equivalent to –273.15 °C.

absorption line spectrum A dark line of a unique wavelength seen in a continuous spectrum.

acceleration The rate of change of an object's velocity:

$$a = \frac{\Delta v}{\Delta t}$$

Unit: $m\,s^{-2}$.

accuracy An accurate value of a measured quantity is one which is close to the true value of the quantity.

acoustic impedance Acoustic impedance Z is the product of the density ρ of a substance and the speed c of sound in that substance ($Z = \rho c$). Unit: $kg\,m^{-2}\,s^{-1}$.

activity The rate of decay or disintegration of nuclei in a radioactive sample.

ampere The SI unit of electric current (abbreviated A).

amplitude The maximum displacement of a particle from its equilibrium position.

amplitude modulation A form of modulation in which the signal causes variations in the amplitude of a carrier wave.

analogue signal A signal that is continuously variable, having a continuum of possible values.

analogue-to-digital conversion (ADC) Conversion of a continuous analogue signal to discrete digital numbers.

angular displacement The angle through which an object moves in a circle.

angular frequency The frequency of a sinusoidal oscillation expressed in radians per second:

$$\text{angular frequency } \omega = \frac{2\pi}{T}$$

angular velocity The rate of change of the angular position of an object as it moves along a curved path.

antinode A point on a stationary wave with maximum amplitude.

atomic mass unit A unit of mass (symbol u) approximately equal to 1.661×10^{-27} kg. The mass of an atom of $^{12}_{6}C$ = 12.000 u exactly.

attenuation The gradual loss in strength or intensity of a signal.

average speed The total distance travelled by an object divided by the total time taken.

Avogadro constant The number of particles in one mole of any substance (approximately 6.02×10^{23} mol^{-1}), denoted N_A.

band theory The idea that electrons in a solid or liquid can have energies within certain ranges or bands, between which are forbidden values.

bandwidth (communications) A measure of the width of a range of frequencies being transmitted.

base units Defined units of the SI system from which all other units are derived.

best fit line A straight line drawn as closely as possible to the points of a graph so that similar numbers of points lie above and below the line.

binding energy The minimum external energy required to separate all the neutrons and protons of a nucleus.

bit A basic unit of information storage, the amount of information stored by a device that exists in only two distinct states, usually given as the binary digits 0 and 1.

Boltzmann constant A fundamental constant given by $k = \frac{R}{N_A}$, where R is the ideal gas constant and N_A is the Avogadro constant.

Boyle's law The pressure exerted by a fixed mass of gas is inversely proportional to its volume, provided the temperature of the gas remains constant.

braking radiation X-rays produced when electrons are decelerated (also called Bremsstrahlung radiation).

capacitance The ratio of charge stored by a capacitor to the potential difference across it.

carrier wave A waveform (usually sinusoidal) which is modulated by an input signal to carry information.

centre of gravity The point where the entire weight of an object appears to act.

centripetal force The resultant force acting on an object moving in a circle; it is always directed towards the centre of the circle.

characteristic radiation Very intense X-rays produced in an X-ray tube, having specific wavelengths that depend on the target metal.

charge carrier Any charged particle, such as an electron, responsible for a current.

Charles's law The volume occupied by a gas at constant pressure is directly proportional to its thermodynamic (absolute) temperature.

closed system A system of interacting objects in which there are no external forces.

coaxial cable An electrical cable with an inner conductor surrounded by a tubular insulating layer and an outside conducting layer.

coherent Two sources are coherent when they emit waves with a constant phase difference.

collimated beam A parallel-sided beam of radiation.

components (of a vector) The magnitudes of a vector quantity in two perpendicular directions.

compression A region in a sound wave where the air pressure is greater than its mean value.

compressive Describes a force that squeezes an object.

computerised axial tomography A technique in which X-rays are used to image the human body in order to produce a computerised 3-D image.

conduction band A range of electron energies in a solid; electrons in the conduction band are free to move throughout the material.

conservation of momentum In a closed system, when bodies interact, the total momentum in any specified direction remains constant.

constructive interference When two waves reinforce to give increased amplitude.

contact force The force an object exerts on another with which it is in contact.

contrast In a high-contrast image, there is a big difference in brightness between bright and dark areas.

contrast media Materials such as barium that easily absorb X-rays. A contrast medium is used to reveal the outlines or edges of soft tissues in an X-ray image.

coulomb The SI unit of electrical charge (abbreviated C). A charge of 1 C passes a point when a current of 1 A flows for 1 s. 1 C = 1 A s.

Coulomb's law Any two point charges exert an electrical force on each other that is proportional to the product of their charges and inversely proportional to the square of the distance between them.

count rate The number of particles (beta or alpha) or gamma-ray photons detected per unit time by a Geiger–Müller tube. Count rate is always a fraction of the activity of a sample.

couple A pair of equal and antiparallel forces having a turning effect but no resultant force.

damped Describes an oscillatory motion in which the amplitude decreases with time due to energy losses.

de Broglie wavelength The wavelength associated with a moving particle, given by the equation:

$$\lambda = \frac{h}{mv}$$

decay constant The constant λ for an isotope that appears in the equation $A = -\lambda N$. It is the probability of an individual nucleus decaying per unit time interval.

decibel A logarithmic unit of measurement that expresses the relative sizes of two powers using the formula $10 \lg(\frac{P_1}{P_2})$.

density The mass per unit volume of a material:

$$\rho = \frac{m}{V}$$

Unit: kg m^{-3}.

dependent variable The variable in an experiment with a value that changes as the independent variable is altered by the experimenter.

derived units Units which are combinations of the base units of the SI system.

destructive interference When two waves cancel to give reduced amplitude.

diffraction The spreading of a wave when it passes through a gap or past the edge of an object.

digital signal A signal that has only a few possible values, often only two.

digital-to-analogue conversion (DAC) Conversion of a series of digital numbers into a continuous analogue signal.

dispersion The splitting of light into its different wavelengths.

displacement The distance moved by an object in a particular direction (measured from a fixed starting point).

Doppler effect The change in frequency or wavelength of a wave observed when the source of the wave is moving towards or away from the observer (or the observer is moving relative to the source).

drag A force that resists the movement of a body through a fluid.

drift velocity, mean The average speed of a collection of charged particles when a current flows.

dynamics The study of motion using quantities such as force and mass.

e.m.f. The total work done when unit charge is moved round a complete circuit. Unit: J C^{-1} or volt (V).

efficiency The ratio of useful output energy to the total input energy for a device, expressed as a percentage:

$$\text{efficiency} = \frac{\text{useful output energy}}{\text{total input energy}} \times 100\%$$

Einstein relation This refers to the equation for the energy of a photon:

$$E = hf \quad \text{or} \quad E = \frac{hc}{\lambda}$$

elastic limit The value of stress beyond which an object will not return to its original dimensions.

elastic potential energy Energy stored in a stretched or compressed material.

electric charge A property of a body that gives rise to a force on the body when it is within an electric field.

electric field A region in which a charged body experiences a force.

electric field strength The force per unit positive charge at a point. Unit: V m^{-1} or N C^{-1}.

electric potential The energy per unit charge due to a charged body's position in an electric field.

electrical resistance The ratio of potential difference to current. Unit: ohm (Ω).

electrolyte An electrically conducting solution. The conduction is due to positive and negative ions in the solution.

electromagnetic spectrum The family of waves that travel through a vacuum at a speed of $3.00 \times 10^8\ \text{m s}^{-1}$.

electronvolt The energy gained by an electron travelling through a p.d. of 1 volt. $1\ \text{eV} = 1.60 \times 10^{-19}\ \text{J}$.

elementary charge The smallest unit of charge that a particle or an object can have. It has a magnitude of $1.60 \times 10^{-19}\ \text{C}$.

emission line spectrum A sharp and bright line of a unique wavelength seen in a spectrum.

energy A calculated quantity which is conserved during any change; that which is transferred when a force does work.

energy band A range of permitted electron energies in a solid.

energy level A quantised energy state of an electron in an atom.

equation of state for an ideal gas:

$pV = nRT$ or $pV = NkT$

(Also known as the ideal gas equation.)

equations of motion Four interrelated equations that can be used to determine the displacement, initial velocity, final velocity and acceleration of a body moving with constant acceleration.

equilibrium An object in equilibrium is either at rest or travelling with a constant velocity because the resultant force on it is zero.

errors Inaccuracies when taking measurements.

evaporation The process by which a liquid becomes a gas at a temperature below its boiling point.

exponential decay Describes the decrease of a quantity where the rate of decrease is proportional to the value of the quantity.

extension The change in the length of a material from its original length.

farad The unit of capacitance (abbreviated F). $1\ \text{F} = 1\ \text{C V}^{-1}$.

Faraday's law of electromagnetic induction The induced e.m.f. is proportional to the rate of change of magnetic flux linkage.

field lines Lines drawn to represent the strength and direction of a field of force.

field of force A region of space where an object feels a force; the force may be gravitational, electric, magnetic, etc.

First law of thermodynamics The increase in internal energy of a body is equal to the thermal energy transferred to it by heating plus the mechanical work done on it.

Fleming's left-hand (motor) rule This rule is used to predict the force experienced by a current-carrying conductor placed in an external magnetic field: thumb → motion, first finger → magnetic field and second finger → conventional current.

Fleming's right-hand (generator) rule This rule is used to predict the direction of the induced current or e.m.f. in a conductor moved at right angles to a magnetic field: thumb → motion, first finger → magnetic field and second finger → induced conventional current.

forbidden gap A range of energy values which an electron in a solid cannot have.

force constant The ratio of force to extension for a spring or a wire. Unit: N m^{-1}.

forced oscillation An oscillation caused by an external driving force; the frequency is determined by the driving force, and is not the natural frequency of the oscillator.

free-body force diagram A diagram showing all the forces acting on an object (but not the forces it exerts on other objects).

free oscillation An oscillation at is the natural frequency of the oscillator.

frequency The number of oscillations per unit time. Unit: hertz (Hz).

frequency modulation A form of modulation in which the signal causes variations in the frequency of a carrier wave.

fundamental frequency The lowest-frequency stationary wave for a particular system.

gain The voltage gain of an amplifier is the ratio of the output voltage to the input voltage.

geostationary orbit The orbit of an artificial satellite which has a period equal to one day so that the satellite remains above the same point on the Earth's equator. From Earth the satellite appears to be stationary.

gravitational field A region where any object with mass experiences a force.

gravitational field strength The gravitational force experienced by an object per unit mass:

$$g = \frac{F}{m}$$

gravitational potential The gravitational potential energy per unit mass at a point in a gravitational field.

gravitational potential energy The energy a body has due to its position in a gravitational field.

ground state The lowest energy state that can be occupied by an electron in an atom.

hadron Any particle which is affected by the strong nuclear force, made from two or three quarks or anti-quarks.

half-life The mean time taken for half the number of active nuclei in a radioactive sample to decay.

half-thickness The mean thickness of an absorbing material required to reduce the intensity of radiation by half.

Hall effect The production of a voltage across a conductor when a current flows through the conductor at right angles to a magnetic field.

Hall voltage The voltage produced across a conductor when a current flows through the conductor at right angles to a magnetic field; used in a Hall probe to measure B since $V_H \propto B$.

harmonic A wave of frequency n times the fundamental frequency, where n is an integer.

Hooke's law The extension produced in an object is proportional to the force producing it.

ideal gas A gas that behaves according to the equation $pV = nRT$ or $pV = NkT$.

ideal gas equation for an ideal gas:

$pV = nRT$ or $pV = NkT$

(Also known as the ideal gas equation.)

image intensifier A device used to change a low-intensity X-ray image into a bright visual image.

impedance matching The reduction in intensity of ultrasound reflected at the boundary between two substances, achieved when the two substances have similar acoustic impedances.

independent variable The variable in an experiment with a value that is altered by the experimenter.

inelastic A collision is inelastic when kinetic energy is not conserved; some is transferred to other forms such as heat. Momentum and total energy are always conserved.

inertia A measure of the mass of an object. A massive object has large inertia.

instantaneous speed The speed of an object measured over a very short period of time.

intensity The power transmitted normally through a surface per unit area:

$$\text{intensity} = \frac{\text{power}}{\text{cross-sectional area}}$$

Unit: W m^{-2}.

interference The formation of points of cancellation and reinforcement where two coherent waves pass through each other.

internal energy The sum of the random distribution of kinetic and potential energies of the atoms or molecules in a system.

internal resistance The resistance of an e.m.f. source. The internal resistance of a battery is due to the chemicals within it.

intrinsic semiconductor A pure substance whose resistivity is intermediate between that of a conductor and an insulator.

inverting amplifier A circuit, involving the use of an amplifier, where the output is 180° out of phase with the input.

ion An atom with a net positive or negative charge.

isotopes Nuclei of the same element with a different number of neutrons but the same number of protons.

***I–V* characteristic** A graph of current against voltage for a particular component of an electrical circuit.

kinematics The study of motion using quantities such as time, distance, displacement, speed, velocity and acceleration.

kinetic energy Energy of an object due to its motion.

kinetic theory of gases A model based on the microscopic motion of atoms or molecules of a gas.

Kirchhoff's first law The sum of the currents entering any point (or junction) in a circuit is equal to the sum of the currents leaving that same point. This law represents the conservation of charge.

Kirchhoff's second law The sum of the e.m.f.s round a closed loop in a circuit is equal to the sum of the p.d.s in that same loop.

Larmor frequency The frequency of precession of a nucleus in an external magnetic field.

Lenz's law An induced current or e.m.f. is in a direction so as to produce effects which oppose the change producing it.

lepton A sub-atomic particle which is not affected by the strong nuclear force.

light-dependent resistor (LDR) A resistor whose resistance decreases as the intensity of light falling on it increases.

light-emitting diode (LED) A semiconductor component that emits light when it conducts electricity.

linear momentum The product of an object's mass and its velocity, $p = mv$. Momentum is a vector quantity.

longitudinal wave A wave in which the particles of the medium oscillate along the direction in which the wave travels.

lost volts The difference between the e.m.f. and the terminal p.d. in a circuit. It is equal to the voltage across the internal resistance.

magnetic field A force field in which a magnet, a wire carrying a current, or a moving charge experiences a force.

magnetic flux The product of magnetic flux density normal to a circuit and the cross-sectional area of the circuit. Unit: weber (Wb).

magnetic flux density The strength of a magnetic field. Magnetic flux density B is defined as:

$$B = \frac{F}{IL}$$

where F is the force experienced by a conductor in the magnetic field, I is the current in the conductor and L is the length of the conductor in the magnetic field. (The conductor is at right angles to the field.)

magnetic flux linkage The product of magnetic flux and the number of turns. Unit: weber (Wb).

magnetic resonance imaging (MRI) A medical imaging technique which uses nuclear magnetic resonance.

mass A measure of the amount of matter within an object. Unit: kilogram (kg).

mass defect The difference between the total mass of the individual, separate nucleons and the mass of the nucleus.

mass excess the difference between the mass of a nuclide (in u) and its mass number.

mean drift velocity The average speed of a collection of charged particles when a current flows.

microwave link A communications system that uses a beam of radio waves in the microwave frequency range to transmit audio, data or video information.

modulation The process of using one waveform to alter the frequency, amplitude or phase of another waveform.

mole The amount of matter which contains 6.02×10^{23} particles.

moment The moment of a force about a point is the magnitude of the force, multiplied by the perpendicular distance of the point from the line of the force. Unit: N m.

monochromatic Describes light of a single frequency.

natural frequency The unforced frequency of oscillation of a freely oscillating object.

negative feedback The output of a system is used to oppose changes to the input of the system, with the result that the changes are reduced.

neutrino A lepton, released during beta-decay.

neutron number The number of neutrons in the nucleus of an atom.

newton The force that will give a 1 kg mass an acceleration of 1 m s^{-2} in the direction of the force. 1 N = 1 kg m s^{-2}.

Newton's first law of motion An object will remain at rest or keep travelling at constant velocity unless it is acted on by a resultant force.

Newton's law of gravitation Any two point masses attract each other with a force that is directly proportional to the product of their masses and inversely proportional to the square of their separation.

Newton's second law of motion The resultant force acting on an object is equal to the rate of change of its momentum. The resultant force and the change in momentum are in the same direction.

Newton's third law of motion When two bodies interact, the forces they exert on each other are equal and opposite.

node A point on a stationary wave with zero amplitude.

noise An unwanted random addition to a transmitted signal.

non-inverting amplifier A circuit, involving the use of an amplifier, in which the output is in phase with the input.

nuclear fission The splitting of a nucleus (e.g. $^{235}_{92}U$) into two large fragments and a small number of neutrons.

nuclear fusion A nuclear reaction in which two light nuclei (e.g. $^{2}_{1}H$) join together to form a heavier but more stable nucleus.

nuclear magnetic resonance A process in which radio waves are absorbed or emitted by nuclei spinning in a magnetic field.

nuclear model of the atom A model of the atom in which negative charges (electrons) are distributed outside a tiny nucleus of positive charge.

nucleon number The number of neutrons and protons in the nucleus of an atom (also called mass number).

nucleon A particle found in an atomic nucleus, i.e. a neutron or a proton.

nucleus The tiny central region of the atom that contains most of the mass of the atom and all of its positive charge.

nuclide A specific combination of protons and neutrons in a nucleus.

number density The number of particles, such as free electrons, per unit volume in a material.

Ohm's law The current in a metallic conductor is directly proportional to the potential difference across its ends, provided its temperature remains constant.

operational amplifier (op-amp) A high-gain electronic d.c. voltage amplifier with differential inputs and, usually, a single output.

optic fibre A glass or plastic fibre that carries light along its length.

oscillation A repetitive back-and-forth or up-and-down motion.

parallel Describes components connected side-by-side in a circuit.

path difference The difference in the distances travelled by two waves from coherent sources at a particular point.

perfectly elastic A collision is perfectly elastic when kinetic energy is conserved. Momentum and total energy are always conserved.

period The time taken by an object (e.g. a planet) to complete one cycle (e.g. an orbit). The period is also the time taken for one complete oscillation of a vibrating object. Unit: second (s).

phase Refers to the point that an oscillating mass has reached in a complete cycle.

phase difference The difference in the phases of two oscillating particles, expressed in degrees or radians.

photoelectric effect An interaction between a photon and an electron in an atom, in which the electron is removed from the atom.

photon A particle of electromagnetic radiation.

piezo-electric crystal A material that produces an e.m.f. when it is stressed, causing its shape to change. Also, when a voltage is applied across it in one direction, it changes its dimensions slightly.

piezo-electric effect The production of an e.m.f. between the faces of a crystal when the crystal is compressed.

Planck constant The constant which links the energy of a photon and its frequency, in the equation:

$$E = hf$$

plum-pudding model A model of the atom in which negative charges are distributed throughout a sphere of positive charge.

positron An anti-electron.

potential difference (p.d.) The energy lost per unit charge by charges passing through a component. Unit: $J\,C^{-1}$ or volt (V).

potential divider A circuit in which two or more components are connected in series to a supply. The output voltage from the circuit is taken across one of the components.

potentiometer A circuit which allows the measurement of an e.m.f. by comparison with a known e.m.f.

power The rate at which energy is transferred or the rate at which work is done. Unit: watt (W).

precession The movement of the axis of a spinning object (proton) around another axis.

precision The smallest change in value that can be measured by an instrument or an operator. A precise measurement is one made several times, giving the same, or very similar, values.

pressure The force acting normally per unit area of a surface:

$$p = \frac{F}{A}$$

Unit: $N\,m^{-2}$ or pascal (Pa).

principle of conservation of energy The idea that, within a closed system, the total amount of energy in all its forms is unchanged during any change.

principle of moments For an object in equilibrium, the sum of clockwise moments about a point is equal to the sum of anticlockwise moments about the same point.

principle of superposition When two or more waves meet at a point, the resultant displacement is the sum of the displacements of the individual waves.

progressive wave A wave that carries energy from one place to another.

projectile Any object thrown in the Earth's gravitational field.

proton number The number of protons in the nucleus of an atom (also called atomic number).

quarks The fundamental particles of which hadrons are made.

radian A unit for measuring angles.
2π radians = 360° or π radians = 180°.

range The horizontal distance covered by an object.

rarefaction A region in a sound wave where the air pressure is less than its mean value.

rectification The process of converting alternating current (a.c.) into direct current (d.c.).

red shift the change in frequency or wavelength of a spectral line observed when the source of light is moving away from the observer; *see* Doppler effect.

reflection The bouncing back of a wave from a surface.

refraction The change in direction of a wave as it crosses an interface between two materials where its speed changes.

regeneration Restoration of a signal to its original form, usually by removing or reducing noise and increasing signal strength.

relative speed The magnitude of the difference in velocities between two objects.

relaxation time The time taken for a nucleus or other excited system to fall back to a lower energy state.

relay An electrically operated switch, caused to open and close by current in a coil.

repeater An electronic device that receives a signal and retransmits it.

resistivity A property of a material, a measure of its electrical resistance, defined by:

$$\rho = \frac{RA}{L}$$

Unit: Ω m.

resistor An electrical component whose resistance in a circuit remains constant, is independent of current or potential difference.

resonance The forced motion of an oscillator characterised by maximum amplitude when the forcing frequency matches the oscillator's natural frequency. A system absorbs maximum energy from a source when the source frequency is equal to the natural frequency of the system.

rest mass The mass of a an isolated stationary particle.

resultant force The single force that has the same effect as all of the forces acting on an object.

right-hand grip rule A rule for finding the direction of the magnetic field inside a solenoid. If the right hand grips the solenoid with the fingers following the direction of the conventional current around the solenoid, then the thumb points in the direction of the magnetic field.

sampling Taking the value of a continuous signal at regular intervals.

scalar quantity A scalar quantity has magnitude but no direction.

semiconductor diode An electrical component made from a semiconductor material (e.g. silicon) that only conducts in one direction. A diode in 'reverse bias' has an infinite resistance.

sensor A device that produces an output (usually a voltage) in response to an input.

series A term used when components are connected end-to-end in a circuit.

sharpness The degree of resolution in an image, which determines the smallest item that can be identified.

simple harmonic motion Motion of an oscillator in which its acceleration is directly proportional to its displacement from its equilibrium position and is directed towards that position.

solenoid A long current-carrying coil used to generate a uniform magnetic field within its core.

specific heat capacity The energy required per unit mass of a substance to raise its temperature by 1 K (or 1 °C). Unit: $J\,kg^{-1}\,K^{-1}$.

specific latent heat of fusion The energy required per unit mass of a substance to change it from solid to liquid without a change in temperature. Unit: $J\,kg^{-1}$.

specific latent heat of vaporisation The energy required per unit mass of a substance to change it from liquid to gas without a change in temperature. Unit: $J\,kg^{-1}$.

speed The rate of change of the distance moved by an object:

$$\text{speed} = \frac{\text{distance}}{\text{time}}$$

Unit: m s^{-1}.

spin A fundamental property of subatomic particles which is conserved during atomic and nuclear reactions.

stationary wave A wave pattern produced when two progressive waves of the same frequency travelling in opposite directions combine. It is characterised by nodes and antinodes. Also known as a standing wave.

strain The extension per unit length produced by tensile or compressive forces:

$$\text{strain} = \frac{\text{extension}}{\text{original length}}$$

strain energy The potential energy stored in an object when it is deformed elastically.

strain gauge A device that contains a fine wire sealed in plastic. Its electrical resistance changes when the object to which it is attached changes shape.

stress The force acting per unit cross-sectional area:

$$\text{stress} = \frac{\text{force}}{\text{cross-sectional area}}$$

strong nuclear force A fundamental force which acts between hadrons.

systematic error An error in readings which is repeated throughout an experiment, producing a constant absolute error or a constant percentage error.

tensile Associated with tension or pulling, e.g. a tensile force.

terminal p.d. The potential difference across an external resistor connected to an e.m.f. source.

terminal velocity The maximum velocity of an object travelling through a fluid. The resultant force on the object is zero.

tesla The SI unit of magnetic flux density (abbreviated T). $1\text{ T} = 1\text{ N A}^{-1}\text{ m}^{-1}$.

thermal energy Energy transferred from one object to another because of a temperature difference; another term for heat energy.

thermal equilibrium A condition when two or more objects in contact have the same temperature so that there is no net flow of energy between them.

thermistor A device whose electrical resistance changes when its temperature changes.

thermocouple A device consisting of wires of two different metals across which an e.m.f. is produced when the two junctions of the wires are at different temperatures.

thermodynamic scale A temperature scale in which temperature is measured in kelvin (K).

threshold frequency The minimum frequency of electromagnetic radiation that will eject electrons from the surface of a metal.

threshold voltage The minimum forward bias voltage across a light-emitting diode (LED) when it starts to conduct and emit light.

time constant The time taken for the current, stored charge or p.d. to fall to 1/e (about 37%) of its original value when a capacitor discharges through a resistor. It is also equal to the product of capacitance and resistance.

torque of a couple The product of one of the forces of a couple and the perpendicular distance between them. Unit: N m.

tracers Radioactive substances used to investigate the function of organs of the body.

transducer A general term used for any device that changes one form of energy into another.

transition When an electron makes a 'jump' between two energy levels.

transverse wave A wave in which the particles of the medium oscillate at right angles to the direction in which the wave travels.

triangle of forces A closed triangle drawn for an object in equilibrium. The sides of the triangle represent the forces in both magnitude and direction.

turns-ratio equation An equation relating the ratio of voltages to the ratio of numbers of turns on the two coils of a transformer:

$$\frac{V_s}{V_p} = \frac{N_s}{N_p}$$

unified atomic mass unit A convenient unit used for the mass of atomic and nuclear particles (1 u is equal to the mass of a ${}^{12}_{6}\text{C}$ carbon atom).

$$1\text{ u} = 1.66 \times 10^{-27}\text{ kg}$$

uniform acceleration Acceleration that remains constant.

uniform motion Motion of an object travelling with a constant velocity.

upthrust The upward force that a liquid exerts on a body floating or immersed in a liquid.

valence band a range of electron energies in a solid; electrons in the valence band are bound to individual atoms.

vector addition Using a drawing, often to scale, to find the resultant of two or more vectors.

vector quantity A quantity which has both magnitude and direction.

vector triangle A triangle drawn to determine the resultant of two vectors.

velocity The rate of change of the displacement of an object:

$$\text{velocity} = \frac{\text{change in displacement}}{\text{time}}$$

Unit: m s^{-1}.

You can think of velocity as 'speed in a certain direction'.

virtual earth approximation An approximation in which the two inputs of an op-amp are nearly at the same potential.

viscous forces Forces that act on a body moving through a fluid that are caused by the resistance of the fluid.

voxel A small cube in a three-dimensional image, the equivalent of a pixel in a two-dimensional image.

wave A periodic disturbance travelling through space, characterised by a vibrating medium.

wavelength The distance between two adjacent peaks or troughs in a wave.

weak nuclear force A fundamental force, involved in radioactive β-decay.

weight The force on an object caused by a gravitational field acting on its mass:

$$\text{weight} = \text{mass} \times \text{acceleration of free fall}$$

Unit: newton (N).

wire-pair A type of electrical wiring in which the two conductors needed to carry a signal are placed close together.

work done The product of the force and the distance moved in the direction of the force.

work function The minimum energy required by a single electron to escape a metal surface.

X-ray tube A device that produces X-rays when accelerated electrons hit a target metal.

Young modulus The ratio of stress to strain for a given material, resulting from tensile forces, provided Hooke's law is obeyed:

$$\text{Young modulus} = \frac{\text{stress}}{\text{strain}}$$

Unit: pascal (Pa; or MPa, GPa).

zero error A systematic error in an instrument that gives a non-zero reading when the true value of a quantity is zero.

Index

absolute zero, 333–4
absorption line spectra, 475–6
acceleration, 15–17, 18–20, 38–9
acceleration, and projectiles, 29–31
acceleration, due to gravity, 25–7, 41–3
acceleration, in equations of motion, 20–3
acceleration, non-uniform, 24
acceleration, of free fall, 41–3
acceleration, uniform, 24
acceleration–time graphs, 292–3
acoustic impedance, 519–20
air columns, and observation of stationary waves, 213–15
air resistance, 42–6, 298, 445
alpha-decay, 229, 231, 232, 490–1, 494, 496–8
alpha-particles, 223–5, 231, 232–3, 234
alpha-radiation, 233, 234
alpha-scattering, 223–5
alternating currents, 416–27, 452–6
alternating currents, from generators, 446
alternating currents, in electricity supply, 457
alternating currents, in transformers, 446, 458–9
alternating currents, rectification of, 460–1
alternating voltages, 453–5, 457, 460
alternating voltages, in electricity supply, 457
aluminium, as absorber of radiation, 234
ammeters, 129, 136
ampere, as SI base unit, 40
ampere, definition of, 131, 416
amplifiers, inverting, 397
amplifiers, non-inverting, 398
amplifiers, operational (op-amps), 393–8
amplitude, and damping, 297–9
amplitude, and intensity, 182
amplitude, and resonance, 299–301
amplitude, angular, 245
amplitude, of alternating current, 452
amplitude, of alternating voltage, 453
amplitude, of oscillations, 288, 291, 293, 295
amplitude, of particle vibrations, 162,
amplitude, of signals on an oscilloscope, 454
amplitude, of stationary waves, 211–12, 213
amplitude, of waves, 179–80, 182–3, 196, 197–8, 521
amplitude modulation (AM), 310–11, 313–4
AM transmissions, 311, 313–14
analogue meters, 400
analogue signals, 314–15
analogue-to-digital conversion (ADC), 315–17
angle, of alpha-scattering, 224
angle, of diffraction, 204, 205–6
angles, in circular motion, 259
angles, measured in radians, 260
angles, uncertainty in measurement of, 243
angular displacement, 259–60, 261
angular frequency, 292, 294, 452, 523
angular momentum, 88
angular velocity, 261–2, 264, 523
antimatter, 230-1
antinodes, 212–3, 216
atomic structure, 223–6
atoms, 42, 130, 476, 478–9
atoms, and elastic potential energy, 108
atoms, and ionising radiation, 509–10
atoms, and radioactivity, 497
atoms, as impurities, 162
attenuation, 315–19, 320, 323, 509–12
attraction, between currents, 415
attraction, electrostatic force of, 117, 361, 366
attraction, gravitational force of, 274, 277, 366
Avogadro constant (or number), 349, 354, 543

balanced equations, 5–6, 96, 490
balanced forces, 44–5, 60, 428–9
balanced potentiometer, 173
bandwidth, 312–4, 320, 321–3, 394, 396–7
Barton's pendulums, 299–300
base stations, 279
base units, 40, 416, 542
beta-decay, 80, 231–2, 490, 494, 500
beta-particles, 231–3
beta-radiation, 234
binding energy, 494–6
Boyle's law, 349–50, 354
braking radiation, 508–9
Bremsstrahlung radiation, 508–9
bridge rectifier, 460
Brownian motion, 346–7

calibration, 242, 335, 400
capacitance, 373–82
capacitors, 129, 373–82, 393, 461, 535
carrier waves, 310–14
cathode-ray oscilloscopes (c.r.o.), 180, 453–4

cell phone network *see* mobile phone network
cells, electrical, 129, 169–70
centre of gravity, 41–2, 59, 381
centripetal acceleration, 262–4, 265
centripetal forces, 262–4, 265–6, 277–9, 425
changes of state, 328–30, 336, 338
channels of communication, 317–23
charge, electrostatic, 117–122
charge, on electron, 132, 231,426, 430, 436
charge, on proton, 231
charge carriers, 130–1, 132–4, 429
charge storage, 374
charge-to-mass ratio of electron, 426, 430
charged particles, 117–9, 122, 430, 470, 509
charges, orbiting, 425–6
charges, point, 360–2
Charles's law, 350–1
circuit design, 144
circuit symbols and diagrams, 128–9
circular motion, 259–66, 292
circular orbits, 264, 278–9, 417, 425
climbing bars, 78–9
clocks, pendulum, 251–2
closed systems, 79, 88
coaxial cables, 317, 319–20
coherence, 199–200
collimated beam, 508, 510, 512
collisions, as a cause of pressure, 48, 352–4
collisions, between gas particles, 346–8, 352–4
collisions, elastic, 89–90
collisions, in particle acclerators, 230
collisions, inelastic, 89–91
collisions, of objects, 86–90
combining displacements, 8–10
combining forces, 54–6
combining velocities, 10
combining waves, 193–4
compressive forces, 296, 332, 349–50, 389–90
computer-aided design (CAD), 144
computerised axial tomography (CAT or CT), 513–6
conductance, 380
conduction, and band theory, 478–9
conduction, of electricity, 117, 130–2, 157, 162, 470 , 478–9
conduction, of thermal energy, 333, 339–40
conduction electrons, 130, 132–4, 162, 388–9, 439, 443
conductors, 129, 132, 136, 157–8, 160, 163, 185–6, 320, 409–15, 429, 437–9, 443–5
conservation, mass–energy, 492–3
conservation of charge, 144, 379
conservation of energy, 79–80, 90, 148, 296, 332, 444, 492
conservation of mass, 491–2
conservation of momentum, 88, 90–1
constant, Avogadro, 349, 354, 543
constant, Boltzmann, 354–5, 543
constant, decay, 498–9, 501, 544
constant, force (of spring), 104–5
constant, gravitational (*G*), 273, 276, 543
constant, Planck, 469, 470–1, 480, 543
constant, universal molar gas, 351
constant of proportionality, 96, 294, 351, 361, 441
constant phase difference, 200–1
constructive interference, 196–9, 201, 202–3, 204
contact force, 47–9, 54, 56–7, 265–6
contrast media, 513
conventional current, 129–30, 374, 410, 423–4, 439, 443
Coulomb's law, 360–7
count rate, 496–7, 498, 500
couple, 63
crash-landings, 91–2
crosslinking, 317
crystals, piezo-electric, 517–8
current, conventional, 129–30, 374, 410, 423–4, 439, 443
current–voltage (*I*–*V*) characteristic, 157–8
currents, alternating, 416–27, 452–6
currents, electric, 40, 117, 129–34, 160, 185, 289, 423, 428, 443, 479, 542
currents, root-mean-square (r.m.s.), 455–7
currents, sinusoidal, 416–27, 452–6
curved trajectory, 122

damped oscillations, 297–9
damping, 298–9, 301, 518
damping, critical, 301
damping, eddy current, 445
de Broglie wavelength, 480
decay, spontaneous, 497
decay constant, 498–9, 501, 544
decay graphs and equations, 298, 499–501
deformation, plastic, 107–8
density, 102, 103, 354, 542
density, and acoustic impedance, 519
density, magnetic flux, 411–3, 414, 416, 424–5, 428, 523, 524
density, nuclear, 226
density, number, 132–4
derived units, 40, 416
destructive interference, 197–9
dielectrics, 373–4
difference, potential (p.d.), 134–7, 146, 151, 157–8, 160, 170, 172–4, 319–20, 363–5, 373–4, 393–6, 470
diffraction, 194, 196, 202–6, 468, 480–1

diffraction, electron, 480–1, 482
diffraction gratings, 202–6, 475
diffraction patterns, 199, 481
digital signals, 314–5
digital-to-analogue conversion (DAC), 315
direct current, 455
dispersion, 205, 475
displacement, 4–10, 17–18, 20
displacement–distance graphs, 179, 193, 212
displacement–time graphs, 6–10, 17–18, 282, 291–2
distance, 4, 8
dividers, potential, 172, 388–90, 391–2, 398
Doppler effect, 184–5
drag, 45–6, 47
driver cells, 172–3

earth, electrical, 119, 129, 363–4, 393–4
Earth, gravitational field due to, 41, 49, 59, 71, 92, 122, 264, 272–9, 346
Earth, magnetic field of, 412, 523
Earth, orbit around Sun, 262–3
earth, virtual, 397
earthquakes, 301
echo sounding, 517, 518–9
eddy currents, 445–7
effect, Doppler, 184–5
effect, photoelectric, 469–74
effect, piezo-electric, 517–8
Einstein, and the photoelectric effect, 469, 472–4
Einstein relation, 469
Einstein's mass–energy equation, 492–3
elastic collisions, 89–90
elastic limit, 105, 108–9
elastic potential energy, 108–9, 296
electric charges, 117–8, 131–2, 134, 232, 360, 363, 366–7, 428
electric currents, 40, 117, 129–34, 160, 185, 289, 423, 428, 443, 479, 542
electric currents, and magnetic fields, 413–6, 436–7
electric currents, induced, 436–7, 443
electric fields, 75, 117–22, 130, 188, 225, 233, 360–7
electric force, 130, 133, 351, 360, 417, 429–30, 484
electric potential, 363–6, 367, 544
electric potential energy, 75, 363
electrical cells, 129, 169–70
electrical components, 129, 136
electrical power, 151, 419–20, 424
electrical resistance, 135, 149–50, 152, 457, 459, 479, 532–3
electricity, generation of, 436, 443, 446, 457
electricity, mains, 452, 455, 457
electricity supply, 457
electrolytes, 130,
electromagnetic energy, 231
electromagnetic fields (EM fields), 320, 417
electromagnetic force, 118
electromagnetic induction, 436–47, 453
electromagnetic radiation, 186–7
electromagnetic spectrum, 314, 469
electromagnetic waves, 181–2, 185–6, 188, 310, 319, 320, 468–9
electromagnets, 26, 157, 251, 399, 407–9, 416, 423–4, 436–7, 443–4, 446
electromotive force (e.m.f.), 134, 169–74, 335, 373–4, 399, 437–9, 453, 458
electromotive force (e.m.f.), and Kirchhoff's laws, 145–8
electromotive force (e.m.f.), induced, 439–47, 517
electron beam tubes, 423, 424, 453–4
electron beams, 122, 423, 427–8, 511
electron diffraction, 480–2
electronic sensors, 387
electronic timers, 26
electrons, 117, 122, 132–4, 223–30, 289, 408, 423–4, 428, 471–4, 483, 491–2
electrons, and the origin of line spectra, 476
electrons, and the photoelectric effect, 471–4
electrons, as beta-particles, 231–3
electrons, charge of, 131
electrons, charge-to-mass ratio of, 426
electrons, conduction, 130, 132–4, 162, 388–9, 439, 443
electrons, discovery of, 429–30
electrons, drift velocity of, 429
electrons, energy of, 162, 478–9, 509
electrons, in capacitors, 373–5
electrons, in the production of X-rays, 186–7, 507–9
electronvolt, 235, 470, 543
electron waves, 480–1
electrostatic forces, 117, 224–5, 229, 232, 363, 472
elementary charge, 131
e.m.f. *see* electromotive force
emission line spectra, 475–6
emission of particles in radioactivity, 490–1, 498
energy, 73–4, 153, 277, 299, 302, 348, 454, 457
energy, atomic, 477
energy, binding, 494–6
energy, conservation of, 79–80, 90, 148, 332, 492–3
energy, electric potential, 363–5
energy, electrical, 138
energy, gravitational potential, 75–9, 276
energy, internal, 90–1, 331–3, 351–5

energy, kinetic, 71, 76–7, 79, 89, 90–2, 233–4, 296–7, 426, 444, 470, 473–4, 494, 508
energy, mass–, 492–3
energy, of electrons, 478–9
energy, of photons, 477
energy, potential, 75
energy, strain, 108
energy, units of, 72, 138, 148, 363, 377, 416, 470, 492–3
energy, wave, 182–3
energy changes, 77–8, 296–7, 329–331, 336–9, 363, 470, 492
energy conservation, 79–80, 90, 148, 332, 492–3
energy graphs, 296–7
energy levels, 476–7
energy storage, 373–6, 494
energy transfers, 78–80,
energy transformations, 70
equation, wave, 183, 213–4, 469
error bars, 537–8
errors, parallax, 240, 242, 250, 426
errors, percentage, 203
errors, random, 26, 242, 244
errors, reducing and eliminating, 217–8
errors, systematic, 26, 27, 218, 242, 244
errors, zero, 241, 242, 253
evaporation, 330–1
explosions, 91–2
exponential decay, 298, 389, 499–500, 510, 520, 523, 535
exponential envelope, 297

Faraday's law of electromagnetic induction, 441–2
feedback, negative, 396–7
field of force, 118, 360, 407
fields, electric, 75, 117–22, 130, 188, 225, 233, 360–7, 417
fields, gravitational, 45, 75, 122, 272–80, 363, 366–7, 417, 523
fields, magnetic, 118, 157, 185–6, 188, 233, 399, 407–17, 522–5
first law of thermodynamics, 332
fission, nuclear, 496
Fleming's left-hand (motor) rule, 409–10, 411, 416, 423–5, 443–4
Fleming's right-hand (generator) rule, 438, 443
fluid resistance, 45–7
fluoroscopy, 511
flux, magnetic, 399, 439–40, 441–2, 446–7
flux density, magnetic, 411–3, 414, 416, 424–6, 428–9, 437, 523, 524
flux linkage, magnetic, 439–40, 441–2, 446–7
FM transmission, 311–12, 313–4
force, and acceleration, 38–9
force, centripetal, 262–6, 278
force, contact, 47, 48, 49, 57
force, drag, 44, 45
force, electric, 130, 133, 351, 360, 367, 417, 429–30, 484
force, electromagnetic, 186
force, electromotive (e.m.f.), 134, 169–74, 335, 373–4, 399, 437–9, 453, 458
force, field of, 118, 360, 407
force, frictional, 43, 266, 299
force, gravitational, 28, 29, 41–3, 59, 75, 272–7, 367
force, magnetic, 407–17, 423–8, 429–30, 438–9
force, moment of a, 60
force, normal, 102
force, nuclear, 229, 230, 232, 491, 494
force, restoring, 290
force, resultant, 38, 44, 54, 55–7, 96, 292
force, strong nuclear, 229, 230, 232, 491, 494
force, viscous, 45
force, weak nuclear, 232
force–extension graphs, 104–9
force meters, 242, 252–3
forced oscillations, 286–7
forces, 28, 49–50, 54–5, 71–2, 76, 88, 95, 302, 348, 353
forces, balanced, 44–5, 60, 61–2, 96, 428–9
forces, compressive, 104
forces, coplanar, 54–8
forces, identifying, 47
forces, tensile, 104–9
forces, triangle of, 55
forces, unbalanced, 44–5, 262, 290
free electrons *see* electrons, conduction
free fall, 535
free fall, acceleration of, 41–3
free oscillations, 286–7
free-body force diagram, 58
frequency, 180, 183, 186, 200, 211–15, 286, 288, 293–5, 299–300, 312–3, 320–1
frequency, angular, 292, 294, 452, 523
frequency, carrier wave, 310
frequency, fundamental, 216
frequency, Larmor, 523
frequency, natural, 286, 287, 299
frequency, observed, 184–5
frequency, resonance, 301
frequency modulation (FM), 196, 311–12, 313–14
friction, 43, 266, 299
fringe separation, 202–3
fundamental frequency, 216
fundamental mode of vibration, 216
fuses, electrical, 137
fusion, nuclear, 496

gain, 393–5
gamma-emission, 231
gamma-radiation, 231, 233, 234, 469
gamma-rays, 231, 469
gas atoms, 351
gases, 74, 328, 346–55
gases, ideal, 346–55
gases, kinetic theory of, 352–5
gases, real, 350–1
Geiger–Müller (GM) tube, 234, 496
generators, 436, 438, 445–6
generators, Van de Graaff, 365
geostationary orbit, 279–80, 322–3
graphs, acceleration–time, 292–3
graphs, displacement–distance, 179, 193, 212
graphs, displacement–time, 6–10, 17–18, 282, 291–2
graphs, force–extension, 104–9
graphs, stress–strain, 106–7
graphs, velocity–time, 17–20, 22–3, 24, 291–2
gravitational constant (*G*), 273, 276, 543
gravitational field strength (*g*), 42, 75, 274–5
gravitational fields, 45, 75, 122, 272–80, 363, 366–7, 417, 523
gravitational force, 28, 29, 41–3, 59, 75, 272–7, 367
gravitational potential, 276–7, 364–5, 367, 544
gravitational potential energy (GPE), 75–9, 276
gravitational pull, 262–3
gravity, 25–30, 41–2, 47, 73, 92, 103, 264, 272–80, 417
gravity, centre of, 41–2, 59, 381
ground state, 477
gyromagnetic ratio, 523

half-life, 298, 499–501
Hall probe, 411–12, 428–9
heating ice, 329–30
Helmholtz coils, 427
high-speed trains, 407
Hooke's law, 104–9, 538
human power, 81

I–V (current–voltage) characteristic, 157–8
ice, 329–30
ideal gases, 346–55
image intensifiers, 511
impedance matching, 519–20
inelastic collisions, 89–91
inertia, 39, 43–4, 426
instantaneous speed, 2
intensifier screens, 511
intensifiers, image, 511
intensity, 182–3, 199, 202, 214, 225, 234, 388, 391, 473–4, 509, 519–20
interference, electrical, 313, 315, 319–20
interference, of waves, 196–9, 468
interference fringes, 201–3
interference pattern, 200–1
internal energy, 90–1, 331–3, 351–5
internal resistance, 134, 169–71, 394
intrinsic semiconductor, 479
inverting amplifiers, 397
ionisation, 232–3
ionising radiation, 232–4
ions, 234
isolated atoms, 329, 478
isotopes, 227–8, 499

joules, 72, 138, 148, 363, 377, 416, 470, 492–3

Kelvin scale, 333–4
kinematics, 2–11
kinetic energy, 71, 76–7, 79, 89, 90–2, 233–4, 296–7, 426, 444, 470, 473–4, 494, 508
kinetic model of a gas, 352–5
kinetic model of matter, 328–32
kinetic theory of gases, 352–5
Kirchhoff 's first law, 144–5, 146–52
Kirchhoff 's second law, 145–6, 146–52
Kundt's dust tube, 216–7

Larmor frequency, 523
LDR *see* light-dependent resistor
LED *see* light-emitting diode
Lenz's law, 441–3, 445
light beams, 3–4, 19, 27, 198, 202–3
light-dependent resistor (LDR), 129, 387, 388, 391–2, 479
light-emitting diode (LED), 129, 323, 374, 399–400, 470–1
light gates, 3, 19, 27, 251
light waves, 186, 193–206
line spectra, 475–80
linear momentum, 86–97, 232, 348, 353, 426, 480
longitudinal waves, 181–2, 214–5, 517
lost volts, 170–1, 173, 394
loudspeakers, 129, 196–7, 198–9, 200, 216–7, 287, 393

magnetic fields, 118, 157, 185–6, 188, 233, 399, 407–17, 522–5
magnetic flux, 399, 439–40, 441–2, 446–7
magnetic flux density, 411–3, 414, 416, 424–6, 428–9, 437, 523, 524
magnetic flux linkage, 439–40, 441–2, 446–7

magnetic forces, 407–17, 423–8, 429–30, 438–9
magnetic resonance imaging (MRI), 522–5
magnets, 97, 157, 407–16, 423, 436–47, 524
magnitude, of electric charge, 131
magnitude, of forces, 48–9, 54–5, 60, 294, 414, 425
magnitude, of induced e.m.f., 441
magnitude, of momentum, 90, 93
magnitude, of scalar and vector quantities, 4, 10, 261
magnitude, orders of, 133, 187, 483
mass, 38–9
mass, and density, 102
mass, and gravitational forces, 272–4
mass, and inertia, 43–4
mass, and momentum, 87–8, 95–6
mass, and weight, 42–3, 75
mass, as SI base unit, 39–40
mass, electron, 223, 231
mass, proton, 226
mass, relative atomic, 228
mass defect, 492
mass–energy conservation, 232, 492–3
mass–energy equation, 492
mass number, 226
mass–spring system, 287–90
masses, point, 272–4, 276
matter, 102–3, 223, 230, 509–10
mechanical waves, 180–3
medical imaging, 507–25
Melde's experiment, 213–5
metallic conductors, 136, 157–8
metal-wire strain gauge, 389–90
microwave links, 317, 320–2
microwave transmitters, 199, 214
microwaves, 187, 196, 214, 320–2, 469, 507
microwaves, interference of, 199
Millikan's oil-drop experiment, 430
mobile phone network, 196
mobile phones, 196, 315, 321, 360
model, plum pudding, 225
molecular energy, 331–2, 354–5
molecules, 331–3, 346–9, 351–4
moments of forces, 60–3
momentum, angular, 88
momentum, linear, 86–97, 232, 348, 353, 426, 480
monochromatic light, 202–3, 204
Moon, 41, 42, 272, 279
motion in two dimensions, 328–31
motion sensors, 4, 19–20, 291
MRI (magnetic resonance imaging), 407–17, 423–8, 429–30, 438–9
musical instruments, 216, 289, 301

natural frequency, 286, 287, 299
negative feedback, 396–7
negative temperature coefficient (NTC) thermistors, 159, 161, 391
neutron number, 226–7, 490–1
neutrons, 117, 132, 225–7, 491–2, 494, 495
newtons, 72–3
Newton's first law of motion, 96, 262
Newton's law of gravitation, 273–4
Newton's second law of motion, 38–9, 57, 96–7, 264–5, 353
Newton's third law of motion, 49, 97, 360, 412, 416
noise, 313–5, 318–9, 394
non-inverting amplifiers, 398
non-uniform acceleration, 24
NTC (negative temperature coefficient) thermistors, 159, 161, 391
nuclear density, 226
nuclear fission, 496
nuclear forces, 229, 230, 232, 491, 494
nuclear fusion, 496
nuclear model of atoms, 223–6
nuclei, 225–8,
nucleon number, 226–9, 232, 490–6
nucleons, 226, 491–6
number, Avogadro, 349, 354, 543
number, neutron, 226–7, 490–1
number, nucleon, 226–9, 232, 490–6
number, proton, 226–9, 232, 490–1

ohm, 135
ohmic components, 158–9
Ohm's law, 158–9
open-ended air columns, 215
operational amplifiers (op-amps), 393–8
optic fibres, 317, 322–3
orbital period, 278–9
orbital speed, 265
orbiting charges, 425–6
orbits, 262–5
oscillations, 286–302
oscilloscopes, 180, 453–4

particle models, 467–9
p.d. (potential difference), *see* potential difference
pendulums, 41, 77, 243, 251–2, 256, 287, 289, 290
pendulums, Barton's, 299–300
penetrating power, 234
people waves, 482

perfectly elastic collisions, 89–90
period, 180–1, 183, 212, 278–9, 288, 291–2, 295–6, 322–3, 452, 453–4
period, orbital, 278–9
Periodic Table, 227, 545
periodic variation, 188
permanent magnets, 407–8, 410
permittivity of free space, 361, 543
phase, 182, 288–9
phase difference, 182, 197–8, 200–1, 212, 288–9, 397
photoelectric effect, 469–74
photon energies, 477
photons, 469–80
piezo-electric crystals, 517–8
piezo-electric effects, 517–8
piezo-electric transducers, 517–8
Planck constant, 469, 470–1, 480, 543
plastic deformation, 107–8
plum pudding model, 223
point charges, 360–2
point masses, 272–4, 276
potential difference (p.d.), 134–7, 146, 151, 157–8, 160, 170, 172–4, 319–20, 363–5, 373–4, 393–6, 470
potential dividers, 172, 388–90, 391–2, 398
potential energy, 75
potential energy, electric, 363–5
potential energy, gravitational, 75–9, 276
potentiometer circuits, 172–4
potentiometers, 172–4
power, 80–2, 136, 138, 182, 313, 317–8, 338, 393, 455–7, 510
power, penetrating, 234
power lines, 323, 457
powers of ten, 6, 121, 542
prefixes, 41, 542
pressure, 47–8, 74, 102–3, 181, 314, 333, 346, 348, 351–4
principle of conservation of energy, 79–80, 90, 148, 296, 332, 444, 492
principle of moments, 60–2
principle of superposition, 193–4
progressive waves, 179, 211–3,
projectiles, 28–31
proton number, 226–9, 232, 490–1
protons, 117, 130–2, 225–232, 491–5, 522–4

quantum physics, 467–83

radiation from radioactive substances, 223, 229, 231, 232–4
radiation penetration, 233–4
radioactive decay, 80, 186–7, 231–4, 490–501
radioactivity, 231–4, 490–501
radioisotopes, 499
radio telescopes, 179
radio waves, 186, 196, 198, 302, 310–14, 320–2, 469
real gases, 350–1
rectification, 460–1
rectifiers, 460–1

relative speed, 90
relays, 398–9
repeaters, 318–9, 323
repulsion, 117–8, 363, 367
repulsion between currents, 415
resistance, air, 42–6, 298, 445
resistance, electrical, 135, 149–50, 152, 457, 459, 479, 532–3
resistance, internal, 134, 169–71, 394
resistivity, 162–3
resistors, 129, 134, 145, 170, 172, 379–80, 388, 544
resonance, 214, 217–8, 299–302
resonance, magnetic, 522–5
resonance tubes, 214
rest mass, 492
resultant force, 38, 44, 54, 55–7, 96, 292
resultant velocity, 10
right-hand grip rule, 409
ripple tanks, 181–4, 195
ripples in water, 195–8
road traffic accidents, 23
root-mean-square (r.m.s.) currents and voltages, 455–7

scalar quantity, 4, 10, 261
semiconductor diodes, 129, 160
semiconductors, intrinsic, 479
sensors, 335, 337, 387–91
sensors, motion, 291
shock waves, 301
sidebands, 312–3
signal attenuation, 315–19, 320, 323
signal-to-noise ratio, 318–9
sign conventions, 29, 54
simple harmonic motion (s.h.m.), 289–302
sinusoidal currents, 416–27, 452–6
SI units, 39–41, 416, 542
slit separation, 201–5
slit-to-screen distance, 201–2
solenoids, 407–9, 440, 444–5
sound waves, 180–1, 183, 184–5, 213–15, 289, 310
specific heat capacity, 336–8
specific latent heat, 338–40
speed, 2–5, 15, 19–20, 43–5, 75
speed, and circular motion, 261–4

speed, and Doppler effect, 183–4
speed, and kinetic energy, 76–7
speed, and momentum, 87
speed, average, 2
speed, in simple harmonic motion, 287–90, 295–6
speed, instantaneous, 2
speed, of atoms in a gas, 346–7, 354–5
speed, of electrons in a current, 132–3
speed, of light, 186–7, 213, 310
speed, of particles emitted during radioactivity, 231
speed, of sound, 216–7, 310
speed, orbital, 265
speed, relative, 90
speed, top, 44–5
speed, wave, 183, 310
speed cameras, 39
spontaneous decay, 497
spring constant, 104
springs, 47, 104–7, 181, 537–8
spring–trolley systems, 291
springy collisions, 86–7
stationary waves, 211–18
step-down transformers, 458–9
step-up transformers, 458–9
sticky collisions, 87
strain, 105–9, 517
strain energy, 108
strain gauges, 389–90
strength, electric field, 119–22, 360, 362, 365, 367, 411
strength, gravitational field, 42, 75, 274–6, 367, 411
strength, magnetic field, 188, 411
stress, 105–7, 389–90, 517
stretched strings, 47, 104–9, 296
stretching materials, 105–7
superconductivity, 157
superconductors, 157
superposition of waves, 193–4

temperature, 328–40
temperature, thermodynamic, 40, 333–4, 350–355
tensile forces, 104–5
tension, 104–5, 262
terminal velocity, 45–6
tesla, 411
thermal energy, 333, 411
thermal equilibrium, 333
thermistors, 129, 159–61, 334–5, 387, 389
thermocouples, 334–5
thermodynamic scale, 40, 333–4, 350–355
thermometers, 332–3, 334–5
threshold frequency, 471–4
threshold voltage, 160–1, 470–1
ticker-timers, 3, 19, 27
top speed, 44–5
torque, 63
transducers, 387–8, 517–20
transducers, piezo-electric, 517–8
transformers, 446–7
transverse waves, 181–2, 188, 193, 211, 215
triangle of forces, 55
turning forces, 59–63
turns-ratio equation for transformer, 458

ultrasound scanning, 516–22
ultraviolet radiation, 106–7, 472
unbalanced forces, 44–5, 60, 428–9
uncertainty, 26, 55, 202, 240–54, 493, 536–8
uniform acceleration, 24–5
uniform electric field, 118–122
uniform motion, 44–5, 262
units, SI, 39–41, 416, 542
upthrust, 47, 48, 103, 346

vapour, 330
variable resistances, 391
vector addition, 10
vector components, 28–9
vector diagrams, 263
vector quantities, 4, 10, 261
velocity, 5–10, 15–24
velocity, angular, 261–2, 264, 523
velocity–time graphs, 17–20, 22–3, 24, 291–2
velocity selectors, 427–8
virtual earth approximation, 397
viscous force, 45
voltage, 49, 120–1, 134–8, 145, 157–8, 159–61, 172–4, 314–7, 335, 391–2
voltage, alternating,
voltage, input, 393–4
voltage, output, 393–4
voltage, root-mean-square (r.m.s.), 455–7
voltmeters, 134
volume, 348–55
volume, of proton, 226

watt, 80–1, 136
wave energy, 182–3
wave equation, 183, 213–14, 469
wavelength, 179–87, 193–6, 213–18, 311, 321, 469, 471, 475–7, 480–3, 507–9

wave models, 468, 474, 480
wave–particle duality, 480
wave speed, 183
waves, 195–6
waves, carrier, 310–14
waves, electromagnetic, 181–2, 185–6, 188, 310, 319, 320, 468–9
waves, light, 186, 193–206
waves, longitudinal, 181–2, 214–5, 517
waves, sound, 180–1, 183, 184–5, 213–15, 289, 310
waves, transverse, 181–2, 188, 193, 211, 215
weight, 39–40, 41–2, 266, 273–5
white light, 161, 202–3, 205–6, 475–6
wire-pair, 319–20
work done, 71–2, 74–6, 78–80, 108, 121, 134, 263, 276, 363, 365, 375–6
work function, 472–4

X-ray attenuation, 509–12
X-ray images, 507–13
X-ray spectrum, 508–9
X-ray tubes, 507–8, 509, 514
X-rays, 186–7, 469, 507–13, 514–16

Young double-slit experiment, 198–9, 200
Young modulus, 106–7

Acknowledgements

The authors and publishers acknowledge the following sources of copyright material and are grateful for the permissions granted. While every effort has been made, it has not always been possible to identify the sources of all the material used, or to trace all copyright holders. If any omissions are brought to our notice, we will be happy to include the appropriate acknowledgements on reprinting.

Cover David Parker/SPL; Chapter 1Liufuyu/Thinkstock; 1.1 Edward Kinsman/SPL; 1.2 Yves Herman/Reuters/Corbis; 1.12 Robert Harding Picture Library Ltd/Alamy; Chapter 2 Vlad Kochelaerskiy/Thinkstock; 2.1 Time & Life Pictures/Getty Images; 2.7 David Scharf/SPL; 2.12 Lockheed Martin Astronautics/SPL; 2.22 Actionplus; 2.27 Loren Winters, Visuals Unlimited/SPL; Chapter 3 Comstock/Thinkstock; 3.1Ondrej Zabransky/ Shutterstock; 3.7 Keith Kent/SPL; 3.9 TBC; 3.11 Getty Images; 3.12 NBAE/Getty Images; Chapter 4 Stockbyte/Thinkstock; 4.1 Getty Images; 4.9 Andrew Lambert Photography/SPL; 4.14 AFP/ Getty Images; 4.16 Corbis; Chapter 5 123ArtistImages/Thinkstock; 5.1 George Steinmetz/Corbis; 5.2, 5.3 AFP/Getty Images; 5.12 Jamalrani/Thinkstock; 5.15 Getty Images; 5.18 Daniel Sambraus/Getty Images; Chapter 6 f9photos/Thinkstock; 6.1 TRL Ltd/SPL; 6.2 ArabianEye/Getty Images; 6.3a,b, 6.4 Andrew Lambert Photography/SPL; 6.7 Motoring Picture Library/Alamy; 6.12 Mehau Kulyk/SPL; 6.20 SPL; Chapter 7 Jerónimo Contreras Flores/Thinkstock; 7.1 Stacy Walsh Rosenstock / Alamy; 7.14 Getty Images; Chapter 8 miluxian/Thinkstock; 8.1 Kent Wood/SPL; 8.5, 8.13 Andrew Lambert Photography/SPL; Chapter 9 gyn9038/Thinkstock; 9.1, 9.2, 9.3, 9.4Adam Hart-Davis/SPL; 9.13 Andrew Lambert Photography/SPL; 9.15 Leslie Garland Picture Library / Alamy; 9.16 Maximilian Stock Ltd/SPL; Chapter 10 Janka Dharmasena/Thinkstock; 10.1 Oscar Malpica/ Alamy; 10.2 VisualField/iStock; 10.19a,b, 10.23 Andrew Lambert Photography/SPL; Chapter 11 Dant/Thinkstock; 11.1 Andy Crump/SPL; 11.3 Richard Megna/Fundamental Photos/SPL; 11.4 imageBROKER/Alamy; Chapter 12 Janka Dharmasena/ Thinkstock; 12.1a SPL; Chapter 13 Epic Stock Media/Thinkstock; 13.1 Jim Reed Photography/SPL; 13.2 Doug Johnson/SPL; 13.4 HERMANN EISENBEISS/SPL; 13.6 sciencephotos/Alamy; 13.10 Chris Pearsall/Alamy; 13.13 Dr Morely Read/SPL; 13.14 Physics Today Collection/American Institute of Physics/SPL; Chapter 14 Steen Wackerhausen/Thinkstock; 14.1 Pascal Goetgheluck/SPL; 14.2 Mr Twister/Shutterstock; 14.5 sciencephotos/Alamy; 14.9 superclic/Alamy; 14.13 Bruce Coleman Inc/ Alamy; 14.21 Phototake Inc/Alamy; 14.24 Photography by Ward / Alamy; Chapter 15 Goodshoot/Thinkstock; 15.1a Alex Bartel/ SPL; 15.1b Bettmann/Corbis; 15.2 Tim Ridley/Dorling Kindersley/Getty Images; 15.7 Andrew Lambert Photography/SPL; 15.13 Ian Sanders/Alamy; Chapter 16 H. Mark Weidman Photography/Alamy; 16.1 Robert Gilhooly/Alamy; 16.2 University of Cambridge, Cavendish Laboratory/SPL; 16.8 Stocktrek Images/Alamy; 16.10 Adam Hart-Davis/SPL; 16.11 Lawrence Berkeley Laboratory/SPL; 16.13 Jean-Loup Charmet/SPL; 16.17 N Feather/SPL; 16.20 Leslie Garland Picture Library/Alamy; Chapter 17 Sergiy Serdyuk/Thinkstock; 17.1 Dinodia Photos / Alamy; 17.12 Marvin Dembinsky Photo Associates / Alamy; 17.16 Gray wall studio/Shutterstock; Chapter 18 eddie toro/Thinkstock; 18.1 Ken Fisher/ The Image Bank/Getty Images; 18.10 Dennis Hallinan / Alamy; Chapter 19 Bram Janssens/Thinkstock; 19.1 Photoservice/iStock; 19.2, 19.3, 19.28, 19.30, 19.31Andrew Lambert Photography/SPL; 19.29 Scott R Barbour/Getty Images; 19.33 Deavid Leah/SPL; 19.36 Simon Fraser/SPL; Chapter 20 gilas/ Thinkstock; 20.1 SPL, 20.2 ONOKY - photononstop/Alamy; 20.13 majorosl/iStock; 20.14 Kpphoto/iStock; 20.15 Peter Burnett/ iStock; 20.17 Peter Ryan/SPL; 20.19 Eyebyte/Alamy; Chapter 21 Janka Dharmasena/Thinkstock; 21.1 Tony Camacho/SPL; 21.7 Martyn F. Chillmaid/SPL; 21.13 Andrew Lambert Photography/SPL; Chapter 22 TongRo Images/Thinkstock; 22.1 Patrick Dumas/Eurelios/SPL; 22.2 David Mack/SPL; Chapter 23 satori13/Thinkstock; 23.1 Health Protection Agency/SPL; Chapter 24 Grzegorz Kula/Thinkstock; 24.1 David Hay Jones/SPL; 24.2 Andrew Lambert Photography/SPL; Chapter 25 nikkytok/ Thinkstock; 25.1 Ria Novosti/SPL; 25.7 David J. Green technology/Alamy; 25.11 Martyn F Chillmaid/SPL; 25.14 Cristian Nitu/ iStock; 25.15 sciencephotos/Alamy; 25.26 Simon Belcher/Alamy; Chapter 26 Rainer Plendl/Thinkstock; 26.1 Alex Bartel/SPL; 26.12 Martyn F Chillmaid/SPL; Chapter 27 SERG_AURORA/Thinkstock; 27.1 OMIKRON/SPL; 27.2, 27.8 Andrew Lambert Photography/SPL; 27.14 SPL; Chapter 28 National Geographic Image Collection/Alamy; 28.1 Sean Gallup/ Getty Images; 28.15 Bill Longchore/SPL; Chapter 29 Stockbyte/Thinkstock; 29.1 Tom Bonaventure/Getty Images; 29.3 Adam Gault/SPL; 29.8 Andrew Lambert Photography/SPL; 29.11 Chris Knapton/SPL; 29.14 Greenshoots Communications /Alamy; Chapter 30 Tee-Photolive/Thinkstock; 30.1 PHOTOTAKE INC/Alamy; 30.2 STphotography/Alamy; 30.3 Paul Broadbent/Alamy; 30.5 Volger Steger/SPL; 30.12a Peter Hermes Furian/Alamy; 30.12b, c, d, 30.14 Dept of Physics Imperial College/SPL; 30.13 untraceable; 30.22, 30.23 SPL; 30.24 Professor Doctor Hannes Lichte, used by permission; 30.26 Dr David Wexler coloured by Dr Jeremy Burgess/SPL; 30.27 Dr Tim Evans/SPL; Chapter 31 ktsimage/Thinkstock; 31.1 US Navy/SPL; 31.7 Tom Pietrasik/SPL; Chapter 32 pixologicstudio/Thinkstock; 32.1 Mauro Fermariello/SPL; 32.3a AJ Photo/HOP Americain/SPL; 32.3b Edward Kinsman/SPL; 32.8 Du Cane MedicalImaging ltd/SPL; 32.14 Zephyr/SPL; 32.15, 32.25 GustoImages/SPL; 32.17 Michelle Del Guerico/SPL; 32.20 Scott Camazine/SPL; 32.21 Abk/BSIP/SPL; 32.28 untraceable; 32.29 Geoff Tomkinson/SPL; 32.35 Alfred Pasieka/SPL; 32.36 Sovereign, ISM/SPL; Chapter Practice Test 1 Janka Dharmasena/Thinkstock ; Chapter Practice Test 2 Juliann/Thinkstock

SPL = Science Photo Library

Advice on how to revise for and approach examinations

Tips for revision

Know the syllabus

The Cambridge International AS and A Level Physics syllabus is divided into **sections**; each section is divided into **Learning outcomes**. The advice here refers to both the AS and A Level examinations. It is useful to familiarise yourself with what you are expected to know for examination. You can access a copy of the syllabus on the Cambridge International Examinations website, and you should always refer to the syllabus for the specific year you are entering the examination. Your class notes may contain extra background information to help you understand topics. Use the syllabus to refer to the Learning outcomes and then extract the relevant points from class notes when you make **learning notes** (see below) and **revision notes**.

When planning your revision, allocate your revision time according to:

- the amount of material you need to learn
- how difficult each section has been for you.

Learn the wording of any Learning outcomes that are definitions of scientific terms, and learn any basic proofs that are required in the Learning outcomes – these could form the basis of some questions. Remember to use 'per' in a definition if a ratio is used; for example, **pressure** is defined as **force per unit area,** and not **force on unit area.** However, it is very important not to simply learn the material by rote, but to be able to apply your learning to usual and unusual situations, as exam questions can present you with a situation in a new context.

Make learning notes

Many students learn by repetition. As you learn, you can reduce your class notes:

class notes → **learning notes** → revision notes

Each time you complete a topic, it's helpful to go back through your class notes and organise them in a way that suits your style of learning, to make your learning notes. Learning notes are personal and individual to you and will emphasise anything that you have found difficult.

You can benefit from rewriting class notes using strategies that work for you – for example, with colours, flow diagrams, bullet points or concept maps, in any way that you find helpful and even humorous. Check your notes against the coursebook and leave out, or reduce to a minimum, any points with which you are thoroughly familiar. Include equations.

To make learning notes:

- produce a shortened version of your class notes
- check your class notes against your coursebook
- leave out background, non-syllabus information
- leave out, or reduce to a minimum, points that you know you will never forget; for example, you know from earlier courses, such as the IGCSE®, the properties of magnets and magnetic fields, but a new formula such as $F = BIL \sin\theta$ needs to be emphasised
- avoid long, flowing sentences that contain many points
- use bullet points or short sentences, each containing one or two points that are likely to gain marks in an exam – 'mark points'

Cambridge International AS and A Level Physics © Cambridge University Press 2014

- use scientific terminology correctly and as specifically as you can.

Your teacher may have covered topics in a different order from that in the published syllabus. This may also be a logical order for you. Make your learning notes for each topic on separate pieces of paper to put into your folder in your preferred order.

Making learning notes has benefits:

- repeating and reformatting class notes helps learning
- it is a first stage in revision for the exams
- notes are in your preferred style to optimise learning
- you can assess how well you understand each Learning outcome.

Make revision notes

It is likely that you will be trying to cope with other subjects as the exam time approaches, so you can reduce your notes further to make revision notes:

class notes → learning notes → **revision notes**

Revision notes should contain only the detail necessary to get your very best grade. Think about how you will set out these notes and discuss them with your friends. Make your revision productive by making it interesting and fun. Make notes, revision cards or topic maps. Revision should be an active process, i.e. you should be 'doing things', not just sitting and reading a book. Revise with a friend so that you can test each other, or try explaining the physics of a topic to a friend – as if you were a teacher! Try not to copy directly from the coursebook or syllabus, but put the ideas in your own words, checking them later. At this stage you should incorporate your experience of practice questions and their mark schemes from past papers.

You might try another process. Using the syllabus as your starting point:

1 Write down on a sheet of paper or card each definition, law and equation. Some equations will be given to you on the examination paper, but you must still learn what the letters stand for, and the units.

2 On another sheet or card, write down the title of each important experiment.

3 On a third sheet, write down the title for each important idea that you need to understand (e.g. the difference between AM and FM modulation).

4 On the next sheet, write down any parts of the syllabus you do not understand.

Now, with your sheets or cards labelled 1, 2, 3 or 4, you should:

1 Write out the definitions, laws and equations **from your own memory**. Tick each one off when you are sure you know it and it is the same as in your notes or in the coursebook.

2 Test your knowledge by drawing a clearly labelled diagram, listing the readings taken and explaining the conclusion. Go back to the page in your notes or the coursebook to check.

3 Test your understanding by drawing labelled diagrams and/or brief notes. Go back to the page in your notes or coursebook to check.

4 Seek help! Ask your teacher, or read the coursebook.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Know what to expect for each exam

Before you enter the exam room, you should know what to expect in the exam paper you are facing. Exam papers have different types of questions, including multiple choice questions, structured questions and practical questions. It is important to know how long the exam lasts and how many questions you need to answer. You can find this information in the syllabus.

The table below, based on the latest syllabus, shows the types of questions in each of the Cambridge International Examinations papers.

Paper	Level	Type of paper and questions it contains
1	AS	written; multiple choice questions
2	AS	written; structured questions
3	AS	practical; testing skills such as manipulation, measurement and observation, presentation of data and observations, analysis, conclusions and evaluation
4	A	written; structured questions
5	A	written; testing the practical skills of planning, analysis and evaluation

Apart from multiple choice questions, it is rare for questions to be worded as straightforward questions with question marks. They generally have **command terms**. A list with explanations of the most common command terms that you should understand is in the syllabus. The most common terms are: **state**, **describe**, **explain** and **calculate**. Make sure you know exactly what is required for each command term. For example, there is no point in writing a long explanation of why something happens if the question merely asks you to state what happens.

Use past papers to help you revise

You will find it helpful to look at or practise past exam papers.

You will also need to learn and practise the practical skills that are tested in the AS practical examination and the A Level paper that concentrates on the skills of planning, analysis and evaluation.

Your teacher may have copies of past papers and mark schemes.

Approaching the examinations

Make sure that you know about the data and formulae given at the front of the question paper. You should be so familiar with this data and these formulae that you almost do not need to look them up, but it is better to refresh your memory each time that you use an equation, to check that you have not made a mistake in writing down a formula or have used a wrong power of ten in a constant.

Suggestions on how to approach multiple choice question papers

Understanding multiple choice questions

Each question may be posed on its own or may be preceded by some information. There are four answer options: **A**, **B**, **C** and **D**. You must choose the one that is the correct answer.

Some questions have a simple construction in which the options **A**, **B**, **C** and **D** appear directly below the question. The options could be one or more sentences long. Other questions have a construction in which each option is a row in a headed table.

During the exam

When tackling questions:

- read questions through more than once
- check whether a positive or negative answer is being asked for, e.g. whether the question says 'which of the following is ...?' or, 'is **not** …?'; for example, when asked for an incorrect ray diagram, be careful not to pick a correct diagram as your answer
- underline or circle important information in the stem of the question to help you understand the important points
- never leave a question unanswered; marks are not deducted for incorrect answers
- try to eliminate some of the possible answers if you are not sure of the answer
- write out your working to numerical questions clearly (on the question paper in the space provided near the question) so you can check it later; do not try to do all the working in your head, as you are likely to make mistakes
- if option **A** appears to be correct, read on and check the others before you confirm your choice
- do not look at the pattern of answers on the answer sheet to predict the correct option – you should not expect there to be an equal number of correct option letters in the whole exam; likewise, if the last answer was an **A**, you cannot assume that the next one must be a **B**, **C** or **D**, for example. For questions that assess topics which you find difficult, try to recall the relevant section of your revision notes before looking at the options (**A**, **B**, **C** and **D**) available.

Suggestions on how to approach papers with structured questions

During a 'real' examination is not the time to work out your preferred strategy for coping with structured questions in exam papers. If you can, practise beforehand by using, for example, past papers, and try different exam approaches to see what is best for you.

Exam papers aren't long enough to test your knowledge and understanding of all the Learning outcomes. Ensure that you have revised well and are prepared for your less-favourite topics to turn up in the exam. In both of the structured-question papers, each main question can assess Learning outcomes from **more than one syllabus section**.

Approaching the exam paper

The first question in a structured-question paper is sometimes slightly easier, so it can be completed with relative ease and gives you confidence. However, make sure that you read through all of the questions thoroughly and then decide which question to tackle first. You should monitor the time that you spend on each question throughout any paper and adjust your working accordingly. This will ensure that towards the end of the examination you will have sufficient time to finish. Practise with past papers under timed conditions to develop your timing skills.

Always show your working when answering a question. You must include the method as well as the answer, and show every step in each calculation. It is also true that you should not give up on any question, but should show what you can do, even if you cannot work the question through to the end.

In calculations, it is best to start by writing down the equation that you will use, then substitute the values in another written expression, and finally give the answer. All of these stages should appear. During the calculation do not prematurely round your answer, and always keep and use more significant figures than are given in the question. You can round your answer at the end, but make sure you do so correctly: for example, 1.86 is 1.9 to two significant figures, **not** 1.8. Always use g as $9.81\,\text{m}\,\text{s}^{-2}$ and not $10\,\text{m}\,\text{s}^{-2}$. Check that the units you have used are consistent; for example, if the distance is given in km and the speed is given in $\text{m}\,\text{s}^{-1}$, then you must convert km to m.

When you have obtained the final answer, make sure that:

- it is a sensible and realistic number; for example, if you have worked out the current in a wire to be 3 000 000 A, then this is likely to be wrong and you should look for the mistake
- it has units; look out, in particular, for questions that ask you to give the units of your answer or that do not have the units on the answer line
- it has the correct number of significant figures; this should be the least number of significant figures in the data in the question – this data is often given to two or three significant figures, so if you write your answer down to one or to four or more figures, then you should have a good reason. Think very carefully indeed before you write down all the figures you see on your calculator!

If you have difficulty in giving an answer in writing, then provide a sketch or diagram; this is sometimes much easier than using many words. Diagrams or sketches should always be labelled with as much detail as you can. Try to make your explanations as clear as possible and use technical terms correctly; for example, it is better to say 'resistance is inversely proportional to area' than 'as area increases resistance decreases'. Sketch graphs should have their axes labelled, and the values of any important points on the axes should be shown, particularly the origin.

Some questions require longer written answers. Before you start these questions, look at the number of marks that they are worth. This will help you decide how many important points you should include in your answer. Then it is best to briefly plan the structure of your answer before starting to write. If there are three marks for the question, then include **at least** three different and valid points if you can. You do not need to start your answer by rewriting the question itself because then there may be too little space left for your answer.

Suggestions on how to approach practical papers

Practical examination papers

Read through each question carefully before you start to take any readings. Do not worry if you have never met the experiment before or if it contains unusual apparatus. The question paper itself should tell you what to do, and you must follow the instructions that you are given.

If there is a problem with your apparatus, for example the power supply does not work, then you can ask the supervisor for help. You may lose some marks if the supervisor has to set the apparatus up for you, but not if there was a genuine failure in the apparatus.

The syllabus gives a lot of information on the skills that you will need. Many of these skills are straightforward but may not be clear from the instructions for the examination. For example, you should choose values for the independent variable that:

- cover the largest possible range of readings that can be obtained using the apparatus, or that are consistent with the instructions given
- give, ideally, an even spread of readings across the range.

For example, if you are asked to measure the resistance of a wire for five different lengths and the wire is stretched along a metre rule, you would not choose to measure lengths of 1, 2, 3, 4 and 5 cm or 10, 20, 30, 40 and 100 cm, but you might suggest 20, 40, 60, 80 and 100 cm.

You may be asked to make a table of your readings. It is best to plan this and to draw the columns and headings **before** you take any readings. Then, as you take each reading you can enter it directly into the table. This will save time.

Before you start:

- look through the question to see whether you need a column for a derived quantity as well the columns for the readings themselves
- think about which readings you will repeat.

For example, if you are measuring the resistance of different lengths of wire, you may record the value of current and voltage twice and calculate the resistance for each length twice. This means that the column for resistance has to have three small sections, two for the values of resistance and one for the average. Do not waste time taking readings more than once unless you are asked to do so. Thus the column headings and first row of your table might be:

Length / cm	Reading 1		Reading 2		Resistance / Ω		
	Current / A	Voltage / V	Current / A	Voltage / V	Reading 1	Reading 2	Average
10.0	1.20	3.45	1.22	3.50	2.88	2.87	2.87

Make sure that each column has a quantity and a unit.

All the readings in a particular column should be recorded to the same number of decimal places, corresponding to the precision of the instrument that you are using. In the example above, every current recorded should be to the nearest 0.01 A if that is the smallest division on the instrument. If the current happens to be 1.20, do not just write down 1.2, as there will be other currents such as 1.51 that have a different number of decimal places.

In the example above, notice that the length was written as 10.0 cm, as it is possible to measure to 1 mm on a rule. Of course the uncertainty in a reading can be larger than the smallest division on the instrument and the ammeter readings may be fluctuating by more than 0.01 A. It is usually best to write down what you see on the instrument. Sometimes the readings will differ when repeated, due to random error.

The number of significant figures in any derived quantity (the resistance in the example above) should be equal to the smallest number of significant figures in the data. If your measurements permit, you may also give one more figure than this. In the example above, the resistance could be given to three or four significant figures.

In most practical examinations you will be required to draw a graph to represent the data you have collected. Make sure before the examination that you have a sharp pencil and a long, clear plastic ruler to draw the graph. You should be thoroughly familiar with the details about drawing graphs, in the section on analysing results in Chapter P1 of the coursebook. To summarise:

- each axis must be labelled with the appropriate quantity and unit
- the plotted points should occupy at least half the grid in both the *x* and *y* directions
- the *x*-axis scale should increase positively to the right
- the *y*-axis scale should increase positively upwards
- the scale should have 1, 2 or 5 units to a 2 cm square
- all plotted points should appear on the grid itself and none should be outside of it
- check that your points are plotted accurately
- when drawing a line of best fit, make sure there is a balance of points on either side of the line along all of its length
- when finding a gradient, draw a triangle that covers more than half of the actual graph line
- show all your working when calculating a gradient
- if you are asked for an intercept, you can read it from the graph; if this is not possible, you can take the equation of the straight line and a point on the line to calculate the intercept.

A practical examination will also ask you to calculate uncertainties and suggest improvements. It is important that these improvements relate to the actual experiment and are possible to carry out, although you will not have to alter the apparatus yourself or take further readings.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Written examination of practical skills

The best preparation for a written examination of practical skills is to have practised the skills you need to answer the examination questions on many occasions, by:

- planning a number of experiments yourself
- taking readings according to your own plan
- evaluating the results.

It may help for you to swap plans with a fellow student and for each to try to follow the other's plan. You will see how difficult it is to describe in detail what is to be done.

As the examination time draws closer, practise by, for example, looking at past papers and their mark schemes under the supervision of your teacher. Remember, you need to have developed and practised the skills you need in 'real' experiments. You need the ability to think for yourself, to know what apparatus is available in a physics laboratory and to know how it is used, so that you can take the correct approach to the questions in the exam paper.

There are three main types of question you will face in an examination of this type.

1 Planning question

First read the question through several times and identify the dependent and independent variables. You may have to consider a number of other issues, which may involve:

- choosing what apparatus you will use to solve the problem set in the question
- describing the experimental arrangement
- describing a sequence of instructions to take the readings
- stating and describing how you will keep other variables constant (remember to include sufficient detail)
- describing how the data is processed, often in the form of a graph (remember to state the axes on the graph and the form you expect the graph to take; if you sketch the graph, make sure the sketch is consistent with your description)
- describing relevant safety precautions; give detail and reasons; do not just say, for example, 'do not touch the resistor' but explain that the resistor may be hot.

The question will suggest a structure, and it is sensible to make sure that your answer follows this structure and includes as much detail as you can about each point as you go along. Marks for additional detail are likely to be offered, so you should answer each section with a wealth of detail and supporting material. For example, you could suggest how additional variables are kept constant, give circuit diagrams or describe additional safety precautions.

The diagram of your apparatus is important:

- it should be labelled with as much detail as you can provide
- you should check that it is possible to take the measurements and alter the independent variable with the actual apparatus that you draw
- you should re-check your diagram, imagining that you are another person who is not familiar with the experiment – could they actually take the necessary readings for the experiment?

2 Analysis and evaluation question

Refresh your memory on the important points in the practical examination paper section guidance above concerning significant figures, graphs, gradients and intercepts; they also apply to a written examination of practical skills.

You should also practise using expressions that involve the exponential function and rearranging them to plot an appropriate graph. More help is given on this topic in Chapter P2 of the coursebook.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Remember that the unit for a logarithmic quantity is written slightly strangely, for example ln (d / cm). The logarithm itself does not have a unit.

You will need plenty of practice in using calculated uncertainties to draw error bars and in using these error bars to draw a 'worst acceptable' line as well as a line of best fit. You should draw your 'worst acceptable' line as a broken line or, preferably, clearly label each line as 'best-fit line' or 'worst acceptable line'.

The skills of data analysis involved in this question are ones that you can practise again and again, for example by using past papers, until you are sure that you can:

- complete tables of results using the right headings
- rearrange equations such as: $y = mx + c$, $y = ax^n$ and $y = ae^{kx}$
- plot ordinary and log graphs, knowing how logarithms are taken of each side in the above equations
- show error bars on graphs
- calculate gradients
- use the correct number of significant figures
- draw curved trend lines and tangents when required to do so
- write sensible conclusions.

If you practise using past papers, ask your teacher for the mark scheme or sample answers, or swap with a friend and discuss your answers, to make sure that you have understood all of the points. If you do this, then the examination itself will not be a great surprise and you will show that you have the determination to succeed.

Finally

Some questions in examination papers are likely to seem strange, and you may be unsure of the answer. There is no need to panic, as this is likely to be the same for every student. If you are in doubt, first make sure that you have understood the question, and then make sure that you are describing when the question says 'describe' or applying physics principles when the question says 'explain'. Try to think carefully to yourself about what physics principle is actually involved, and perhaps roughly plan your approach before starting to write the answer. Avoid repeating the question or writing out pages from your notes or from the coursebook, but make it your aim to relate the principle to the question itself. An examination is trying to discover what you really know and can do. If all has gone well during your course and when using the coursebook, you will have ensured that you not only understand these principles but can apply them to a range of questions. If that is the case, then you can be confident of giving a good performance.

The questions, example answers, marks and comments that appear in the Coursebook and this accompanying CD-ROM were written by the authors. In examination, the way in which marks would be awarded to answers may be different.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Introduction to the examination and changes to the syllabus

The Cambridge International Examinations Physics examination – a student's perspective

You should find out from your teacher which examination you will take and when. The following details refer to the latest syllabus, but it is important that you refer to the syllabus for the year in which you are taking the examination, in case details change.

If you are entered for the Advanced Subsidiary (AS) Physics examination, then you will take three papers, Papers 1, 2 and 3, in a single examination series in either June or November, or in March for candidates in India.

After your AS level examination, you may, if you wish, continue to study physics and then take the full Advanced Level Physics qualification. If you do this, your AS level marks in Papers 1, 2 and 3 are carried forward and you take two more papers, Papers 4 and 5, in a single session. Your final grade for the full A level is based on all of the papers you have taken, Papers 1 to 5.

However, you may take the complete Advanced Level Physics qualification in a single examination series. In this case you take all the five papers in one single session. This may not be as hard as you think, as some of the questions in Papers 4 and 5 rely on material taught for the previous papers and you will be much more experienced when answering Papers 1, 2 and 3 if you take them later. On the other hand, there are fewer papers to take in any single series if you take the AS level examination before the final two papers.

The coursebook is divided into two sections to help you prepare for the AS level and the full A level separately, but you will need to know all the material from the AS level when you take Papers 4 and 5.

Details about the papers

Paper	How long is the paper and with how many marks?	What is in the paper?	What is the paper worth for the AS level exam?	What is the paper worth for the full A level exam?
Paper 1 Multiple choice	1 hour and 15 minutes 40 marks	This paper has 40 multiple choice questions, each with four options and based on the AS syllabus. You answer all the questions on an answer grid.	31%	15.5%
Paper 2 AS structured questions	1 hour and 15 minutes 60 marks	This paper has a variable number of structured questions of variable value. You answer all the questions and write on the question paper.	46%	23%

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Paper 3 Practical test	2 hours 40 marks	You carry out practical work under timed conditions, with two experiments from different areas of physics, using the apparatus for each experiment for 1 hour. The experiments test your practical skills rather than theory and may be based on physics not included in the syllabus content. You write on the question paper itself.	23%	11.5%
Paper 4 A level structured questions	2 hours 100 marks	This paper has a variable number of structured questions of variable mark value. All questions are based on the A level syllabus content, but knowledge of AS level content is assumed for this paper. You answer all the questions and write on the question paper.	–	38.5%
Paper 5 Planning, analysis and evaluation	1 hour and 15 minutes 30 marks	This paper has two questions, each worth equal marks and based on the practical skills of planning, analysis and evaluation. You are tested on these skills rather than on your knowledge of any theory; the context of the questions may be outside of the syllabus content. You answer both questions on the question paper itself.	–	11.5%

Changes for the 2016 syllabus

You should make sure that you use the correct syllabus for the session that you are taking the examination. You will find the syllabus on the Cambridge International Examinations (CIE) website at www.cie.org.uk, or you should ask your teacher.

For 2016 examinations, there were a number of changes made to the syllabus. It may be important that you know about these changes.

Changes to the examinations

The main change to the examinations themselves is that in Paper 4 there is now only one section. For examinations taken before 2016, Paper 4 contained two sections, with section B based on applications of physics. Although these applications are largely still in the syllabus, they are now more closely related to the appropriate area of physics. For example, nuclear magnetic resonance imaging (MRI) is now part of the section on magnetic fields, rather than being in a separate applications section called Remote sensing. Questions on these applications may appear at any place within Paper 4.

The second edition of the coursebook now incorporates many of these applications alongside the appropriate topic, rather than as a separate section.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Changes to the syllabus

New material in the 2016 syllabus

The 2016 syllabus includes some new material, with the introduction of extra topics:

- the Doppler shift for a moving source of waves
- basic ideas about fundamental particles such as quarks and neutrinos
- the band theory of energy levels in solids
- the derivation of the Hall voltage
- the derivation and use of the equation $I = nAvq$ for current-carrying conductors
- the conservation of momentum applied to problems in two dimensions.

Minor additions or clarifications

There have been some clarifications and minor additions:

- sources of energy loss in a transformer
- the effect of internal resistance of a source of e.m.f. on the terminal p.d.
- capacitance applied to both isolated conductors and parallel plate capacitors
- the charge on a sphere and the mass of a sphere to be taken at its centre
- thermistor thermometers rather than resistance thermometers
- the definition of the radian
- the introduction of the Boltzmann constant into the kinetic theory of gases.

Material moved from AS to A level

The following have been moved from AS to A level:

- the concept of internal energy and the simple kinetic model for solids, liquids and gases
- the difference in structure and density related to the spacing, ordering and motion of molecules
- Brownian motion
- the temperature characteristic of a thermistor
- thermistors and light-dependent resistors in potential dividers providing a p.d. dependent on temperature and illumination
- simple nuclear reactions and nuclear equations
- the spontaneous and random nature of nuclear decay.

Material removed from the syllabus

The following material has been removed from the syllabus:

- the structure of crystalline and non-crystalline solids
- distinguishing between the processes of melting, boiling and evaporation
- the force–extension graphs for typical ductile, brittle and polymeric materials, including an understanding of ultimate tensile stress
- polarisation as a phenomenon associated with transverse waves
- Millikan's experiment and the evidence for quantisation of charge
- the need for remote sensing in medicine
- the mobile-phone system, the public switched telephone network and the mobile phone handset.

All of these changes have been incorporated, where necessary, into the coursebook.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Answers to SAQs

Chapter 1

1 average speed = $\frac{10\,000}{1577.53} = 6.34\ \text{m s}^{-1}$

2 **a** mm s^{-1}
b mph
c km s^{-1}
d m s^{-1}
e km h^{-1}

3 distance = 12 cm = 120 mm, so
average speed = $\frac{120}{60} = 2.0\ \text{mm s}^{-1}$

4 average speed = $\frac{0.050\ \text{m}}{0.40\ \text{s}} = 0.125\ \text{m s}^{-1}$
$\approx 0.13\ \text{m s}^{-1}$

5 **a** Constant speed
b Increasing speed (accelerating)

6 For example, attach a card to a weight and drop it through a light gate. Alternatively, attach ticker-tape to the falling mass.

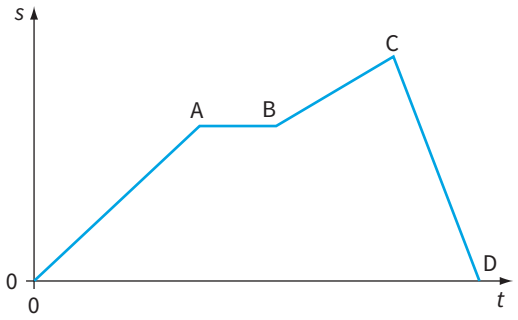
7 **a** Displacement
b Speed
c Velocity
d Distance

8 Distance $d = v \times t = 1500 \times 0.2 = 300$ m.
(Remember the 0.4 s total time is that taken for the sound waves to travel out and be reflected back from the surface of the water.)

9 Time taken for orbit is one year
$= 1 \times 365.25 \times 24 \times 60 \times 60 = 31\,557\,600$ s.
Distance travelled = circumference of orbit
$= 2 \times \pi \times 1.5 \times 10^{11} = 9.425 \times 10^{11}$ m.
So the Earth's speed = $29.9\ \text{km s}^{-1} \approx 30\ \text{km s}^{-1}$.
As the Earth orbits the Sun, its direction of motion keeps changing. Hence its velocity keeps changing. In the course of one year, its displacement is zero, so its average velocity is zero.

10 Sloping sections: bus moving. Horizontal sections: bus stationary (e.g. at bus stops).

11 OA: constant speed.
AB: stationary.
BC: reduced constant speed.
CD: running back to gate.

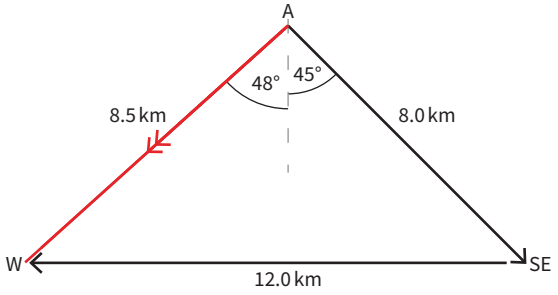


12 **a** $85\ \text{m s}^{-1}$
b Graph is a straight line through the origin, with gradient = $85\ \text{m s}^{-1}$.

13 **a** Graph is a straight line for the first 3 h; then less steep for the last hour.
b Car's speed in first three hours = $23\ \text{km h}^{-1}$
c Car's average speed in first four hours
$= \frac{84}{4} = 21\ \text{km h}^{-1}$

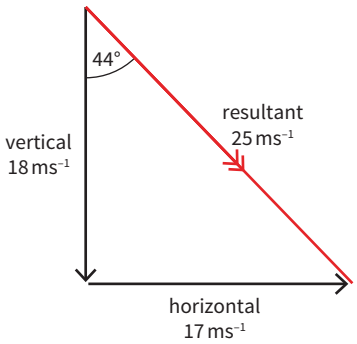
14 **a** Total distance travelled = 3.0 + 4.0 = 7.0 km
b, c The two parts of the journey are at 90° to each other, so the resultant displacement is given by Pythagoras' theorem.
displacement^2 $= 3.0^2 + 4.0^2 = 25.0$, so
displacement $= 5.0$ km
angle $= \tan^{-1}\left(\frac{4.0}{3.0}\right) = 53°$ E of N (or 37° N of E)

15 **a, b** 8.5 km; 48° W of S



16 Swimmer aims directly across river; river flows at right angles to where she aims. So resultant velocity is given by geometry:
magnitude$^2 = 2.0^2 + 0.8^2 = 4.64$ so
magnitude $= \sqrt{4.64} = 2.154 \approx 2.2\,\text{m s}^{-1}$
direction $= \tan^{-1}\left(\frac{0.8}{2}\right) \approx 22°$ to the direct route (68° to the river bank)

17 **a**



b $17.3\,\text{m s}^{-1} \approx 17\,\text{m s}^{-1}$

c 43.9° ≈ 44° to the vertical

Answers to SAQs

Chapter 2

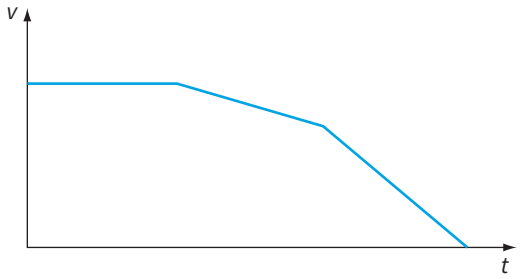
1 Change in velocity $\Delta v = (18 - 0) = 18\ \text{m s}^{-1}$
Time taken $\Delta t = 6.0\ \text{s}$
So acceleration $a = \frac{\Delta v}{\Delta t} = \frac{18}{6.0} = 3.0\ \text{m s}^{-2}$

2 Change in velocity $\Delta v = (11 - 23) = -12\ \text{m s}^{-1}$
Time taken $\Delta t = 20\ \text{s}$
So acceleration $a = \frac{\Delta v}{\Delta t} = \frac{-12}{20} = -0.60\ \text{m s}^{-2}$
The magnitude of the deceleration is $0.60\ \text{m s}^{-2}$

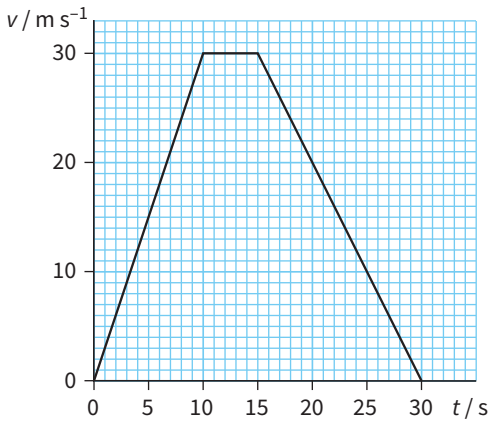
3 **a** Rearranging the equation $a = \frac{v - u}{t}$ gives $v = u + at$
So after 1 s, $v = 0.0 + (9.81 \times 1) = 9.81 \approx 9.8\ \text{m s}^{-1}$

b After 3 s, $v = 0.0 + (9.81 \times 3)$
$= 29.4\ \text{m s}^{-1} \approx 29\ \text{m s}^{-1}$

4



5 **a**



b, c During first 10 s,
acceleration $a = \frac{\Delta v}{\Delta t} = \frac{30}{10} = 3.0\ \text{m s}^{-2}$

d During last 15 s,
acceleration $a = \frac{\Delta v}{\Delta t} = \frac{(0 - 30)}{(30 - 15)} = \frac{-30}{15} = -2.0\ \text{m s}^{-2}$

e From area under graph: 525 m

6



Dots evenly spaced, then getting steadily closer together.

7 If l_1 = length of first section of interrupt card,
t_1 = time when first section enters light gate,
t_2 = time when first section exits light gate,
l_2 = length of second section of interrupt card,
t_3 = time when second section enters light gate,
t_4 = time when second section exits light gate,
then:
initial velocity $u = \frac{l_1}{t_2 - t_1} = \frac{0.05}{0.20 - 0.0} = 0.25\ \text{m s}^{-1}$
final velocity $u = \frac{l_2}{t_4 - t_3} = \frac{0.05}{0.35 - 0.30} = 1.0\ \text{m s}^{-1}$
$\Delta t = t_3 - t_1 = 0.30 - 0.0 = 0.30\ \text{s}$
So acceleration $a = \frac{\Delta v}{\Delta t} = \frac{1.0 - 0.25}{0.30} = 2.5\ \text{m s}^{-2}$

8 For first ticker-tape section, length $l_1 = 10\ \text{cm}$,
time taken $t_1 = 5 \times 0.02 = 0.10\ \text{s}$
so initial velocity $u = \frac{l_1}{t_1} = \frac{0.10}{0.10} = 1.0\ \text{m s}^{-1}$
For second ticker-tape section, length $l_2 = 16\ \text{cm}$, time taken $t_2 = 5 \times 0.02 = 0.10\ \text{s}$
so final velocity $v = \frac{l_2}{t_2} = \frac{0.16}{0.10} = 1.6\ \text{m s}^{-1}$
Sections of tape are adjacent, so time between start of first section and start of final section, Δt = time taken by first section
$= 5 \times 0.02 = 0.10\ \text{s}$
so acceleration $a = \frac{\Delta v}{\Delta t} = \frac{1.6 - 1.0}{0.10} = 6.0\ \text{m s}^{-2}$

9 **a** We know u, a and t and we want to know v, so we use the equation
velocity $v = u + at = 0.0 + (2.0 \times 10) = 20\ \text{m s}^{-1}$

b We know u, a and t and we want to know s, so we use the equation
distance $s = ut + \frac{1}{2}at^2 = 0.0 + \frac{1}{2} \times 2.0 \times 10 \times 10$
$= 100\ \text{m}$

c We know u, v and a and we want to know t, so we rearrange the equation $v = u + at$ so that
time $t = \frac{v - u}{a} = \frac{24 - 0}{2.0} = 12\ \text{s}$

10 **a** We know u, v and t and we want to know a, so we use the equation
acceleration, $a = \frac{v - u}{t} = \frac{20 - 4.0}{100} = 0.16\ \text{m s}^{-2}$

b Average velocity, $v_{\text{avg}} = \frac{v + u}{2} = \frac{20 + 4.0}{2} = 12\ \text{m s}^{-1}$

c We could use $s = ut + \frac{1}{2}at^2$ but given that we have worked out the average speed, it is simpler to use
distance, $s = v_{avg} \times t = 12 \times 100 = 1200\,\text{m}$

11 We know u, a and s and we want to know v, so we use the equation
$v^2 = u^2 + 2as$
So final velocity $v = \sqrt{u^2 + 2as}$
$= \sqrt{(8.0)^2 + 2 \times 1.0 \times 18} = \sqrt{100} = 10\,\text{m s}^{-1}$

12 We know u, v and a and we want to know s, so we rearrange the equation $v^2 = u^2 + 2as$, so that distance
$s = \frac{v^2 - u^2}{2a} = \frac{(0)^2 - (30)^2}{2 \times (-7)} = \frac{900}{14} = 64.3\,\text{m} \approx 64\,\text{m}$

13 We know v, a and s and we want to know u, so we rearrange the equation $v^2 = u^2 + 2as$ into $u^2 = v^2 - 2as$, so initial speed
$u = \sqrt{v^2 - 2as} = \sqrt{(0.0)^2 + 2 \times (-6.5) \times 50} = \sqrt{650}$
$= 25.5\,\text{m s}^{-1}$
This is just over the speed limit.

14 **a** $t = 7.5\,\text{s}$; $v = 220\,\text{m s}^{-1}$

b Draw a tangent to the curve at point P. Read off two sets of values from the tangent to find the gradient. For example:
at time $t_1 = 0\,\text{s}$, $v_1 \approx 60\,\text{m s}^{-1}$
at time $t_2 = 12\,\text{s}$, $v_2 \approx 300\,\text{m s}^{-1}$
So, approximately, acceleration
$a = \frac{\Delta v}{\Delta t} = \frac{300 - 60}{12 - 0} = 20\,\text{m s}^{-2}$

15 **a** The car is slowing down with constant (uniform) deceleration.

b Initial velocity $v_1 = 20\,\text{m s}^{-1}$
final velocity $v_2 = 8\,\text{m s}^{-1}$

c Acceleration $a = \frac{\Delta v}{\Delta t} = \frac{8 - 20}{30 - 0} = -0.40\,\text{m s}^{-2}$

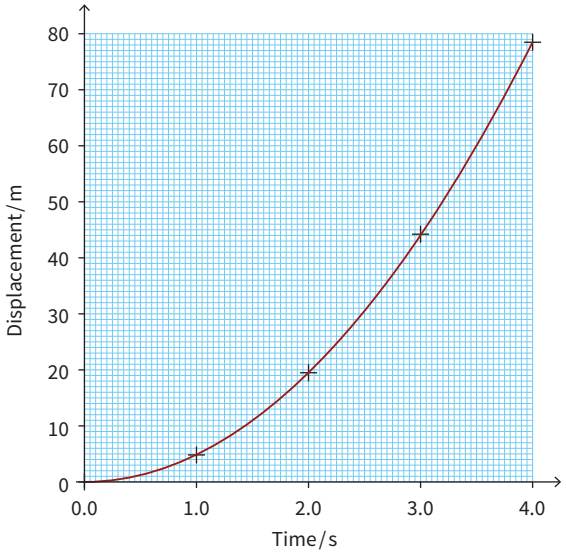
d Displacement of car = area under graph
= (area of rectangle with side $8\,\text{m s}^{-1}$ and length 30 s) + (area of triangle with side $12\,\text{m s}^{-1}$ and base 30 s)
$= (8 \times 30) + \left(\frac{1}{2} \times 12 \times 30\right) = 420\,\text{m}$

e Displacement of car, $s = ut + \frac{1}{2}at^2$
$= (20 \times 30) + \left(\frac{1}{2} \times (-0.40) \times 30 \times 30\right) = 600 - 180$
$= 420\,\text{m}$

16 **a** Calculate distance fallen for each time using $s = ut + \frac{1}{2}at^2$, with $u = 0$.

Time / s	0	1.0	2.0	3.0	4.0
Displacement / m	0	4.9	19.6	44.1	78.5

b Graph is a parabola through the origin.



c In 2.5 s, stone falls 30.6 m ≈ 31 m. Check using $s = ut + \frac{1}{2}at^2 = 0 + \left(\frac{1}{2} \times 9.81 \times 2.5 \times 2.5\right) \approx 31\,\text{m}$

d Time taken = 2.86 s ≈ 2.9 s
Check by rearranging , remembering that $u = 0$, so that time $t = 2.86\,\text{s} \approx 2.9\,\text{s}$

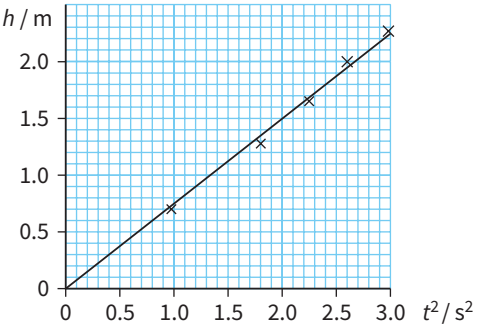
17 **a** We know s and a, and that $u = 0$, and we need to find t.
Rearrange $s = ut + \frac{1}{2}at^2$, remembering that $u = 0$, so that time $t = \sqrt{\frac{2s}{a}} = \sqrt{\frac{2 \times 0.8}{9.81}} \approx 0.40\,\text{s}$

b We know s and a, and that $u = 0$, and we need to find v.
Use $v^2 = u^2 + 2as$ so that impact velocity
$v = \sqrt{u^2 + 2as} = \sqrt{(0)^2 + 2 \times 9.81 \times 0.8} = \sqrt{15.7}$
$\approx 4.0\,\text{m s}^{-1}$

18 **a** Using the method in the worked example, calculate the average speed of the steel ball
$= \frac{s}{t} = \frac{2.10}{0.67} = 3.134\,\text{m s}^{-1}$
Then find the values of v and u:
final speed, $v = 2 \times 3.134\,\text{m s}^{-1} = 6.268\,\text{m s}^{-1}$
initial speed, $u = 0.0\,\text{m s}^{-1}$
Substitute these values into the equation for acceleration
$a = \frac{v - u}{t} = \frac{6.268}{0.67} = 9.355\,\text{m s}^{-2} \approx 9.4\,\text{m s}^{-2}$

b Air resistance; delay in release of ball.

19 **a**



b Because $s = \frac{1}{2}at^2$ the gradient is g^2, the acceleration of free fall, $g \approx 1.6\,\text{m s}^{-2}$

c This object is not falling on the Earth, perhaps on the Moon.

20 Drop an object towards the sensor, but take care not to break it. A better method is to use a sloping ramp with a trolley; gradually increase the angle of slope. Deduce the value of the acceleration when the ramp is vertical.

21 **a** $F_x = 17.3\,\text{N} \approx 17\,\text{N}$; $F_y \approx 10\,\text{N}$

b $v_x = 1.7\,\text{m s}^{-1}$; $v_y = -4.7\,\text{m s}^{-1}$

c $a_x = -5.2\,\text{m s}^{-2}$; $a_y = -3.0\,\text{m s}^{-2}$

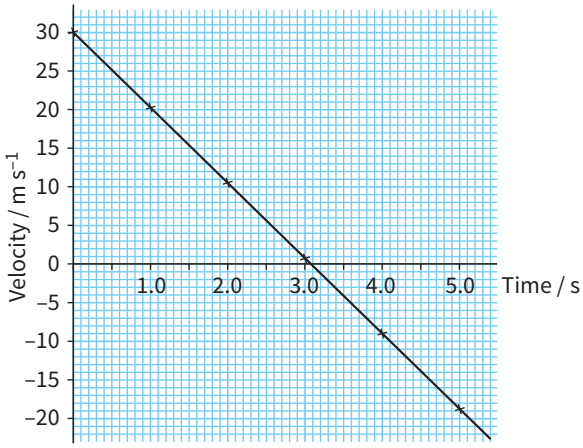
d $F_x = 77.3\,\text{N} \approx 77\,\text{N}$; $F_y = 20.7\,\text{N} \approx 21\,\text{N}$

22 The stone's displacement now is $s = -25\,\text{m}$.
Substituting in $s = ut + \frac{1}{2}at^2$ gives
$-25 = 20t + \frac{1}{2} \times (-9.81) \times t^2$
So $4.9t^2 - 20t - 25 = 0$ or approximately
$5t^2 - 20t - 25 = 0$, which can be simplified to
$t^2 - 4t - 5 = (t - 5)(t + 1) = 0$.
So time taken to reach the foot of the cliff = 5 s (i.e. 1 s more). Accurate answer is 5.08 ≈ 5.1 s.
In solving the quadratic equation, you will have found a second solution, $t = -1$ s. Obviously, the stone could not take a negative time to reach the foot of the cliff. However, this solution does have a meaning: it tells us that, if the stone had been thrown upwards from the foot of the cliff at the correct speed, it would have been travelling upwards at $20\,\text{m s}^{-1}$ as it passed the top of the cliff at $t = 0$ s.

23 **a** Use $v = u + at$ to calculate v, remembering that $a = -9.81\,\text{m s}^{-2}$

Velocity / m s^{-1}	30	20.19	10.38	0.57	−9.24	−19.05
Time / s	0	1.0	2.0	3.0	4.0	5.0

b



c 3.1 s

24 **a** Horizontal speed remains constant after being thrown (ignoring air resistance), so horizontal velocity $= \frac{s}{t} = \frac{12.0}{4.0} = 3.0\,\text{m s}^{-1}$

b For vertical distance, use $s = ut + \frac{1}{2}at^2$, remembering that $u = 0$.
$s = ut + \frac{1}{2}at^2 = 0 + \frac{1}{2} \times (-9.81) \times 4.0 \times 4.0$
$= -78.5$ m, so height of cliff is 78.5 m.

25 **a** Vertical component of velocity $= 8 \times \sin 40°$
$= 5.14 \approx 5.1\,\text{m s}^{-1}$

b Vertical component of velocity $= 0\,\text{m s}^{-1}$

c Rearrange $v = u + at$, so that time $t = \frac{v - u}{a}$
$= \frac{0 - 5.14}{-9.81} = 0.524 \approx 0.52$ s

d Horizontal component of velocity $= 8 \times \cos 40°$
$= 6.13 \approx 6.1\,\text{m s}^{-1}$

e Assume horizontal component of velocity is constant and use
distance $s = ut + \frac{1}{2}at^2 = 6.1 \times 0.52 + 0$
$= 3.21 \approx 3.2$ m

26 First calculate the time taken for the projectile to return to the ground.
Initial vertical velocity,
$u_{ver} = 40 \times \sin 45° = 28.3\,\text{m s}^{-1}$
We know the vertical distance travelled when the projectile hits the ground = 0 m.
So rearrange $s = ut + \frac{1}{2}at^2$ to find t.
$0 = 28.3t + \frac{1}{2} \times 9.81t^2 = 28.3t + 4.905t^2$
So $t = 0$ (when the projectile is launched) or $t = 5.77$ s (when it returns to the ground).
Assume horizontal velocity is constant,
$u_{hor} = 40 \times \cos 45° = 28.3\,\text{m s}^{-1}$
So horizontal distance
$s = ut = 28.3 \times 5.77 = 163\,\text{m} \approx 160\,\text{m}$

Answers to SAQs

Chapter 3

1 Force, $F = ma = 800 \times 2.0 = 1600\ \text{N}$

2 Acceleration, $a = \frac{F}{m} = \frac{200\,000}{5000} = 40\ \text{m s}^{-2}$

3 Acceleration, $a = \frac{F}{m} = \frac{200}{(60 + 40)} = 2.0\ \text{m s}^{-2}$
Bike has zero initial velocity, so final velocity $v = at = 2.0 \times 5.0 = 10\ \text{m s}^{-1}$

4 Apples vary in mass; the acceleration due to gravity varies from place to place.

5 **a** Pressure = $\frac{F}{A}$, $F = ma$, so pressure has base units $\frac{\text{kg m s}^{-2}}{\text{m}^2} = \text{kg m}^{-1}\text{ s}^{-2}$

b Energy = force × distance so energy has base units $\text{kg m}^2\text{ s}^{-2}$

c Density = $\frac{\text{mass}}{\text{volume}}$, so density has base units kg m^{-3}

6 **a** Base units of pressure = $\text{kg m}^{-1}\text{ s}^{-2}$
Base units of ρgh = $(\text{kg m}^{-3}) \times (\text{m s}^{-2}) \times (\text{m})$
= $\text{kg m}^{-1}\text{ s}^{-2}$
Since the base units are the same, the equation is homogeneous.

b Base unit of speed × time = $(\text{m s}^{-1}) \times (\text{s}) = \text{m}$
Base unit of $at^2 = (\text{m s}^{-2}) \times (\text{s}^2) = \text{m}$
Since both terms have base unit m, the base unit of distance, the equation is homogeneous.

7 Sides are 27.5 cm and 21.8 cm (note: your coursebook may have a slightly different trimmed page size)
Area = 27.5×21.8
= $599.5 \approx 600\ \text{cm}^2$ to 3 sig. figs
= $0.0600\ \text{m}^2$ to 3 sig. figs

8 **a** $6 \times 10^{-8}\ \text{A}$

b $5 \times 10^{8}\ \text{W}$

c $20 = 2 \times 10^{1}\ \text{m}$

9 Estimated masses are shown in brackets. Note that g has been rounded appropriately to $10\ \text{m s}^{-2}$ in these estimates.

a (1.0 kg), so weight = $mg = 1.0 \times 10 = 10\ \text{N}$

b (1.0 kg), so weight = 10 N

c (60 kg), so weight = 600 N

d (0.025 kg), so weight = 0.25 N

e (40 000 kg), so weight = 400 000 N

10 The greater the mass of the car, the greater the force needed to slow it down with a given deceleration. For large cars, it is less demanding on the driver if the engine supplies some of the force needed to brake the car.

11 Due to inertia, the driver continues to move forward, although the car stops. A seat belt provides the force needed to overcome this inertia.

12 The large one; its weight is greater, so it reaches a greater speed before air resistance is sufficient to equal its weight.

13 **a** Lubricate the skis to reduce friction.

b Wear tight-fitting, smooth clothing to reduce air resistance.

c Develop powerful muscles to provide a large forward force.

d The steeper the slope the better, to maximise the effect of gravity.

14 **a** The lighter one: lower terminal velocity.

b Turn head-first and pull in his arms and legs to produce a streamlined shape and reduce air resistance.

15 **a** Upthrust

b Friction

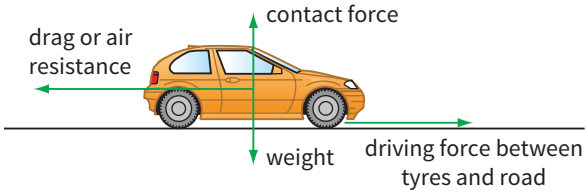
c Weight (= force of gravity)

d Contact force (normal reaction)

e Tension

f Drag

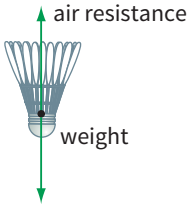
16



17



a Going up



b Going down

18 **a** Force up on your foot and down on the foot that you tread on. Both forces are contact forces (normal reactions).

b Force backwards on the car and forwards on the wall. Both forces are contact forces (normal reactions).

c Backwards force on car and forwards force on ground. Both forces are frictional forces.

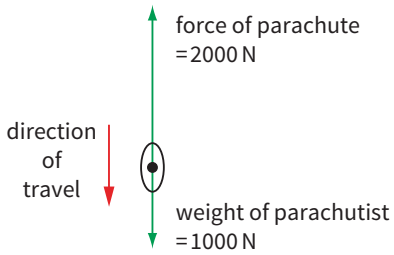
d Upwards force on ball and downwards force on your hand. Both forces are contact forces (normal reactions).

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Answers to SAQs

Chapter 4

1 a



b Force upwards = 2000 – 1000
= 1000 N upwards

c She will accelerate upwards (i.e. decelerate).

2 a Yes, the ship is in equilibrium, because it travels at a constant velocity (not accelerating, so no resultant force acting on it).

b Upthrust is equal and opposite to weight of boat, as it is floating, so = 1000 kN

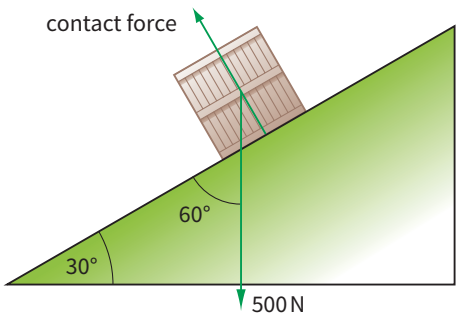
c Because the velocity is constant, we know that the drag is equal and opposite to the force of the engines, so = 50 kN

3 a Vertical component of force
= weight – upthrust
= 2.5 – 0.5 = 2.0 N downwards
Horizontal component of force = 1.5 N
So resultant force is obtained from
$R^2 = (2.0)^2 + (1.5)^2 = 6.25$
So $R = 2.5$ N
Angle = $\tan^{-1}\frac{1.5}{2.0} = 37°$ to vertical.

b No, there is a net force acting upon it.

4 With rope horizontal, the force pulling the box is F. With the rope at an angle θ to the horizontal, the horizontal component ($= F\cos\theta$) is less, since $\cos\theta$ is less than 1.

5 a



b Component of weight down slope
= 500 sin 30° = 250 N

c The contact force of the slope is a normal reaction, so it is at 90° to the slope.

d Friction; up the slope.

6 a Component of acceleration parallel to slope
$= 9.81 \times \sin 25° = 4.1\ \text{m s}^{-2}$

b Net force down slope = 40 × 9.81 × sin25° – 80
= 85.8
acceleration = $\frac{85.8}{40} = 2.1\ \text{m s}^{-2}$

7 a sum of clockwise moments = sum of anticlockwise moments
$400 \times 0.20 = F \times 1.20$
So force required is $F = \frac{400 \times 0.20}{1.20} = 67$ N

b sum of clockwise moments = sum of anticlockwise moments
$400 \times 0.20 = F \times 0.50$
So force on legs of wheelbarrow is
$F = \frac{400 \times 0.20}{0.50} = 160$ N

8 a Remember that weight = mg, and that the acceleration g is the same for all the masses; in our moments equation, g cancels out from both sides.
sum of clockwise moments = sum of anticlockwise moments
$(100 \times 30) + (10 \times 45) = M \times 20$
so mass $M = \frac{3000 + 450}{20} = 172.5 \approx 173$ g

b By this method, weighing could be carried out with a limited selection of relatively small masses.

9 a, b $F_1 = 0$ N m
$F_2 = 10 \times 0.25 = 2.5$ N m clockwise
$F_3 = 10 \sin 30° \times 0.50 = 2.5$ N m clockwise
$F_4 = 5 \times 1.0 = 5$ N m anticlockwise

c Sum of clockwise moments = 2.5 + 2.5 = 5 N
= sum of anticlockwise moments. So, yes, the moments are balanced.

10 Torque = force × radius, so
force = $\frac{\text{torque}}{\text{radius}} = \frac{137}{0.18} = 761\ \text{N} \approx 760\ \text{N}$

Answers to SAQs

Chapter 5

1 **a** Yes, work done against friction.

b Yes, gravity does work in making you go faster.

c No, because the stone remains at a constant distance from the centre of the circle.

d No, because you do not move.

2 Work done, $W = F \times s = mg \times s$
$= 70 \times 9.81 \times 2.5 = 1720 \approx 1700$ J

3 **a** Work done, $W = F \times s = 10 \times 250 = 2500$ J

b 2500 J (ignoring work done against air resistance)

4 The crane does work to overcome the downward force of gravity, so all that matters is the vertical height through which the crane lifts the load.
Work done, $W = F \times s = 500 \times 40$
$= 20\,000$ J $= 20$ kJ

5 Work done by force up slope,
$W = F \times s = 100 \times 0.50 = 50$ J
Work done by contact force = 0 J
Work done by force down slope,
$W = F \times s = -(30 \times 0.50) = -15$ J
(note the movement is in the opposite direction to the force)
Work done by gravity,
$W = F \times s = -(100 \sin 45° \times 0.50) = -35$ J

6 Work done, $W = p\Delta v = 1.0 \times 10^5 \times 0.002 = 200$ J

7 Change in gravitational potential energy,
$E_p = mgh = 52 \times 9.81 \times 2.5 = 1275$ J ≈ 1300 J

8 Change in gravitational potential energy,
$E_p = mgh = 100 \times 9.81 \times 5500$
$= 5400$ kJ or 5.4 MJ

9 **a** Elastic potential energy.

b Work is done on the magnets in pulling them apart. The potential energy transferred to the magnets has come from chemical energy in the student.

10 Kinetic energy of car,
$E_k = \frac{1}{2}mv^2 = \frac{1}{2} \times 500 \times (15)^2 = 56$ kJ
Kinetic energy of motorcycle,
$E_k = \frac{1}{2}mv^2 = \frac{1}{2} \times 250 \times (30)^2 = 156$ kJ
The motorcycle has more k.e.

11 Change in kinetic energy
= k.e. before – k.e. after:
$= \frac{1}{2} \times 0.200 \times (15.8)^2 - \frac{1}{2} \times 0.200 \times (12.2)^2$
$= 25 - 15 = 10$ J

12 Calculate the loss in g.p.e. as the sphere falls from its highest position.
$E_p = mgh = 10.0 \times 9.81 \times 0.15 = 14.7$ J
The gain in the sphere's k.e. is 14.7 J. Using this to calculate the sphere's speed,
$E_k = \frac{1}{2}mv^2 = \frac{1}{2} \times 10.0 \times v^2 = 14.7$ J
So $v^2 = \frac{14.7}{5.0} = 2.94\ \text{m}^2\,\text{s}^{-2}$ and $v = 1.7\ \text{m}\,\text{s}^{-1}$,
the same as for the sphere of mass 5.0 kg; the result is unchanged for any value of mass.

13 Reduction in gravitational potential energy,
$\Delta E_p = mg\Delta h = 80\,000 \times 9.81 \times (10\,000 - 1000)$
$= 7.1 \times 10^9$ J
This energy becomes increased kinetic energy of air molecules – the air temperature rises.

14 We can say that the change in g.p.e. from the top of the jump to the point just before she enters the water is equal to the change in kinetic energy, so that: $mgh = \frac{1}{2}mv^2$
This means that her speed just before the point where she enters the water is
$v = \sqrt{2gh} = \sqrt{2 \times 9.81 \times 10} = \sqrt{196} = 14\ \text{m}\,\text{s}^{-1}$

15 **a** Kinetic energy as stone reaches foot of cliff,
$E_k = \frac{1}{2}mv^2 = \frac{1}{2} \times m \times (38)^2 = m \times 722$ J
Gravitational potential energy change,
$E_p = mgh = m \times 9.81 \times 80 = m \times 785$ J
So proportion of g.p.e. converted to k.e.
$\frac{\text{k.e.}}{\text{g.p.e.}} = \frac{722}{785} = 0.92 = 92\%$

b The rest of the stone's initial energy is converted to heat (because work is done against air resistance).

16 Rearranging power = $\frac{\text{work done}}{\text{time taken}}$ gives
work done = power × time taken
= 50 000 W × 60 s = 3.0×10^6 J (or 3.0 MJ)

17 Power = $\frac{W}{t} = \frac{4\,200\,000}{60}$ = 70 000 W (or 70 kW)

18 **a** Work done in one second,
$W = Fs$ = 700 × 40 = 28 000 J (28 kJ)

b 28 kW

19 Work done against gravity,
$W = Fs$ = weight × height moved
$= mgh$ = 55 × 9.81 × 28 × 0.20 = 3020 J
So her useful power = $\frac{W}{t} = \frac{3020}{5.4}$ = 560 W

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Answers to SAQs

Chapter 6

1 **a** Ball B has greater mass.

b Trolley B has greater mass.

2 **a** Momentum, $p = mv = 0.50 \times 20 = 10\ \text{kg m s}^{-1}$

b Momentum, $p = mv = 25\,000 \times 20$
$= 5.0 \times 10^5\ \text{kg m s}^{-1}$

c Momentum, $p = mv = 9.1 \times 10^{-31} \times 2.0 \times 10^7$
$= 1.82 \times 10^{-23}$
$\approx 1.8 \times 10^{-23}\ \text{kg m s}^{-1}$

3 Momentum before
$= m_A u_A + m_B u_B = (0.50 \times 2.0) + (0.50 \times (-3.0))$
$= 1.0 - 1.5 = -0.5\ \text{kg m s}^{-1}$ (i.e. to the left)
Momentum after $= m_A v_A + m_B v_B$
$= (0.50 \times (-2.0)) + (0.50 \times 1.0) = -1.0 + 0.5$
$= -0.5\ \text{kg m s}^{-1}$ (i.e. to the left)

4

Type of collision	perfectly elastic	inelastic
Momentum	conserved	conserved
Kinetic energy	conserved	not conserved
Total energy	conserved	conserved

5 **a** Before collision:
momentum of ball A,
$p_A = m_A u_A = 4.0 \times 2.5 = +10\ \text{kg m s}^{-1}$
momentum of ball B,
$p_B = m_B u_B = 4.0 \times (-1.5) = -6\ \text{kg m s}^{-1}$

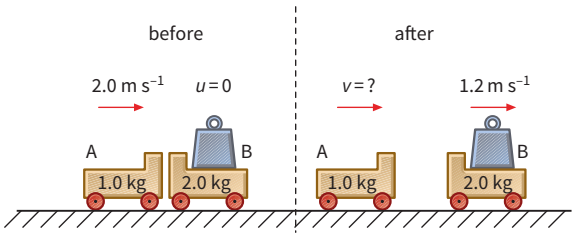
b After collision:
momentum of ball A,
$p_A = m_A u_A = 4.0 \times (-1.5) = -6\ \text{kg m s}^{-1}$
momentum of ball B,
$p_B = m_B u_B = 4.0 \times 2.5 = +10\ \text{kg m s}^{-1}$

c Total momentum before collision
$= p_A + p_B = +4\ \text{kg m s}^{-1}$
Total momentum after collision
$= p_A + p_B = +4\ \text{kg m s}^{-1}$
So momentum is conserved.

d Kinetic energy before collision
$= \frac{1}{2} m_A u_A^2 + \frac{1}{2} m_B u_B^2$
$= \left(\frac{1}{2} \times 4.0 \times (2.5)^2\right) + \left(\frac{1}{2} \times 4.0 \times (1.5)^2\right)$
$= 12.5 + 4.5 = 17\ \text{J}$
Kinetic energy after collision
$= \frac{1}{2} m_A v_A^2 + \frac{1}{2} m_B v_B^2$
$= \left(\frac{1}{2} \times 4.0 \times (1.5)^2\right) + \left(\frac{1}{2} \times 4.0 \times (2.5)^2\right)$
$= 4.5 + 12.5 = 17\ \text{J}$
So k.e. before = k.e. after.

e Relative speed before $= 2.5 - (-1.5) = 4.0\ \text{m s}^{-1}$
Relative speed after $= 2.5 - (-1.5) = 4.0\ \text{m s}^{-1}$

6 **a**



b Call 1.0 kg trolley A.
Call 2.0 kg trolley B.
Conservation of momentum means:
momentum before collision
= momentum after collision
So $m_A u_A + m_B u_B = m_A v_A + m_B v_B$
Rearrange to find speed of first trolley after collision, v_A.
$m_A u_A + m_B u_B - m_B v_B = m_A v_A$

$$v_A = \frac{m_A u_A + m_B u_B - m_B v_B}{m_A}$$

$$= \frac{(1.0 \times 2.0) + (2.0 \times 0.0) - (2.0 \times 1.2)}{1.0}$$

$= -0.40\ \text{m s}^{-1}$
The minus sign indicates that the first trolley reverses direction.

7 **a** If you consider the star to be stationary before exploding, the star has zero momentum. After the explosion, matter flies off in all directions – equal amounts of momentum are created in all directions, so their (vector) sum is zero. Momentum is conserved.

b You give downward momentum to the Earth; as you slow down, so does the Earth; as you start to fall back down, the Earth starts to 'fall' back up towards you. At all times, your momentum is equal and opposite to that of the Earth, so combined momentum is zero, i.e. conserved.

8 If u = initial velocity and v = final velocity, change in momentum of ball

$\Delta p = mv - mu = m(v - u) = 0.40 \times (1.5 - (-1.2))$
$= 1.08\ \text{kg m s}^{-1}$

Change in k.e.

$\Delta E_k = \frac{1}{2}mv^2 - \frac{1}{2}mu^2 = \frac{1}{2}m(v^2 - u^2)$
$= \frac{1}{2} \times 0.40 \times ((1.5)^2 - (1.2)^2) = -0.162\ \text{J}$

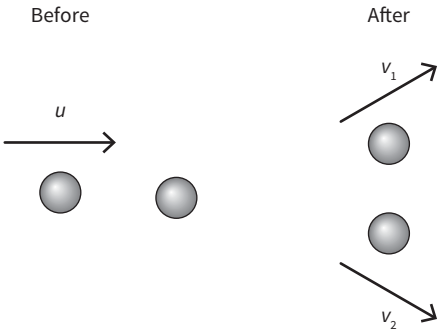
The wall has gained momentum. The ball has lost kinetic energy, which has become thermal energy (heat) of the ball and air.

9 Consider the movement of the moving ball before the collision. If we take its direction of movement as the x-axis, and the perpendicular to its direction of movement as the y-axis, then compare before and after the collision.

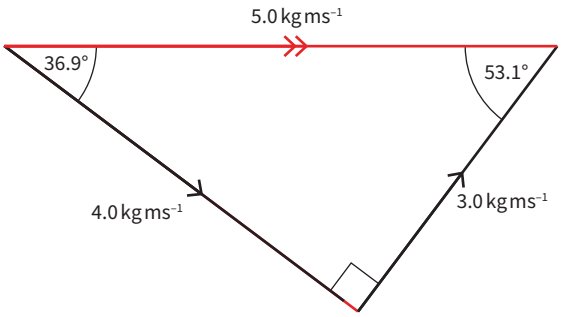
Before collision: component of momentum along x-axis only; no component along y-axis.

After collision: the second ball has a component of momentum along the y-axis (as it moves away at an angle to the x-axis).

Therefore, to conserve momentum along the y-axis, after the collision the first ball must also have an equal and opposite component of momentum along the y-axis. So the first ball must change direction.



10



11 Consider momentum changes in the y-direction.

Before collision:
momentum = 0

After collision:
component of momentum of particle 1
$= 2.40 \sin 60° = 2.08\ \text{kg m s}^{-1}$ upwards
component of momentum of particle 2
$= 2.40 \sin 60° = 2.08\ \text{kg m s}^{-1}$ downwards
These components are equal and opposite and hence their sum is zero.

Consider momentum changes in the x-direction.

Before collision:
momentum $= 2.40\ \text{kg m s}^{-1}$ to the right

After collision:
component of momentum of particle 1
$= 2.40 \cos 60° = 1.20\ \text{kg m s}^{-1}$ to the right
component of momentum of particle 2
$= 2.40 \cos 60° = 1.20\ \text{kg m s}^{-1}$ to the right
Total momentum
$= 1.20 + 1.20 = 2.40\ \text{kg m s}^{-1}$ to the right

Hence momentum is conserved in both the x- and y-directions, so total momentum is conserved.

12 **a** Component of velocity of first ball in x-direction $= 1.00\ \text{m s}^{-1}$
Component of velocity of first ball in y-direction = 0

b Assume that each ball has mass m and that the second ball has velocity V at an angle θ to the x-direction (with components V_x in the x-direction and V_y in the y-direction)

Consider momentum conservation in the x-direction:
$m \times 1.00 = m \times 0.80 \cos 20° + m \times V_x$
Cancel m from all terms, so
$V_x = 1.00 - 0.80 \cos 20° = 0.25\ \text{m s}^{-1}$

Consider momentum conservation in the y-direction (taking 'upwards' as positive):
$0 = -m \times 0.80 \sin 20° + m \times V_y$
So $V_y = 0.80 \sin 20° = 0.27\ \text{m s}^{-1}$

c Magnitude of velocity of second ball, V, is given by Pythagoras' theorem:
$V^2 = (V_x)^2 + (V_y)^2$
so $V = \sqrt{[(0.25)^2 + (0.27)^2]} = 0.37\ \text{m s}^{-1}$

Direction of velocity of second ball, angle θ to the x-direction

$= \tan^{-1}\left(\frac{V_y}{V_x}\right) = \tan^{-1}\left(\frac{0.27}{0.25}\right) = 47°$

13 **a** Change in momentum

$\Delta p = mv - mu = m(v - u) = 1000 \times (24 - 10)$

$= 1.4 \times 10^4 \text{ kg m s}^{-1}$

b Force $F = \frac{\Delta p}{\Delta t} = \frac{1.4 \times 10^4}{15} = 933 \text{ N} \approx 930 \text{ N}$

14 **a** Rearrange $F = \frac{\Delta p}{\Delta t}$ to give change in momentum,

$\Delta p = F\Delta t = 240 \times 0.25 = 60 \text{ kg m s}^{-1}$ (or 60 N s)

b In the direction of the kicking force.

15 Force applied to water,

$F = \frac{\Delta p}{\Delta t} = \frac{m(v - u)}{\Delta t}$

$= \frac{m}{\Delta t}(v - u)$

$= 10 \times (0 - 5.0) = -50 \text{ N}$

(negative because force applied to water is against the direction of water flow)

So force of water on wall = 50 N

If the water bounces, a greater force is applied because of a greater change in momentum.

16 Force exerted by golf club,

$F = \frac{\Delta p}{\Delta t} = \frac{m(v - u)}{\Delta t} = \frac{0.046 \times (50 - 0)}{0.0013}$

$= 1.77 \times 10^3 \text{ N} \approx 1.8 \text{ kN}$

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Answers to SAQs

Chapter 7

1 Volume of cube = $3.0 \times 3.0 \times 3.0 = 27\,\text{cm}^3$

Density = $\frac{\text{mass}}{\text{volume}} = \frac{240}{27} = 8.89\,\text{g cm}^{-3} = 8890\,\text{kg m}^{-3}$

2 Volume of sphere

$V = \frac{4}{3}\pi r^3 = \frac{4}{3} \times \pi \times (0.15)^3 = 0.0141\,\text{m}^3$

Rearrange density = $\frac{\text{mass}}{\text{volume}}$ so

mass = density × volume

$= 7850 \times 0.0141 = 111\,\text{kg}$

3 Pressure, $p = \frac{F}{A} = \frac{80}{4 \times 0.0010} = 20\,\text{kPa}$

4 Estimate weight = 600 N,

area of feet = $500\,\text{cm}^2 = 0.05\,\text{m}^2$

So pressure $p = \frac{F}{A} = \frac{600}{0.05} = 12\,\text{kPa}$.

5 Pressure at depth 0.8 m is

$p = \rho g h_1 = 1000 \times 9.81 \times 0.8 = 7.85 \times 10^3\,\text{Pa}$

Pressure at depth 2.4 m is

$p = \rho g h_2 = 1000 \times 9.81 \times 2.4 = 2.35 \times 10^4\,\text{Pa}$

Maximum total pressure, $p = p_{\text{atm}} + p_{\text{water}}$

$= 1.01 \times 10^5 + 2.35 \times 10^4 = 1.25 \times 10^5\,\text{Pa}$

6 Rearrange $p = \rho g h$ to give height

$h = \frac{p}{\rho g} = \frac{1.01 \times 10^5}{1.29 \times 9.81} = 7980\,\text{m} \approx 8000\,\text{m}.$

This figure is too small because it assumes the density of the air is constant. In fact, density decreases with height.

7 a Spring D has the greatest value of force constant (the graph has the steepest gradient).

b Spring A is the least stiff (it extends the most for each unit of force applied).

c Spring C does not obey Hooke's law – there is no section of the graph that forms a straight line.

8 Metals from stiffest to least stiff:

	Metal	Young modulus / GPa
Most stiff	steel	210
	iron (wrought)	200
	copper	130
	brass	90–110
	aluminium	70
	tin	50
Least stiff	lead	18

9 Stiffest non-metal is glass (Young modulus = 70–80 GPa).

10 For material A, Young modulus

$E_A = \frac{\text{stress}}{\text{strain}} = \frac{\sigma}{\varepsilon} = \frac{15 \times 10^6}{0.001} = 1.5 \times 10^{10}\,\text{Pa} = 1.5\,\text{GPa}$

For material B, Young modulus,

$E_B = \frac{\text{stress}}{\text{strain}} = \frac{\sigma}{\varepsilon} = \frac{12 \times 10^6}{0.0024} = 5.0 \times 10^9\,\text{Pa} = 5.0\,\text{GPa}$

11 Stress = $\frac{\text{force}}{\text{cross-sectional area}}$

$= \frac{50}{0.5 \times 10^{-6}} = 1.0 \times 10^8\,\text{Pa}$

(remember that $0.5\,\text{mm}^2 = 0.5 \times 10^{-6}\,\text{m}$)

Strain = $\frac{\text{extension}}{\text{original length}} = \frac{0.1}{200.0}$

$= 5.0 \times 10^{-4}$ (0.05%)

Young modulus = $\frac{\text{stress}}{\text{strain}} = \frac{1.0 \times 10^8}{5.0 \times 10^{-4}}$

$= 2.0 \times 10^{11}\,\text{Pa}$

12 Young modulus, $E = \frac{\text{stress}}{\text{strain}}$

Rearrange so strain = $\frac{\text{stress}}{\text{Young modulus}}$

Then insert formulae for stress and strain

$\frac{x}{L} = \frac{F}{A \times E}$

and cross-sectional area, $A = \frac{\pi d^2}{4}$

This gives extension = $\frac{4FL}{\pi d^2 \times E}$

$= \frac{4 \times 10 \times 1.00}{\pi \times (0.001)^2 \times 130 \times 10^9}$

$= 9.796 \times 10^{-5}\,\text{m}$

$\approx 9.8 \times 10^{-5}\,\text{m}$

13 Stress = $\frac{\text{force}}{\text{cross-sectional area}} = \frac{4F}{\pi d^2}$

$= \frac{4 \times 1.00}{\pi \times (0.0004)^2} = 8.0 \times 10^6\,\text{Pa}$

Strain = $\frac{\text{extension}}{\text{original length}} = \frac{0.001}{0.800}$

$= 1.25 \times 10^{-3}$ (at most)

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Young modulus = $\frac{\text{stress}}{\text{strain}} = \frac{8.0 \times 10^6}{1.25 \times 10^{-3}}$
$= 6.4 \times 10^9$ Pa (but could be more, because extension may be less than 1 mm).

14 **a** Young modulus = $\frac{\text{stress}}{\text{strain}} = \frac{150 \times 10^6}{0.003} = 50$ GPa, breaking stress = 150 MPa

b Young modulus = $\frac{\text{stress}}{\text{strain}} = \frac{100 \times 10^6}{0.001} = 100$ GPa, breaking stress = 130 MPa
Note that the Young modulus is only found for the straight portion of the stress–strain graph.

c Young modulus = $\frac{\text{stress}}{\text{strain}} = \frac{100 \times 10^6}{0.004} = 25$ GPa, breaking stress = 100 MPa

15 Elastic potential energy,
$E = \frac{1}{2}Fx = \frac{1}{2} \times 12 \times 0.18 = 1.08\text{ J} \approx 1.1\text{ J}$
The rubber band is assumed to obey Hooke's law; hence the answer is an estimate.

16 Elastic potential energy, $E = \frac{1}{2}Fx = \frac{1}{2}kx^2$
$= \frac{1}{2} \times 4800 \times (0.0020)^2 = 9.6 \times 10^{-3}$ J

17 **a** A has greater stiffness (less extension per unit force).

b A requires greater force to break (line continues to higher force value).

c B requires greater amount of work done to break (larger area under graph).

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Answers to SAQs

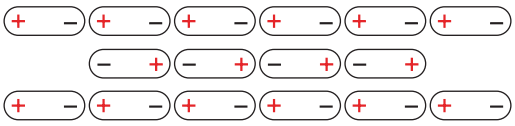
Chapter 8

1 **a** Diagram **i** shows positive charges repelling.

b Diagram **iii** shows negative charges repelling.

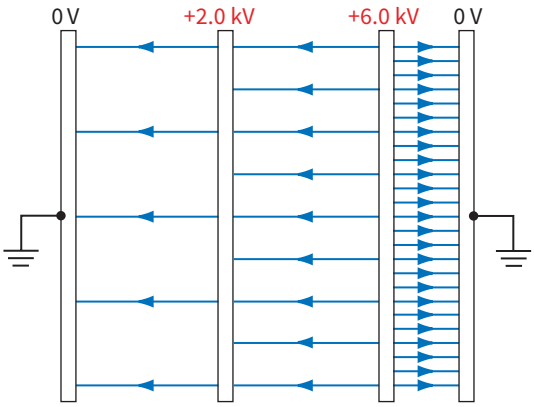
c Diagram **ii** shows opposite charges attracting.

2



3 The field strength is greatest at the top (pointed) part of the building. The electric field lines are closest together there.

4



5 Electric field strength

$E = \frac{F}{Q} = \frac{150}{20 \times 10^{-3}} = 7500\ \text{N C}^{-1}$ downwards

6 Electric field strength

$E = \frac{V}{d} = \frac{1000}{0.40} = 2500\ \text{V m}^{-1} = 2500\ \text{N C}^{-1}$

7 Electric field strength

$E = \frac{F}{Q} = \frac{8 \times 10^{-16}}{1.6 \times 10^{-19}} = 5000\ \text{V m}^{-1}$ **or** N C^{-1}

8 **a** Rearrange $E = \frac{V}{d}$ so that voltage to cause sparks is given by

$V = Ed = 40\,000 \times 4 = 160\,000\ \text{V}$

b Rearrange $E = \frac{V}{d}$ so that minimum distance is given by

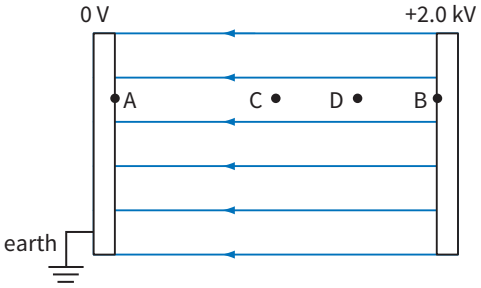
$d = \frac{V}{E} = \frac{325\ \text{V}}{40\,000\ \text{V cm}^{-1}} = 0.008\ \text{cm} = 0.08\ \text{mm}$

c Rearrange $E = \frac{V}{d}$ so that voltage to cause lightning is given by

$V = Ed = 40\,000 \times 10\,000 = 400\ \text{MV}$

(Remember that the field strength was given in volts per centimetre.)

9 **a**



b Potential difference = 2.0 kV

c Electric field strength between parallel plates is uniform, so has the same value at both points.

$E = \frac{V}{d} = \frac{2.0 \times 10^{3}}{0.25} = 8.0\ \text{kV m}^{-1}$

d $F = QE = +5 \times 10^{-6} \times 8.0 \times 10^{3} = 0.04\ \text{N}$ to the left

10 $E = \frac{V}{d} = \frac{5.0 \times 10^{3}}{0.10} = 5.0 \times 10^{4}\ \text{V m}^{-1}$ **or** N C^{-1}

$F = QE = +2 \times 10^{-6} \times 5.0 \times 10^{4} = 0.10\ \text{N}$

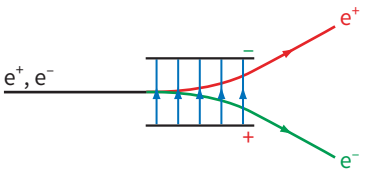
11 Force $F = QE = -1.6 \times 10^{-19} \times 5.0 \times 10^{6}$

$= -8.0 \times 10^{-13}\ \text{N}$

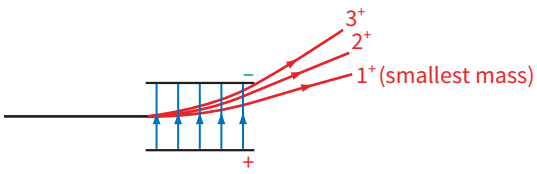
Rearrange $F = ma$ to give magnitude of acceleration:

$a = \frac{F}{m} = \frac{8.0 \times 10^{-13}}{9.11 \times 10^{-31}} = 8.8 \times 10^{17}\ \text{m s}^{-2}$

12 **a**



b



Ions with a greater mass will show smaller deflection.

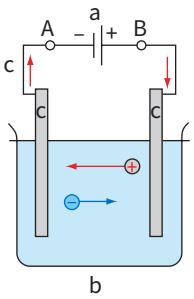
Ions with greater charge will have greater deflection.

Answers to SAQs

Chapter 9

1 Towards the right.

2 **a, b, c**



3 Charge $\Delta Q = I\Delta t = 0.4 \times 15 = 6.0\,\text{C}$

4 Rearrange $\Delta Q = I\Delta t$ so $I = \frac{\Delta Q}{\Delta t} = \frac{150}{30} = 5\,\text{A}$

5 Rearrange $\Delta Q = I\Delta t$ so $I = \frac{\Delta Q}{\Delta t} = \frac{50}{20} = 2.5\,\text{A}$

6 **a** Rearrange $\Delta Q = I\Delta t$ and use 50 A per hour as the charge available, so
$\Delta t = \frac{\Delta Q}{I} = \frac{50}{200} = 0.25$ hours = 15 minutes

b $\Delta Q = I\Delta t$; for a current of 200 A,
$\Delta Q = 200 \times 15 \times 60 = 180\,000\,\text{C}$

7 Number of protons in 1 C
$= \frac{1}{1.6 \times 10^{-19}} = 6.2 \times 10^{18}$ protons

8 The smallest possible unit of electric charge is the elementary charge, 1.6×10^{-19} C. So the only possible electrical charges are whole multiples of this elementary charge. The only value from the list that is a whole multiple is 8.0×10^{-19} C.

9 Current $I = nAve$
$= 5.9 \times 10^{28} \times 2.0 \times 10^{-6} \times 0.1 \times 10^{-3} \times 1.6 \times 10^{-19}$
$= 1.9\,\text{A}$

10 Cross-sectional area of copper wire
$A = \pi r^2 = \frac{\pi d^2}{4} = \frac{\pi \times (1.0 \times 10^{-3})^2}{4} = 0.79 \times 10^{-6}\,\text{m}^2$
Rearrange $I = nAve$ to give drift velocity
$v = \frac{I}{nAe} = \frac{5.0}{8.5 \times 10^{28} \times 0.79 \times 10^{-6} \times 1.6 \times 10^{-19}}$
$= 0.47 \times 10^{-3}\,\text{m s}^{-1} = 0.47\,\text{mm s}^{-1}$

11 Current $I = nAve$, but remember that, in a series circuit, current is the same all the way round the circuit. Also, A and e are the same for both lengths of wire. This means that $n_1Av_1e = n_2Av_2e$, where the 1 denotes the copper wire and the 2 denotes the silver wire.
So that means $\frac{V_2}{V_1} = \frac{n_1}{n_2} = \frac{8.5 \times 10^{28}}{5.9 \times 10^{28}} = 1.4$.
This means that the drift velocity in the silver wire is 1.4 times the drift velocity in the copper wire.

12 Current $I = \frac{V}{R} = \frac{12}{36} = 0.33\,\text{A}$

13 A 60 W lamp has higher resistance. It allows less current through for the same potential difference (power is current × potential difference, so, applying Ohm's law, power is inversely proportional to resistance).

14 **a** Potential difference, $V = IR = 1.0 \times 50 = 50\,\text{V}$

b Potential difference, $V = IR = 2.0 \times 50 = 100\,\text{V}$

15 Resistance $R = \frac{V}{I} = \frac{230}{0.40} = 580\,\Omega$ (to 2 sig. figs)

16 Resistance $R = \frac{V}{I} = \frac{6.0}{2.4} = 2.5\,\Omega$

17 Power $P = VI$, so rearrange to give current:
$I = \frac{P}{V} = \frac{60}{230} = 0.26\,\text{A}$

18 Power $P = VI = 25 \times 10^3 \times 40 \times 10^3 = 1.0 \times 10^9\,\text{W}$ (1000 MW or 1 GW)

19 **a** Power $P = VI$ so rearrange to give current:
$I = \frac{P}{V} = \frac{10^4}{230} = 43\,\text{A}$

b Fuse needs to be rated above, but close to, full normal current draw, so a fuse of 45 to 50 A would be appropriate.

20 Power $P = \frac{V^2}{R} = \frac{(3.0)^2}{(20 \times 10^3)} = 0.45\,\text{mW}$

21 **a** Current $I = \frac{P}{V} = \frac{15}{230} = 0.065\,\text{A}$

b Resistance $R = \frac{V}{I} = \frac{230}{0.065} = 3500\,\Omega$

22 Rearrange power $P = I^2R$ to give resistance:
$R = \frac{P}{I^2} = \frac{100}{(0.43)^2} = 540\,\Omega$

23 Energy transferred
$W = IV\Delta t = 10 \times 12 \times (5.0 \times 60 \times 60)$
$= 2200\,000$ J, or 2.2 MJ

24 **a** Charge through lamp
$\Delta Q = I\Delta t = 10 \times 20 = 200$ C

b Per coulomb of charge,
energy $= \frac{W}{\Delta Q} = \frac{400}{200} = 2.0\ \text{J C}^{-1}$

c Rearrange energy transferred, $W = IV\Delta t$,
to give potential difference: $V = \frac{\Delta W}{I\Delta t}$
$= \frac{400}{(10 \times 20)} = 2.0$ V
or use potential difference $= \frac{\text{energy}}{\text{charge}}$
$= \frac{400}{200} = 2.0$ V

Answers to SAQs

Chapter 10

1 Kirchhoff's first law means that
$I + 3.0\,A = 7.5\,A$, so $I = 7.5 - 3.0 = 4.5\,A$

2 Current flowing into P equals current flowing out of P, so if current in wire X is I_x, then
$I_x + 3.0\,A + 2.5\,A = 7.0\,A$
So $I_x = 7.0 - 3.0 - 2.5 = 1.5\,A$, towards P

3 Current flowing into junction
$\Sigma I_{in} = 3.0 + 2.5 + 1.0 = 6.5\,A$
Current flowing out of junction
$\Sigma I_{out} = 4.0 + 0.5 + 2.0 = 6.5\,A$
Yes, Kirchhoff's law is satisfied.

4 $I_x = 3.0 + 2.0 - 7.0 = -2.0\,A$, which means the current is towards P.

5 Sum of e.m.f.s around any loop in a circuit is equal to the sum of the p.d.s around the loop.
So e.m.f. of power supply = Σp.d.s across resistors, meaning p.d. across resistor R
= e.m.f. of power supply − p.d. across 20 Ω resistor
$= 10 - (0.1 \times 20) = 8.0\,V$
$V = IR$ so resistance $R = \frac{V}{I} = \frac{8.0}{0.1} = 80\,\Omega$

6 **a** Choose the loop containing the 5 V cell at the top, the 10 Ω resistor with current I, and the central 5 V cell, as the only current involved is I.

b Sum of e.m.f.s of cells in loop = 5.0 + 5.0 = 10 V = p.d. across resistor.
$V = IR$ so $I = \frac{V}{R} = \frac{10}{10} = 1.0\,A$

7 In loop, sum of p.d.f.s = 30 − 10 = 20 V, which by Kirchhoff's second law must equal the sum of the p.d.s across the resistors, given by $V = IR$.
Sum of p.d.s across resistors
$= (0.5 \times R) + (0.5 \times 10) + (0.2 \times 10) + (0.2 \times 20)$
So $20 = (0.5 \times R) + 11$, giving
$R = \frac{(20 - 11)}{0.5} = 18\,\Omega$

8 In series, the 1 C charge passes through both batteries and gains or loses 6 J in each. If the batteries are connected so that both of them move the charge in the same direction, the total e.m.f. = 6 + 6 = 12 V.
If the batteries are connected back to front, the charge gains energy in one cell but loses it in the other, so total e.m.f. = 0 V.
In parallel, half the charge flows through one battery and half through the other, so the total energy gained is 6 J, meaning the total e.m.f. = 6 V.

9 Consider the circuit loop at the top, containing the 10 V cell and a 20 Ω resistor.
Use Kirchhoff's second law and $V = IR$ to give
$10\,V = I_1 \times 20\,\Omega$, so current through A_1 is
$I_1 = \frac{10}{20} = 0.50\,A$
Consider the circuit loop at the bottom, containing the 5 V cell and a 20 Ω resistor.
Use Kirchhoff's second law and $V = IR$ to give
$5\,V = I_1 \times 20\,\Omega$, so current through A_3 is
$I_3 = \frac{5}{20} = 0.25\,A$
Now use Kirchhoff's first law at the circuit junction to the right of ammeter A_2 to give
$I_1 = I_2 + I_3$
So current through A_2 is
$I_2 = I_1 - I_3 = 0.50 - 0.25\,A$

10 Total resistance = 5 + 5 + 10 = 20 Ω

11 Use Kirchhoff's second law to give
e.m.f. = $V_1 + V_2$
So V_2 = e.m.f. − V_1 = 2.0 − 1.2 = 0.8 V

12 **a** All five in series and pointing the same way, so
e.m.f. = 1.5 + 1.5 + 1.5 + 1.5 + 1.5 = 7.5 V

b Five in series, with two facing in the opposite direction, so
e.m.f. = 1.5 + 1.5 + 1.5 − 1.5 − 1.5 = 1.5 V
or all five in parallel

c Five in series, with one facing in the opposite direction, so
e.m.f. = 1.5 + 1.5 + 1.5 + 1.5 − 1.5 = 4.5 V
or two in parallel to give e.m.f. of 1.5 V, connected in series to two more in parallel (also giving e.m.f. of 1.5 V), then connected in series to the single remaining cell with e.m.f. of 1.5 V.
Sum of e.m.f.s = 1.5 + 1.5 + 1.5 = 4.5 V

Cambridge International AS and A Level Physics © Cambridge University Press 2014

13 $\frac{1}{R_{Total}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \frac{1}{R_4} = \frac{4}{10}$
So $R_{Total} = \frac{10}{4} = 2.5\,\Omega$

14 **a** $R = R_1 + R_2 = 100 + 200 = 300\,\Omega$
b $\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} = \frac{1}{100} + \frac{1}{200} = \frac{3}{200}$, so $R = \frac{200}{3} = 67\,\Omega$
c Resistance of the series combination is given in part **a**, 300 Ω
So for full combination,
$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} = \frac{1}{300} + \frac{1}{200} = \frac{5}{600}$
So $R = \frac{600}{5} = 120\,\Omega$

15 **a** Rearrange $V = IR$ to give $I = \frac{V}{R} = \frac{12}{500} = 0.024\,A$
b Total resistance
$R = R_1 + R_2 = 500 + 1000 = 1500\,\Omega$
Rearrange $V = IR$ to give $I = \frac{V}{R} = \frac{12}{1500} = 0.008\,A$
c Total resistance R is given by
$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} = \frac{1}{500} + \frac{1}{1000} = \frac{3}{1000}$
So $R = \frac{1000}{3} = 333\,\Omega$
Rearrange $V = IR$ to give $I = \frac{V}{R} = \frac{12 \times 3}{1000} = 0.036\,A$

16 Total resistances possible are: 40 Ω, 50 Ω, 67 Ω, 75 Ω, 100 Ω (two ways), 167 Ω, 200 Ω (two ways), 250 Ω, 300 Ω and 400 Ω. In detail:
i One 100 Ω on its own
ii One 200 Ω on its own
iii Both 100 Ω in series, $R = 200\,\Omega$
iv 100 Ω + 200 Ω in series, $R = 300\,\Omega$
v All in series, $R = 200 + 100 + 100 = 400\,\Omega$
vi Both 100 Ω in parallel,
$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} = \frac{1}{100} + \frac{1}{100} = \frac{2}{100}$ so $R = 50\,\Omega$
vii 100 Ω and 200 Ω in parallel,
$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} = \frac{1}{100} + \frac{1}{200} = \frac{3}{200}$ so $R = 67\,\Omega$
viii 100 Ω and 200 Ω in parallel, plus 100 Ω in series, $R = 67 + 100 = 167\,\Omega$
ix 100 Ω and 100 Ω in parallel, plus 200 Ω in series, $R = 50 + 200 = 250\,\Omega$
x 100 Ω and 100 Ω in series, connected in parallel with 200 Ω, so
$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} = \frac{1}{200} + \frac{1}{200} = \frac{2}{200}$ so $R = 100\,\Omega$
xi 100 Ω and 200 Ω in series, connected in parallel with 100 Ω, so
$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} = \frac{1}{300} + \frac{1}{100} = \frac{4}{300}$ so $R = 75\,\Omega$
xii All in parallel, so $\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$
$= \frac{1}{100} + \frac{1}{100} + \frac{1}{200} = \frac{5}{200}$ so $R = 40\,\Omega$

17 10 Ω (remember, for resistors connected in parallel, their combined resistance is smaller than any of the individual resistances).

18 The p.d. across each resistor is the same as the e.m.f. of the battery. Rearrange $V = IR$ to give current, $I = \frac{V}{R} = \frac{10}{20} = 0.50\,A$

19 Combined resistance of all resistors, R, is given by $\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} = \frac{1}{20} + \frac{1}{40} + \frac{1}{50} = \frac{19}{200}$ so
$R = \frac{200}{19} = 10.5\,\Omega$
Rearrange $V = IR$ to give current
$I = \frac{V}{R} = \frac{10}{10.5} = 0.95\,A$
But an easier way to approach this is to calculate the current through each resistor using $I = \frac{V}{R}$, given that the p.d. across each resistor is the same and equals the e.m.f. of the battery.
Sum of currents $= \frac{10}{20} + \frac{10}{40} + \frac{10}{50} = \frac{190}{200} = 0.95\,A$

20 $\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2}$ so $\frac{1}{10} = \frac{1}{20} + \frac{1}{R_2}$
So $\frac{1}{R_2} = \frac{1}{10} - \frac{1}{20} = \frac{1}{20}$, hence second resistor $R_2 = 20\,\Omega$

21 Two in parallel, connected in series with a further two.
For the parallel combination,
$\frac{1}{R_P} = \frac{1}{100} + \frac{1}{100} = \frac{1}{50}$ so $R_P = 50\,\Omega$
Thus the total resistance of the series combination is $R_S = 100 + 100 + 50 = 250\,\Omega$

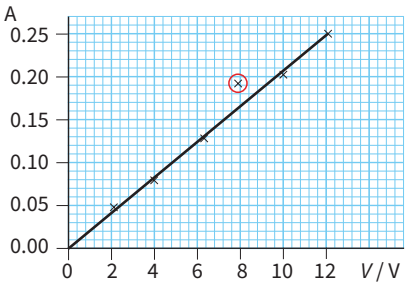
22 Resistance of parallel combination given by
$\frac{1}{R_P} = \frac{1}{300} + \frac{1}{60} = \frac{1}{50}$ so $R_P = 50\,\Omega$
So total resistance of circuit
$R = 50 + 50 = 100\,\Omega$
Rearrange $V = IR$ to give current, $I = \frac{V}{R}$
Current at A, $I = \frac{600}{100} = 6.0\,A$
Current at B is the same as at A, 6.0 A
The p.d. across parallel combination
= e.m.f. – p.d. across 50 Ω resistor
$= 600 - (6.0 \times 50) = 300\,V$
Current at C is $\frac{V}{R} = \frac{300}{300} = 1.0\,A$
Current at D is $\frac{300}{60} = 5.0\,A$
Current at E = current at A = 6.0 A

23 **a** Current $= \frac{V}{R} = \frac{10}{100} = 0.10\,A$
b Current $= \frac{V}{R} = \frac{10}{100 + 5.0} = 0.095\,A$

Answers to SAQs

Chapter 11

1 a



b All except point at 7.9; this point is anomalous and can be ignored.

c 48 Ω

d Yes

2 At 2.0 V, $R = \frac{V}{I} = \frac{2.0}{0.010} = 200\,\Omega$;
at 8.0 V, $R = \frac{8.0}{0.060} = 133\,\Omega$.
No, it does not obey Ohm's law.

3 a At θ_1, $R = \frac{V}{I} = \frac{30}{2.4} = 12.5\,\Omega$; at θ_2, $R = \frac{15}{1.5} = 10\,\Omega$.

b θ_1 is the higher temperature.

4 a Filament lamp is A, steel wire is B.

b 8.0 V

c Resistance, $R = \frac{V}{I} = \frac{8.0}{3.4} = 2.4\,\Omega$

5 a i 3.1 kΩ

ii 1.5 kΩ

b i 5 °C

ii 36 °C

6 The lamp will become brighter because the resistance of the thermistor decreases. This reduces the total resistance in the circuit and therefore the current increases.

7 a A thermistor's resistance changes more per degree of temperature change than a metal wire. This makes the thermometer more sensitive.

b A metal wire will work over a much wider range than a thermistor. **Or**, the metal's resistance increases almost linearly, making the thermometer more linear.

8 a Rearrange resistance, $R = \frac{\rho L}{A}$ to give length $L = \frac{RA}{\rho}$, where cross-sectional area
$A = \frac{1}{4}\pi d^2 = \frac{1}{4} \times \pi \times (0.5 \times 10^{-3})^2 = 1.96 \times 10^{-7}\,\text{m}^2$
So length for 1.0 Ω resistance
$= \frac{RA}{\rho} = \frac{1.0 \times 0.2 \times 10^{-6}}{44.0 \times 10^{-8}} = 0.45\,\text{m}$

b Length for 5.0 Ω resistance
$= \frac{RA}{\rho} = \frac{5.0 \times 0.2 \times 10^{-6}}{44.0 \times 10^{-8}} = 2.2\,\text{m}$

c Length for 10 Ω resistance
$= \frac{RA}{\rho} = \frac{10 \times 0.2 \times 10^{-6}}{44.0 \times 10^{-8}} = 4.5\,\text{m}$

9 Volume of copper, $V = 1.0\,\text{cm}^3$
= length L × cross-sectional area A
So length $L = \frac{V}{A}$
Thus resistance
$R = \frac{\rho L}{A^2} = \frac{1.69 \times 10^{-8} \times 1.0 \times 10^{-6}}{(4.0 \times 10^{-7})^2} = 0.11\,\Omega$

10 a Resistance, $R = \frac{\rho L}{A}$. We know the first length, $L_1 = 1.0$ m, as well as the resistivity of copper and the resistance R_1. So the cross-sectional area of the wire is $A = \frac{\rho L_1}{R_1}$.
The resistance of the 5.0 m length of wire is
$R_2 = \frac{\rho L_2}{A} = \frac{L_2 R_1}{L_1} = \frac{5.0 \times 0.50}{1.0} = 2.5\,\Omega$

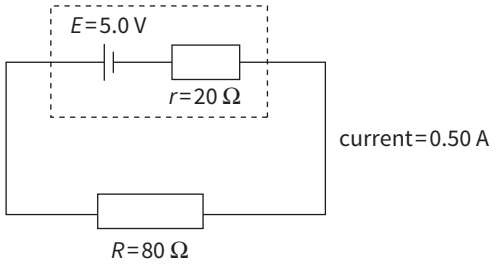
b Area $= \frac{1}{4}\pi d^2$, so halving the diameter reduces the area by a factor of 4. Resistance is inversely proportional to area. Therefore halving the diameter increases the resistance by a factor of 4. So resistance $R = 0.50 \times 4 = 2.0\,\Omega$.

11 40 Ω. The resistance increases by a factor of 4 (because cross-sectional area has halved and length has doubled).

Answers to SAQs

Chapter 12

1



Rearrange e.m.f. $E = I(R+r)$ to give current

$I = \frac{E}{R+r} = \frac{5.0}{8.0+2.0} = 0.50\,\text{A}$

2 **a** For circuit (i), rearrange e.m.f. $E = I(R+r)$ to give current

$I = \frac{E}{R+r} = \frac{3.0}{10+10+4.0} = 0.125 \approx 0.13\,\text{A}$

For circuit (ii), external resistance R is given by $\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} = \frac{1}{10} + \frac{1}{10}$, so $R = 5.0\,\Omega$.

So current $I = \frac{E}{R+r} = \frac{3.0}{5.0+4.0} = 0.33\,\text{A}$.

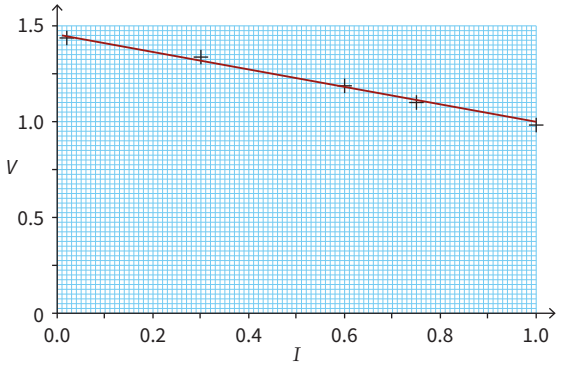
b For circuit (i),
lost volts = $Ir = 0.125 \times 4.0 = 0.5\,\text{V}$
terminal p.d. = e.m.f. – lost volts
= 3.0 – 0.5 = 2.5 V
For circuit (ii),
lost volts = $Ir = 0.33 \times 4.0 = 1.33\,\text{V}$
terminal p.d. = e.m.f. – lost volts
= 3.0 – 1.33 = 1.67 V

3 Rearrange e.m.f., $E = I(R+r)$ to give current,

$I = \frac{E}{R+r} = \frac{1.5 \times 4}{2.0 + (0.1 \times 4)} = 2.5\,\text{A}$

4 $E = 3.0\,\text{V}$.
P.d. across 10 Ω resistor = 2.8 V so current in circuit with resistor connected
$= \frac{V}{R} = \frac{2.8}{10} = 0.28\,\text{A}$
Rearrange e.m.f. $E = IR + Ir$ to give internal resistance of battery:
$r = \frac{E - IR}{I} = \frac{3.0 - 2.8}{0.28} = 0.71\,\Omega$

5



$E = 1.5\,\text{V}$, $r = 0.5\,\Omega$

6 **a** Terminal p.d. = $E - Ir = 12 - (100 \times 0.04) = 8\,\text{V}$

b Power, $P = VI$, and p.d. $V = IR$ means current, $I = \frac{V}{R}$, so $P = \frac{V^2}{R}$
Rearrange to give resistance,
$R = \frac{V^2}{R} = \frac{(12)^2}{36} = 4.0\,\Omega$

c Power, $P = \frac{V^2}{R} = \frac{(8)^2}{4} = 16\,\text{W}$

7 When resistor is set to 0 Ω, $V_{out} = 0\,\text{V}$
When resistor is set to 40 Ω,
$V_{out} = \frac{R_2}{(R_1+R_2)} \times V_{in} = \frac{40}{(10+40)} \times 10 = 8\,\text{V}$

8 **a** For a 1 cm length of wire,
potential difference = $\frac{4.0}{100} = 0.04\,\text{V}$
Length needed for 1.0 V = $\frac{1.0}{0.04} = 25\,\text{cm}$

b A length of 37.0 cm has a p.d. across it of $37.0 \times 0.04 = 1.48\,\text{V}$. The driver cell will have internal resistance and it is supplying current to the potentiometer wire. Therefore the p.d. across its terminals and the wire will be slightly less than the e.m.f. (4.0 V) of the cell.

c If a balance length of 31.2 cm is required by a cell of e.m.f. 1.230 V, then p.d. supplied by unknown e.m.f. cell = $1.230 \times \frac{37.0}{31.2}$
= 1.459 V ≈ 1.46 V

Answers to SAQs

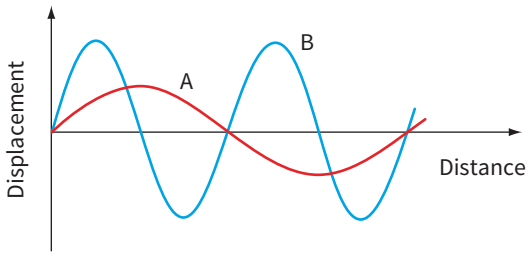
Chapter 13

1 **a** Wavelength 15 cm, amplitude 4.0 cm

b Wavelength 20 cm, amplitude 2.0 cm

2 One complete wave occupies 2.5 scale divisions. One scale division represents 0.005 s, so the period of the wave $T = 2.5 \times 0.005 = 0.0125$ s. So the frequency $f = \frac{1}{T} = \frac{1}{0.0125} = 80$ Hz

3



4 **a** Intensity $= \frac{\text{power}}{\text{cross-sectional area}} = \frac{P}{A}$

Surface area of a sphere $= 4\pi r^2$

So the intensity at 1.0 m from the lamp

$= \frac{P}{4\pi r^2} = \frac{100}{4 \times \pi \times (1.0)^2}$

$= 7.96\,\text{W m}^{-2} \approx 8.0\,\text{W m}^{-2}$

b Intensity at 2.0 m from the lamp

$= \frac{P}{4\pi r^2} = \frac{100}{4 \times \pi \times (2.0)^2}$

$= 1.99\,\text{W m}^{-2} \approx 2.0\,\text{W m}^{-2}$

5 **a** Intensity is proportional to (amplitude)^2, so doubling amplitude multiplies intensity by 4, to 1600 W m^{-2}.

b Intensity is reduced by a factor of 4, so the amplitude decreases by a factor of 2, to 2.5 cm.

6 Rearrange speed of wave, $v = f\lambda$, to give frequency $f = \frac{v}{\lambda} = \frac{5060}{0.25} = 20\,240$ Hz ≈ 20 kHz

7 Speed $v = f\lambda = 64 \times 1.40 = 89.6\,\text{m s}^{-1} \approx 90\,\text{m s}^{-1}$

8 **a** Wavelength $\lambda = 20/4 = 5.0$ cm

b Frequency $f = 30$ Hz

c Speed $v = f\lambda = 30 \times 0.050 = 1.5\,\text{m s}^{-1}$

9

Station	Wavelength / m	Frequency / MHz
Radio A (FM)	$= \frac{v}{f} = \frac{3.0 \times 10^8}{97.6 \times 10^6}$ $= 3.07$	97.6
Radio B (FM)	$= \frac{v}{f} = \frac{3.0 \times 10^8}{94.6 \times 10^6}$ $= 3.17$	94.6
Radio B (LW)	1515	$= \frac{v}{\lambda} = \frac{3.0 \times 10^8}{1515\,\text{Hz}} = 0.198$ (note: MHz)
Radio C (MW)	693	$= \frac{v}{\lambda} = \frac{3.0 \times 10^8}{693\,\text{Hz}} = 0.433$

10 **a** Observed wavelength

$\lambda_{obs} = \frac{(v + v_s)}{f} = \frac{(330 - 80)}{120} = 2.1$ m

$f_{obs} = \frac{c}{\lambda_{obs}} = \frac{330}{2.1} = 160$ Hz

b $f_{obs} = \frac{f_s \times v}{(v + v_s)} = \frac{120 \times 330}{(330 + 80)} = 97$ Hz

11 **a** In vacuum, frequency

$f = \frac{v}{\lambda} = \frac{3.0 \times 10^8}{700 \times 10^{-9}} = 4.3 \times 10^{14}$ Hz

b In glass, frequency is the same,

$f = 4.3 \times 10^{14}$ Hz

Wavelength $\lambda = \frac{v}{f}$

$= \frac{2.0 \times 10^8}{4.3 \times 10^{14}} = 465 \times 10^{-9} \approx 470$ nm

12 **a** Change in wavelength, $\Delta\lambda = 550 - 535 = 15$ nm

b $f_0 = \frac{3.0 \times 10^8}{550 \times 10^{-9}} = 5.45455 \times 10^{14}$ Hz

$f_s = \frac{3.0 \times 10^8}{535 \times 10^{-9}} = 5.60748 \times 10^{14}$ Hz

More sig. figs are included for accuracy because the difference in these values is needed.

$f_0 = \frac{f_s \times v}{(v + v_s)}$ gives

$5.45455 = \frac{5.60748 \times 3 \times 10^8}{(3 \times 10^8 + v)}$

giving $v = 8.4 \times 10^6\,\text{m s}^{-1}$

c Observed wavelength is larger than that observed in the laboratory, **or** frequency is lower than that observed in laboratory, so the star is moving away from the astronomer.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

13 Calculate frequencies using $f = \frac{v}{\lambda}$, with $v = 3.0 \times 10^8\ \text{m s}^{-1}$

Radiation	Wavelength range / m	Frequency / Hz
radio waves	$>10^6$ to 10^{-1}	300 to 3×10^9
microwaves	10^{-1} to 10^{-3}	3×10^9 to 3×10^{11}
infrared	10^{-3} to 7×10^{-7}	3×10^{11} to 4.3×10^{14}
visible	7×10^{-7} (red) to 4×10^{-7} (violet)	4.3×10^{14} to 7.5×10^{14}
ultraviolet	4×10^{-7} to 10^{-8}	7.5×10^{14} to 3×10^{16}
X-rays	10^{-8} to 10^{-13}	3×10^{16} to 3×10^{21}
γ-rays	10^{-10} to 10^{-16}	3×10^{18} to 3×10^{24}

14 **a** Visible
b Ultraviolet
c 1–100 mm
d 400–700 nm
e 4.3×10^{14} Hz to 7.5×10^{14} Hz

15 **a** radio waves
b microwaves
c infrared
d visible light
e ultraviolet
f X-rays or γ-rays

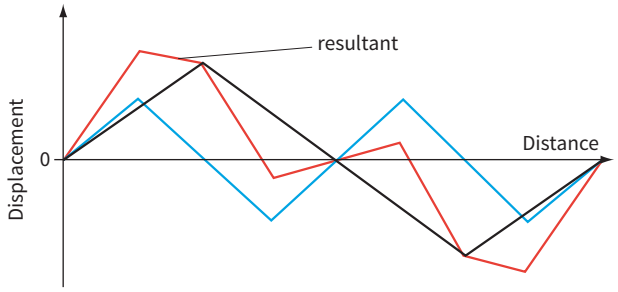
16 Use table of answers to question 13:
a radio waves
b radio waves
c visible light
d X-rays or γ-rays

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Answers to SAQs

Chapter 14

1

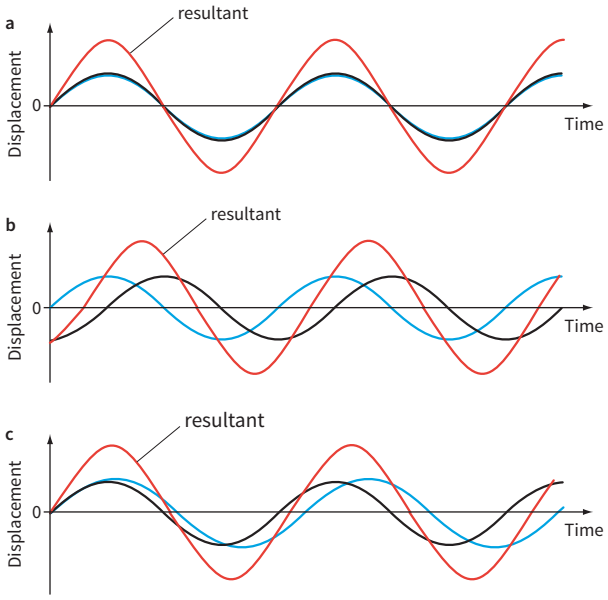


2 The grid spacing is much smaller than the wavelength of the microwaves, so the waves do not pass through. However, the wavelength of light is much smaller, so it can pass through unaffected.

3 Two loudspeakers with slightly different frequencies might start off in step, but they would soon go out of step. The interference at a particular point might be constructive at first, but would become destructive.

4 The intensity would increase.

5



6 D: dark fringe, because rays from slits 1 and 2 differ in path length by one-and-a-half wavelengths ($1\frac{1}{2}\lambda$).
E: bright fringe, because the path difference is 2λ.

7 Wavelength of light used, λ, and separation of slits, a, remain the same.
This means that $\frac{x_1}{D_1} = \frac{x_2}{D_2}$. Doubling D means x must also double, so separation of fringes $x_2 = 3.0\,\text{mm}$.

8 **a** $x = \frac{\lambda D}{a}$. Therefore $x \propto \frac{1}{a}$, so decreasing a gives increased x.

b Blue light has shorter wavelength, so x is less (x is proportional to λ).

c For larger D, x is greater, so there is a smaller percentage uncertainty in x (x is proportional to D).

9 Rearrange $\lambda = \frac{ax}{D}$ to give

$$x = \frac{\lambda D}{a} = \frac{589 \times 10^{-9} \times 1.20}{0.0002} = 3.5\,\text{mm}$$

10 D and a are fixed. So $\frac{\lambda_1}{x_1} = \frac{\lambda_2}{x_2}$ and so

$$x_2 = \frac{4.5 \times 10^{-7} \times 2.4 \times 10^{-3}}{6.0 \times 10^{-7}} = 1.8 \times 10^{-3}\,\text{m} = 1.8\,\text{mm}$$

(**or** wavelength is $\frac{3}{4}$ of previous value, so spacing of fringes is $\frac{3}{4}$ of previous value)

11 For the second-order maximum, rays from adjacent slits have a path difference of 2λ, so they are in phase.

12 **a** Rearrange $d \sin\theta = n\lambda$ so

$$\sin\theta = \frac{n\lambda}{d} = \frac{2 \times 580 \times 10^{-9}}{3.33 \times 10^{-6}} = 0.348, \text{ so } \theta = 20.4°$$

b For $n = 3$, $\sin\theta = 0.522$, so $\theta = 31.5°$
For $n = 4$, $\sin\theta = 0.697$, so $\theta = 44.2°$
For $n = 5$, $\sin\theta = 0.871$, so $\theta = 60.6°$
You cannot have $\sin\theta > 1$. So there are 11 maxima.

13 **a** θ increases, so the maxima are more spread out and there may be fewer of them. (Note: $\sin\theta \propto \lambda$.)

b d decreases, so again θ increases, the maxima are more spread out and there may be fewer of them. (Note: $\sin\theta \propto 1/d$.)

Cambridge International AS and A Level Physics © Cambridge University Press 2014

14 **a** By calculation, use $\lambda = \frac{ax}{D}$ so fringe separation

$x = \frac{\lambda D}{a} = \frac{546 \times 10^{-9} \times 0.80}{0.50 \times 10^{-3}} = 0.87\,\text{mm}$

so ten fringes have a total width of 8.7 mm.
But using a ruler, the student will measure 9 mm.

b Separation of lines in grating

$= \frac{1}{3000}\,\text{cm}$

$= 3.33 \times 10^{-6}\,\text{m}$

By calculation, rearrange $d \sin\theta = n\lambda$ so

$\sin\theta = \frac{n\lambda}{d} = \frac{2 \times 546 \times 10^{-9}}{3.33 \times 10^{-6}} = 0.328$

so $\theta = 19.1°$, but the student will measure 19°.

c For the double-slit experiment, a measured width of 9 mm for ten fringes will give a wavelength of 562 nm. For the diffraction grating experiment, the measured second-order angle of 19° will give 543 nm. Hence the diffraction grating method is more accurate. In practice, it is also much more precise because the fringes are bright and sharp (well-defined).

15 **a** For red light, rearrange $d \sin\theta = n\lambda$ so

$\sin\theta = \frac{n\lambda}{d} = \frac{1 \times 700 \times 10^{-9}}{2.00 \times 10^{-6}} = 0.350$, so

$\theta_{red} = 20.5°$

For violet light, rearrange $d \sin\theta = n\lambda$ so

$\sin\theta = \frac{n\lambda}{d} = \frac{1 \times 400 \times 10^{-9}}{2.00 \times 10^{-6}} = 0.200$, so

$\theta_{violet} = 11.5°$

Therefore angular separation
$= 20.5° - 11.5° = 9.0°$.

b The third-order maximum for violet light is deflected through a smaller angle than the second-order maximum for red light.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Answers to SAQs

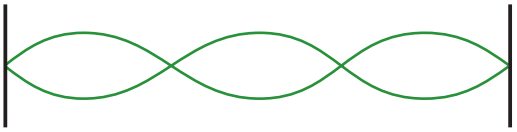
Chapter 15

1 **a** Wavelength of stationary wave = 2 × distance between nodes = 50 cm

b Distance from node to adjacent antinode = 0.5 × distance between nodes = 12.5 cm

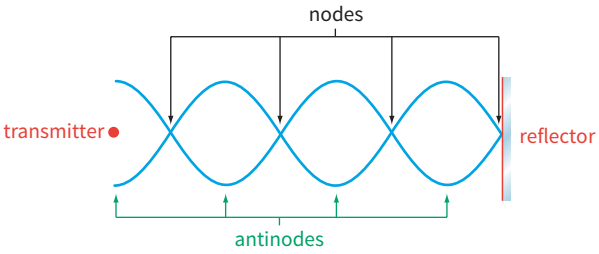
2 **a** One complete wavelength is shown in the image, so wavelength = 60 cm
Separation of two antinodes = $\lambda/2$ = 30 cm

b



c Length of string = 60 cm, so to produce three antinodes, $3\lambda/2$ = 60 cm, meaning λ = 40 cm

3 **a**



b Distance between antinodes = 14 mm = $\lambda/2$, so wavelength λ = 28 mm. Frequency
$f = \frac{v}{\lambda} = \frac{3.0 \times 10^8}{0.028} = 1.07 \times 10^{10}$ Hz ≈ 11 GHz

4 In both cases, waves are reflected (by the metal sheet or by the water). The outgoing and reflected waves combine to produce a stationary wave pattern.

5 **a** It is much easier to detect where sound falls to zero than where sound is a maximum.

b Increased accuracy – if the wavelength is short, it is difficult to measure just one wavelength.

6 **a** Three antinodes between them means the distance between the two measured nodes = $3\lambda/2$ = 20 cm, so wavelength
λ = 13.3 cm ≈ 13 cm

b Speed $v = f\lambda = 2500 \times 0.13 = 325\ \text{m s}^{-1} \approx 330\ \text{m s}^{-1}$

7 **a** For the shorter tube, $\lambda/4 = l_1 + c$, where l_1 is the length of the tube and c is the end correction.
For the longer tube, $3\lambda/4 = l_2 + c$
Subtract the first equation from the second to give $\lambda/2 = l_2 - l_1$, meaning
$\lambda = 2 \times (0.388 - 0.126) = 0.524$ m ≈ 52 cm

b For the shorter tube, the end correction
$c = \frac{\lambda}{4} - l_1 = \frac{52.4}{4 - 12.6} = 0.50$ cm

c Speed $v = f\lambda = 630 \times 0.524 = 330\ \text{m s}^{-1}$

Answers to SAQs

Chapter 16

1 If there were air molecules in the container, the α-particles would scatter off them as well and distort the results. The α-particles may also be absorbed by 6 cm of air.

2 **a** More back-scattered, because of greater chance of close approach to gold nucleus.

b Fewer back-scattered, because their inertia would tend to carry them forwards.

c Fewer back-scattered, because the repulsive force would be smaller. (Note: gold and silver atoms occupy roughly equal volumes.)

3 Density = $\frac{\text{mass}}{\text{volume}}$, so 193 g of gold occupy

volume = $\frac{\text{mass}}{\text{density}} = \frac{0.193}{19\,700} = 9.8 \times 10^{-6}\,\text{m}^3$

Therefore the volume occupied by one atom

$= \frac{\text{volume of 193 g}}{\text{number of atoms in 193 g}}$

$= \frac{9.8 \times 10^{-6}}{6.02 \times 10^{23}} = 1.6 \times 10^{-29}\,\text{m}^3$

Volume of one atom, $V = \frac{4}{3}\pi r^3$, so

radius = $\sqrt[3]{\frac{3V}{4\pi}} = \sqrt[3]{\frac{3 \times 1.6 \times 10^{-29}}{4 \times \pi}} = 1.6 \times 10^{-10}\,\text{m}$

This assumes there is little empty space between atoms.

4 Number of neutrons
= nucleon number – proton number

a 7

b 44

c 60

d 118

e 122

5 **a** $+e$

b no charge

c $+Ze$, where Z is the proton number

d no charge

e $+2e$

6 Number of neutrons
= nucleon number – proton (or atomic) number
So isotopes of uranium with nucleon numbers 235 and 238 contain 143 and 146 neutrons, respectively.

7 **a** Proton number 80 for all.
Neutron numbers 116, 118, 119, 120, 121, 122, 124.

b Average relative atomic mass
= average nucleon number

$= \frac{[(196 \times 0.2) + (198 \times 10) + (199 \times 16.8) + (200 \times 23.1) + (201 \times 13.2) + (202 \times 29.8) + (204 \times 6.9)]}{100}$

= 200.6

8 They are grouped into isotopes as follows:
A and E; C; D, F and G; B and H.

A = $^{44}_{20}\text{Ca}$ isotope of calcium
B = $^{50}_{23}\text{V}$ isotope of vanadium
C = $^{46}_{21}\text{Sc}$ isotope of scandium
D = $^{46}_{22}\text{Ti}$ isotope of titanium
E = $^{46}_{20}\text{Ca}$ isotope of calcium
F = $^{48}_{22}\text{Ti}$ isotope of titanium
G = $^{50}_{22}\text{Ti}$ isotope of titanium
H = $^{51}_{23}\text{V}$ isotope of vanadium

9 **a** Gravitational force acts between all protons and neutrons in a nucleus, as both types of particle have mass.

b Electrostatic force acts only between protons in a nucleus, as neutrons are uncharged.

c Strong nuclear force acts between all protons and neutrons in a nucleus.

10 Electric charge is conserved in the decay, as indicated by the $+e$ charge on the proton and the $+e$ charge on the emitted positron. If the neutrino were charged, this would violate conservation of electric charge.
The mass of the proton ≈ mass of the neutron, and each has much greater mass than a positron. For mass to be conserved in the decay, the neutrino must have a very small mass indeed.

11 Mass: a proton has much greater mass (about 1840×) than the positron; a photon has no mass.
Charge: a proton has charge +*e*, a positron has charge +*e* and a photon has no charge.
Particle or anti-particle:
proton = particle, positron = anti-particle, photon = electromagnetic radiation.

12 **a** Hadrons: proton, neutron, pion, kaon, pi^+ meson, phi meson

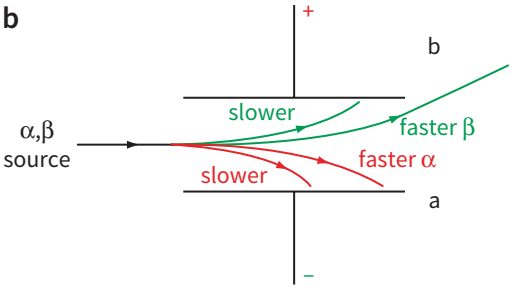
b Leptons: electron, positron, muon, electron neutrino, electron antineutrino

13 Hadrons are affected by the strong nuclear force, leptons are not.
Hadrons are composed of quarks and antiquarks, leptons are not.

14 **a** A β-particle has less charge, is smaller and travels faster, so is less likely to cause ionisation of an atom and thus travels further before losing all its energy.

b Air is much less dense and so less ionisation is caused per unit distance travelled.

15 **a, b**



The α-particles are deflected much less than the β-particles.

16 Most strongly ionising implies that many more collisions occur, so there is greater loss of momentum and therefore less penetration.

17 The α-particles pass between two plates connected to an electric power supply.
The α-particles cause ionisation in the air between the plates and thus a small current flows. When smoke enters the device, the α-particles are absorbed and there is less ionisation of the air and thus a smaller current in the external circuit. The circuit then switches on the alarm. Alpha-radiation is most suitable because it is the most strongly ionising. It causes a greater current to flow, and also α-particles are more likely to be absorbed by smoke. Also, the range of the α-particles is so small that they are unlikely to constitute a hazard to the user.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Answers to SAQs

Chapter 17

1 a Full circular face of clock = 360°
Clock face divided into twelve sections, so angular displacement of hour hand per hour = 360°/12 = 30°

b i Angular displacement of minute hand = half the clock face = 180°

ii Angular displacement of hour hand = 3.5 × 30° = 105°

2 a Angle of 30° = $\frac{30 \times \pi}{180}$ = 0.52 rad

Angle of 90° = $\frac{90 \times \pi}{180}$ = 1.57 rad

Angle of 105° = $\frac{105 \times \pi}{180}$ = 1.83 rad

b Angle of 0.5 rad = $\frac{0.5 \times 180}{\pi}$ = 28.6°

Angle of 0.75 rad = $\frac{0.75 \times 180}{\pi}$ = 43.0°

Angle of π rad = $\frac{\pi \times 180}{\pi}$ = 180°

Angle of $\frac{1}{2}\pi$ rad = $\frac{\frac{1}{2}\pi \times 180}{\pi}$ = 90°

c Angle of 30° = $\frac{30 \times \pi}{180} = \frac{\pi}{6}$ rad

Angle of 120° = $\frac{120 \times \pi}{180} = \frac{2\pi}{3}$ rad

Angle of 270° = $\frac{270 \times \pi}{180} = \frac{3\pi}{2}$ rad

Angle of 7200° = $\frac{720 \times \pi}{180}$ = 4π rad

3 The magnitude of the velocity remains the same (the speed is constant).

4 a Speed is a scalar quantity and is constant, so change in speed = 0 m s^{-1}

b Velocity changes to be in the opposite direction, so change in velocity = 0.2 – (–0.2) = 0.2 + 0.2 = 0.4 m s^{-1}

5 Second hand turns through 360° in 1 minute, or 2π rad in 1 minute, which is $\frac{2\pi}{60}$ rad in 1 second.
Angular velocity $\omega = \frac{2\pi}{60}$ = 0.105 rad s^{-1}

6 a Number of revolutions per second = 1200/60 = 20 rev s^{-1}

b Angular velocity = 20 × 2π = 40π rad s^{-1} = 130 rad s^{-1}

7 Speed $v = \omega r$ = 0.105 × 1.8 = 0.19 cm s^{-1}

8 a Angular velocity
$\omega = \frac{\Delta\theta}{\Delta t} = \frac{90 \times \pi}{180 \times 15}$ = 0.105 rad s^{-1}

b Speed $v = \omega r$ = 0.105 × 50 = 5.2 m s^{-1}

9 Rearrange $v = \omega r$ to give angular velocity
$\omega = \frac{v}{r} = \frac{7800}{7000\,000} = 1.1 \times 10^{-3}$ rad s^{-1}

10 a The gravitational pull of the Earth on the Moon.

b The frictional force of the road on the wheels.

c Tension in the string supporting the pendulum.

11 There will be no frictional force between the road and the wheels. If the driver turns the steering wheel, the car will continue in a straight line.

12 Speed and kinetic energy are scalar quantities, the others are vectors. Speed is constant; velocity has a constant magnitude but continuously changing direction (the direction is tangential to the circle); kinetic energy is constant; momentum has a constant magnitude but continuously changing direction (the direction is tangential to the circle); the centripetal force has a constant magnitude but continuously changing direction (the direction is always towards the centre of the circle); the centripetal acceleration behaves in the same way as the centripetal force.

13 Acceleration $a = \frac{v^2}{r}$
$v = \omega r$
$a = \frac{(\omega r)^2}{r}$
so $a = \omega^2 r$

14 Distance travelled for one complete orbit
$s = 2\pi r = 2 \times \pi \times 6400\,000 = 4.02 \times 10^7$ m
Rearrange speed $v = \frac{s}{t}$ to give time
$t = \frac{s}{v} = \frac{4.02 \times 10^7}{7920} = 5.08 \times 10^3$ s (84.6 min)

15 Rearrange centripetal force $F = \frac{mv^2}{r}$ to give

maximum speed $v = \sqrt{\frac{Fr}{m}} = \sqrt{\frac{8.0 \times 0.30}{0.2}}$

$= 3.46\,\text{m s}^{-1} \approx 3.5\,\text{m s}^{-1}$

16 **a** Centripetal force

$F = ma = 350 \times 10^3 \times 8.8 = 3.1 \times 10^6\,\text{N}$

b Rearrange centripetal force $F = \frac{mv^2}{r}$ to give

speed $\sqrt{\frac{Fr}{m}} = \sqrt{\frac{3.1 \times 10^6 \times (340 + 6400) \times 10^3}{350 \times 10^3}}$

$= 7.7 \times 10^3\,\text{m s}^{-1}$

c Rearrange speed $v = \frac{s}{t}$ to give

time $t = \frac{s}{v} = \frac{2\pi r}{v} = \frac{2 \times \pi \times (340 + 6400) \times 10^3}{7.7 \times 10^3}$

$= 5.50 \times 10^3\,\text{s} \approx 1.5\,\text{h}$

d Number of orbits $= \frac{\text{length of day}}{\text{orbit time}} = \frac{24\,\text{h}}{5.5 \times 10^3\,\text{s}}$

$= 15.7$ times

17 **a** Time taken for one revolution

$t = \frac{1}{3} = 0.333\,\text{s}$

Speed, $v = \frac{s}{t} = \frac{2\pi r}{t} = \frac{2 \times \pi \times 0.5}{0.333}$

$= 9.42\,\text{m s}^{-1} \approx 9.4\,\text{m s}^{-1}$

b Centripetal acceleration

$a = \frac{v^2}{r} = \frac{(9.42)^2}{0.5} = 178\,\text{m s}^{-2} \approx 180\,\text{m s}^{-2}$

c Tension in string, $F = ma = 0.40 \times 178 = 71\,\text{N}$

18 **a** Speed, $v = \frac{s}{t} = \frac{2\pi r}{t} = \frac{2 \times \pi \times 2.3 \times 10^{11}}{687 \times 24 \times 60 \times 60}$

$= 2.43 \times 10^4\,\text{m s}^{-1} \approx 24\,\text{km s}^{-1}$

b Centripetal acceleration

$a = \frac{v^2}{r} = \frac{(2.43 \times 10^4)^2}{2.3 \times 10^{11}} = 2.57 \times 10^{-3} \approx 2.6 \times 10^{-3}\,\text{m s}^{-2}$

c Gravitational force

$F = ma = 6.4 \times 10^{23} \times 2.57 \times 10^{-3} = 1.6 \times 10^{21}\,\text{N}$

19 The tension in the string must have a vertical component to balance the weight of the bung.

20 In level flight, lift balances the weight. During banking, the vertical component of the lift is less than the weight, so the aeroplane loses height unless the speed can be increased to provide more lift.

21 The normal contact force of the wall of the slide has a horizontal component, which provides the centripetal force. If you are going fast, you need a bigger force, so the horizontal component must be greater. This happens as you move up the curved wall of the slide.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Answers to SAQs

Chapter 18

1 a Size of gravitational force of attraction
$= \frac{GMm}{r^2} = \frac{6.67\times10^{-11}\times0.1\times0.1}{(0.01)^2} = 6.67\times10^{-9}\,\text{N}$

b Size of gravitational force of attraction
$= \frac{GMm}{r^2} = \frac{6.67\times10^{-11}\times5.0\times10^{10}}{(4.0\times10^{9})^2} = 1.0\times10^{-8}\,\text{N}$

c Size of gravitational force of attraction
$= \frac{GMm}{r^2} = \frac{6.67\times10^{-11}\times6.0\times10^{24}\times1.4\times10^{4}}{(6.8\times10^{6})^2}$
$= 1.2\times10^{5}\,\text{N}$

2 Estimate that each person has mass 70 kg and that they are seated 0.5 m apart.
Size of gravitational force of attraction
$= \frac{GMm}{r^2} = \frac{6.67\times10^{-11}\times70\times70}{(0.5)^2} = 1.3\times10^{-6} \approx 10^{-6}\,\text{N}.$
A human with mass 70 kg weighs about 700 N on Earth. Their weight is greater than their mutual attraction by a factor of approximately 10^9.

3 Difference in weight ΔF = weight at surface of Earth – weight at top of Everest $= mg - \frac{GMm}{r^2}$
$= (100\times9.81) - \frac{6.67\times10^{-11}\times6.0\times10^{24}\times100}{(6.409\times10^{6})^2} \approx 7\,\text{N}.$
This would be just about detectable, although other factors such as dehydration would be more significant.

4 a i Gravitational field strength $g = -\frac{GM}{r^2}$
$= -\frac{6.67\times10^{-11}\times7.4\times10^{22}}{(1.74\times10^{6})^2} = -1.6\,\text{N kg}^{-1}$

ii Gravitational field strength $g = -\frac{GM}{r^2}$
$= -\frac{6.67\times10^{-11}\times2.0\times10^{30}}{(7.0\times10^{8})^2} = -270\,\text{N kg}^{-1}$

b Gravitational field strength is very weak on the Moon, so gas molecules will have enough energy to escape from the Moon, whereas the Sun has a very high field strength and therefore pulls gas molecules very close together.

5 a Gravitational field strength $g = -\frac{GM}{r^2}$
$= -\frac{6.67\times10^{-11}\times6.0\times10^{24}}{(3.8\times10^{8})^2} = -2.8\times10^{-3}\,\text{N kg}^{-1}$

b Force $= mg = 7.4\times10^{22}\times(-2.8\times10^{-3})$
$= -2.1\times10^{20}\,\text{N}$
Acceleration $= \frac{F}{m} = 2.8\times10^{-3}\,\text{m s}^{-2}$

6 Gravitational field strength $g = -\frac{GM}{r^2}$ so if the gravitational field strength of Earth at its surface is $g_E = -9.81\,\text{N kg}^{-1}$, the gravitational field strength of Jupiter at its surface is
$g_J = \frac{9.81\times320}{(11.2)^2} = -25\,\text{N kg}^{-1}$

7 Gravitational field strength due to the Sun, on the surface of the Earth, is
$g_S = -\frac{GM}{r^2} = -\frac{6.67\times10^{-11}\times2.0\times10^{30}}{(1.5\times10^{11})^2} = -5.9\times10^{-3}\,\text{N kg}^{-1}$
Gravitational field strength due to the Moon, on the surface of the Earth, is
$G_M = -\frac{GM}{r^2} = -\frac{6.67\times10^{-11}\times7.4\times10^{22}}{(3.8\times10^{8})^2} = -3.4\times10^{-5}\,\text{N kg}^{-1}$
So the Sun has a greater pull on each kilogram of the seawater.

8 a Size of gravitational force on baby caused by Mars
$F_M = \frac{GMm}{r^2} = \frac{6.67\times10^{-11}\times6.4\times10^{23}\times4.0}{(1.0\times10^{11})^2} = 1.7\times10^{-8}\,\text{N}$

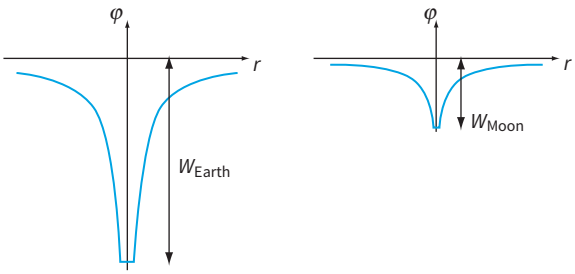
b Size of gravitational force on baby caused by mother
$F_m = \frac{GMm}{r^2} = \frac{6.67\times10^{-11}\times50\times4.0}{(0.4)^2} = 8.3\times10^{-8}\,\text{N}$

9 Closer to the Moon, because the Moon has less mass than the Earth.
We are looking for a distance x from the centre of the Earth where the two gravitational fields balance out. If the Earth's mass is M_E and the Moon's mass is M_M, and the Earth-to-Moon distance is d, then
$\frac{GM_E}{x^2} = \frac{GM_M}{(d-x)^2}$
Rearranging this equation gives $M_E(d-x)^2 = M_M x^2$
meaning $x = \frac{d\sqrt{M_E}}{\sqrt{M_M}+\sqrt{M_E}} = \frac{3.8\times10^{8}\times\sqrt{6.0\times10^{24}}}{\sqrt{6.0\times10^{24}}+\sqrt{7.4\times10^{22}}}$
$= 3.4\times10^{5}\,\text{km}$ from the Earth.

10 a Gravitational potential $\varphi = -\frac{GM}{r}$
$= -\frac{6.67\times10^{-11}\times6.0\times10^{24}}{6.4\times10^{6}} = -6.3\times10^{7}\,\text{J kg}^{-1}$

b Gravitational potential $\varphi = -\frac{GM}{r}$
$= -\frac{6.67\times10^{-11}\times7.4\times10^{22}}{1.74\times10^{6}} = -2.8\times10^{6}\,\text{J kg}^{-1}$

c



Potential well for Earth Potential well for Moon

d W_{Earth} is the energy needed for each kilogram, initially on the Earth's surface, to escape from the Earth's field, which can be seen to be much greater than W_{Moon}, the energy needed for each kilogram, initially on the Moon's surface, to escape from its field. So the rocket does not need to carry so much fuel.

11 Orbital radius = radius of Earth + altitude of satellite above surface

$= 6.4 \times 10^6 + 2.0 \times 10^5 = 6.6 \times 10^6$ m

Orbital speed $v = \sqrt{\left(\frac{GM}{r}\right)}$

$= \sqrt{\frac{6.67 \times 10^{-11} \times 6.0 \times 10^{24}}{6.6 \times 10^6}} = 7.8 \text{ km s}^{-1}$

12

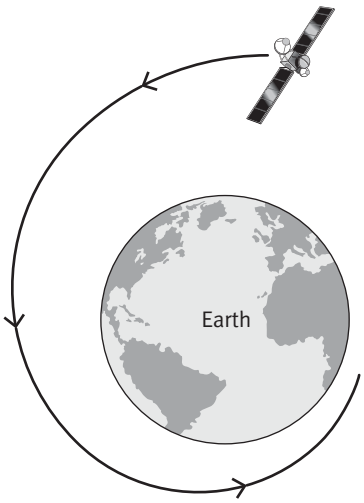


Diagram showing the satellite spiralling in towards Earth. The satellite needs to fire small thruster rockets to maintain its speed and orbit.

13 Radius r of a geostationary orbit is given by

$r^3 = \frac{GMT^2}{4\pi^2} = \frac{6.67 \times 10^{-11} \times 6.4 \times 10^{23} \times (24.6 \times 60 \times 60)^2}{4 \times \pi^2}$

$= 8.7 \times 10^{21}$ m^3, so $r = 20\,600$ km

14 Distance travelled by signal sent to and returned by satellite

$= 2 \times (42\,300\,000 - 6\,400\,000) = 7.18 \times 10^7$ m

So extra time taken by signal travelling via satellite $= \frac{\text{distance}}{\text{speed}} = \frac{7.18 \times 10^7}{3.0 \times 10^8} = 0.24$ s.

So time delay = 0.24 s, signals travel 30% slower in cables but the distance is much shorter.

Answers to SAQs

Chapter 19

1 **a** Forced
b Free
c Free
d Forced

2 Curved (Figure 19.6a)

3 Amplitude = 10 cm, period = 120 ms (0.12 s), frequency = $\frac{1}{T} = \frac{1}{0.120} = 8.3$ Hz

4 **a** Half an oscillation
b The waves have different frequencies so the phase difference is continuously changing.

5 The trolley is the mass; the central position of the trolley is the equilibrium position; the resultant restoring force of the springs is the force.

6 The restoring force is not proportional to the distance from the equilibrium point. When the person is not in contact with the trampoline the restoring force is equal to the person's weight, which is constant.

7 **a** Amplitude = 0.02 m
b Time period = 0.40 s
c Maximum velocity = 0.31 m s^{-1}
d Maximum acceleration = 5.0 m s^{-2}

8 At the extreme left of the oscillation (i.e. maximum negative displacement), the acceleration is positive (towards the right).

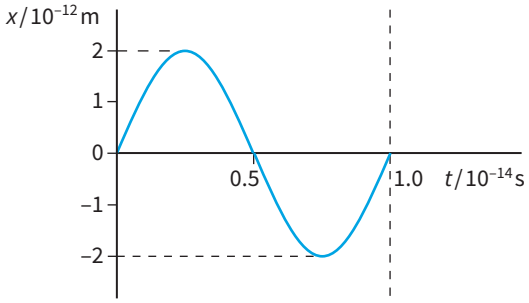
9 Gradient = 0, velocity = 0

10 **a** 0 cm s^{-1}
b 47 cm s^{-1}
c 0 cm s^{-2}

11 **a** Period $T = \frac{1}{f} = \frac{1.0}{2} = 0.5$ s
b 2 Hz
c Angular frequency $\omega = 2\pi f = 4\pi$ rad s^{-1}

12 **a** 0.20 m
b 0.4 s
c Frequency = $\frac{1}{T} = \frac{1}{0.4} = 2.5$ Hz
d Angular frequency $\omega = 2\pi f = 5\pi$ rad s^{-1}
e –0.1 m
f 0 m s^{-1}
g 3.1 m s^{-1}

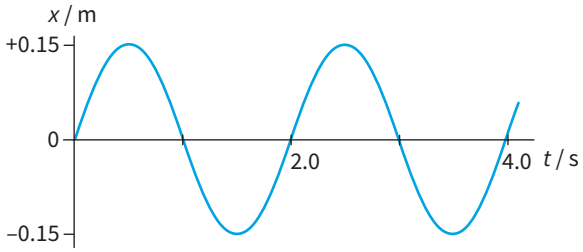
13 **a**



b Gradient at steepest point, giving approximately 1.3×10^{3} m s^{-1}

14 **a** 3×10^{-4} m (0.3 mm)
b Frequency = $\frac{240\pi}{2\pi} = 120$ Hz
c Period = $\frac{1}{f} = \frac{1}{120} = 8.3 \times 10^{-3}$ s

15 **a** $x = 0.15\cos(\pi t)$
b



16 **a** Angular frequency
$\omega = 2\pi f = 2 \times \pi \times 1.4 = 8.80$ s^{-1}
So $a = -\omega^2 x = -77.4x$
b Acceleration = $-77.4 \times 0.050 = 3.9$ m s^{-2}

17 Angular frequency $\omega = 2\pi f$ so, in acceleration equation, $a = -\omega^2 x = -4\pi^2 f^2 x$
meaning $4\pi^2 f^2 = 300$ s^{-2}, therefore
$f = \sqrt{\frac{300}{4\pi^2}} = 2.76$ Hz ≈ 2.8 Hz

18 **a** **i** 2.0 s

ii Frequency $= \frac{1}{T} = \frac{1}{2.0} = 0.5$ Hz

iii Angular frequency, $\omega = 2\pi f = \pi$ rad s^{-1} or 3.14 rad s^{-1}

b $a = -\omega^2 x = -9.87x$ or $a = -\pi^2 x$

c Maximum speed $= \omega x_0 = 3.14 \times 12$
$= 37.6$ cm s$^{-1} \approx 38$ cm s^{-1}

d Speed at $x = 6$ cm is $v = \omega\sqrt{[(x_0)^2 - x^2]}$
$= 3.14 \times \sqrt{[(12)^2 - (6)^2]} = 32.6$ cm s$^{-1} \approx 33$ cm s^{-1}

19 **a** The restoring force $= kx$ (from Hooke's law), $a \propto F$, therefore $a \propto x$. The force acts in the opposite direction to the displacement.

b $a = -\frac{F}{m} = -\frac{kx}{m} = -\omega^2 x$

$\omega^2 = \frac{k}{m}$

$\omega = \sqrt{\frac{k}{m}}$

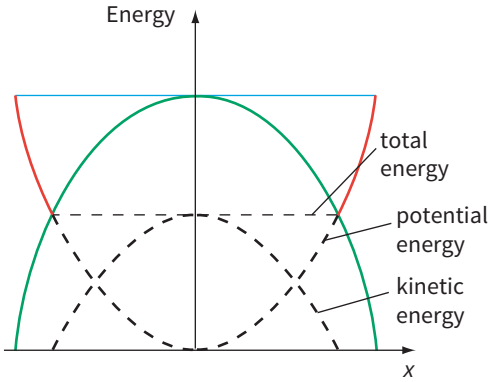
$f = \frac{\omega}{2\pi} = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$

$T = \frac{1}{f} = 2\pi\sqrt{\frac{m}{k}}$

20 **a** Gravitational potential energy

b Gravitational potential energy is transferred to kinetic energy, which reaches a maximum when the bob passes through the lowest point; then k.e. is converted to g.p.e. once more.

21



22 **a** 0.35 m s^{-1}

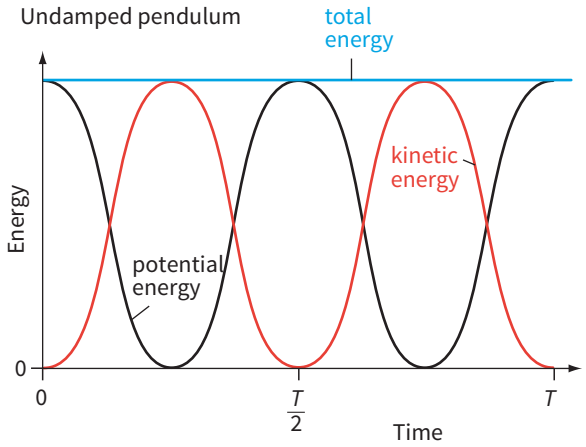
b Maximum kinetic energy $= \frac{1}{2}mv^2$
$= \frac{1}{2} \times 2.0 \times (0.35)^2 = 0.12$ J

c Maximum potential energy
= maximum kinetic energy = 0.12 J

d Maximum acceleration 0.17 m s^{-2}

e Maximum restoring force $= ma$
$= 2.0 \times 0.17 = 3.4$ N

23 **a**



b The total energy of the pendulum would decrease gradually and hence the maximum values of k.e. and p.e. would also decrease.

24 *Any three from:*

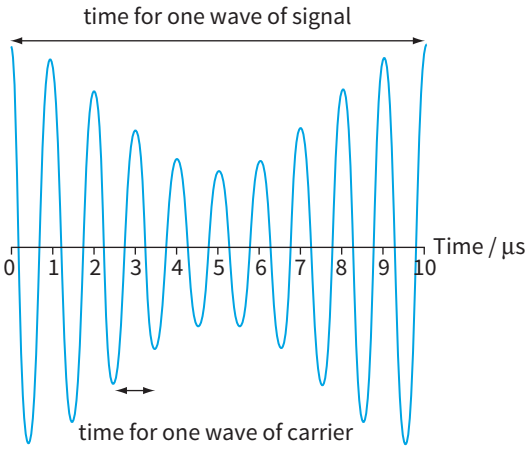
Example	Useful or problem?	What is resonating?
Buildings in earthquake	Problem	Mechanical structure is forced to oscillate by energy from waves of earthquake.
Components in engines	Problem	At certain rates of rotation, parts of an engine may resonate mechanically; the resonance is driven by the energy output of the engine. This can lead to components cracking or disintegrating, with dangerous consequences.
Positive feedback in amplification systems (high-pitched squealing sound)	Problem	Microphone is held too close to a loudspeaker that is emitting waves of the same frequency as the microphone is tuned to, so the waves from the loudspeaker force the amplifier to resonate.
Tuned radio	Useful	Electric signal in circuit is forced to oscillate by incoming radio waves.
Microwave cooker	Useful	Water molecules are forced by microwaves.
Magnetic resonance in atoms	Useful	Nuclei in atoms behave as magnets; they can be made to resonate by electromagnetic waves. Each nucleus resonates at a different frequency, so the structures of molecules can be determined.

Answers to SAQs

Chapter 20

1 Carrier wave frequency is halved, $f_c = 50\,\text{kHz}$
Signal wave frequency is halved, $f_s = 5\,\text{kHz}$

2



3 **a** The difference in amplitude of the carrier wave increases, i.e. the heights of the lower peaks of the waveform fall and the heights of the higher peaks rise.

b The time interval decreases between successive instances when the signal waveform is at its very highest.

4 **a** The maximum difference in frequency between the signal and the carrier wave increases, i.e. the signal has higher and lower frequencies.

b The signal increases and decreases in frequency more times per second.

5 **a** Maximum frequency shift
= frequency deviation × amplitude
= 8.0 × 3.0 = 24 kHz

b Maximum frequency of modulated carrier wave = frequency of carrier wave + maximum frequency shift = 500 + 24 = 524 kHz

c Number of times per second the modulated carrier wave increases and decreases in frequency = frequency of signal
= 16 000 times a second

6 **a** Amplitude modulation

b Carrier frequency = 40 kHz

c Frequency of signal used to modulate carrier wave = 5 kHz

d Bandwidth = 45 – 35 = 10 kHz

7 **a** Width of frequency spectrum
= 1700 – 530 = 1170 kHz
Number of stations possible
$= \frac{\text{width of frequency spectrum}}{\text{bandwidth per station}} = \frac{1170}{9} = 130$

b Very few FM stations are possible. Number $= \frac{1170}{200} = 5.85$, so only five are possible.

8 If you are concerned about the quality of the signal that you are listening to, then FM is better. If you want more stations and are not in line of sight with the transmitter, then AM is better.

9 AM covers a larger area than FM. FM can reach all the inhabitants of a town but the expense of aerials and transmitters means this is not possible in rural areas.

10 **a** $14 = (1 \times 8) + (1 \times 4) + (1 \times 2)$; in binary, 1110

b $16 = (1 \times 16)$; in binary, 10000

11 **a** Binary 1111 is $(1 \times 8) + (1 \times 4) + (1 \times 2) + (1 \times 1) = 15$

b Binary 0001011 is $(1 \times 8) + (1 \times 2) + (1 \times 1) = 11$

12 **a** The value on the vertical axis (the voltage) has only one of two values, 0 or 1.

b Digital signals can be more easily regenerated, can be stored and processed more easily, errors in transmission can be detected and digital electronic systems are more reliable.

c Noise is introduced, e.g. by random thermal motion of electrons. The signal is attenuated due to energy loss (I^2Rt) in the wires.

13 $\text{Attenuation} = 10 \lg\left(\frac{P_1}{P_2}\right) = 10 \lg\left(\frac{5.0}{0.0002}\right)$
$= 44\,\text{dB}$ (or a gain of $-44\,\text{dB}$).

14 Attenuation = $10 \lg\left(\frac{P_1}{P_2}\right)$

Substituting known values gives

$30 = 10 \lg\left(\frac{6.0 \times 10^{-3}}{P_2}\right)$ meaning $3 = \lg\left(\frac{6.0 \times 10^{-3}}{P_2}\right)$

Taking inverse logs, $10^3 = \left(\frac{6.0 \times 10^{-3}}{P_2}\right)$

Rearranging gives $P_2 = \left(\frac{6.0 \times 10^{-3}}{103}\right) = 6.0 \times 10^{-6}\,\text{W}$

$= 6.0 \times 10^{-3}\,\text{mW}$

15 Signal-to-noise ratio = $10 \lg(1.0) = 0.0\,\text{dB}$

16 **a** Given that gain and attenuation expressed in decibels are logarithmic, we can say that overall gain

= gain in amplifier + gain in cable

= gain in amplifier – attenuation in cable

= 30 – 18 = 12 dB

b Substitute values in gain = $10 \lg\left(\frac{P_2}{P_1}\right)$

So $12 = 10 \lg\left(\frac{P_2}{1.0 \times 10^{-3}}\right)$,

therefore $\lg\left(\frac{P_2}{1.0 \times 10^{-3}}\right) = 1.2$

Take inverse logs to give output power

$P_2 = 10^{1.2} \times 1.0 \times 10^{-3} = 15.8\,\text{mW} \approx 16\,\text{mW}$

17 **a** Coaxial cable contains a central copper core surrounded by an insulator and a mesh, or braid, made of copper wires made into a tube.

An analogue signal can have any value, within limits, and is not restricted to just a few values.

A space wave is an electromagnetic wave above 30 MHz that is not reflected or refracted back to Earth by the ionosphere and may be received by a satellite in space or be detected by line of sight within the Earth's atmosphere.

A sky wave is an electromagnetic wave that is reflected or refracted back to the Earth by the ionosphere.

b Space wave frequency greater than 30 MHz, wavelength less than 10 m

Sky wave frequency between 3 and 30 MHz, wavelength between 10 and 100 m

c

Method	Advantage	Disadvantage
coaxial	less attenuation, noise, interference than wire pair	more expensive than wire pair
analogue signal	no ADC and DAC conversion required	more noise, interference than digital
space wave	can be received by satellite	satellite use expensive
sky wave	worldwide reception possible with multiple reflections	reflection by the ionosphere is not reliable

Other answers are possible.

Answers to SAQs

Chapter 21

1 Bonds are still relatively strong (stronger than gas but weaker than a solid) and the molecules are free to move through the body of the liquid.

2 **a** Energy has to be supplied to convert the liquid into vapour – this takes time.

b When ice is converted to liquid water, only a few intermolecular bonds are broken, whereas when liquid is converted to vapour all the molecules are totally separated. This requires more energy.

c Dogs do not sweat through their skin like humans; instead they sweat from their tongues. The sweat evaporates and cools the dog. The panting blows air across the tongue, increasing the rate of evaporation.

3 **a** Change in internal energy = energy supplied by heating + energy supplied by doing work
= 250 + 500 = +750 kJ

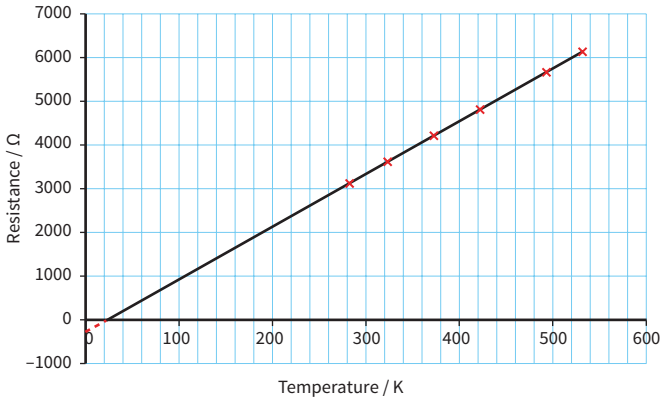
b Change in internal energy = energy supplied by heating + energy supplied by doing work
= energy supplied by heating – work done by the gas on its surroundings
= 250 – 200 = +50 kJ

4 **a** T(K) = θ(°C) + 273.15, so
0 °C ≈ 273 K
20 °C ≈ 293 K
120 °C ≈ 393 K
500 °C ≈ 773 K
–23 °C ≈ 250 K
–200 °C ≈ 73 K

b θ(°C) = T(K) – 273.15, so
0 K ≈ –273 °C
20 K ≈ –253 °C
100 K ≈ –173 °C
300 K ≈ +27 °C
373 K ≈ +100 °C
500 K ≈ +227 °C

5

Temperature / °C	Resistance / Ω	Temperature / K
10	3120	283
50	3600	323
75	3900	348
100	4200	373
150	4800	423
220	5640	493
260	6120	533



Absolute zero is the temperature at which all substances have the minimum internal energy; the kinetic energy of the atoms of copper is zero and their electrical potential energy is at a minimum. Therefore we would expect the resistance to be zero.

6 **a** Calibration

b Range

c Linearity

d Sensitivity

7 Energy $E = mc\Delta\theta = 5.0 \times 4180 \times (100 - 20)$
= 1.67 MJ ≈ 1.7 MJ

8 Energy required for lead
$E_{lead} = mc\Delta\theta = 2.0 \times 126 \times 30 = 7.56$ kJ
Energy required for copper
$E_{copper} = mc\Delta\theta = 4.0 \times 380 \times 5.0 = 7.60$ kJ
So the copper block requires more energy.

9 Energy supplied in the time is
E = power × time = 50 × 4.0 × 60 = 12 kJ
Rearrange energy $E = mc\Delta\theta$ to give specific heat capacity:

$$c = \frac{E}{m\Delta\theta} = \frac{12\,000}{1.2} \times (45 - 22)$$
$$= 435 \text{ J kg}^{-1}\text{ K}^{-1} \approx 440 \text{ J kg}^{-1}\text{ K}^{-1}$$

Cambridge International AS and A Level Physics © Cambridge University Press 2014

10 At higher temperatures, the rate of energy loss to the surroundings is greater, therefore the temperature rise is slower.

11 Systematic – it can (theoretically) be removed by perfect insulation.

12 Mass of water used,
m = mass of (beaker + water) – mass of beaker
= 0.672 – 0.150 = 0.522 kg
Power of heater, $P = VI = 11.4 \times 3.9 = 44.5$ W
So in 13.0 min, energy supplied by heater
$= P \times t = 44.5 \times 13.0 \times 60 = 34.7$ kJ
Temperature rise, $\Delta\theta = 30.2 - 18.5 = 11.7$ K
So, from this experiment, the heat capacity of water is

$$c = \frac{E}{m\Delta\theta} = \frac{34\,700}{0.522} \times 11.7 = 5680 \approx 5700 \text{ J kg}^{-1}\text{ K}^{-1}$$

The biggest source of error will be energy loss due to poor insulation. There will also be an error because we have ignored the specific heat capacity of the beaker.

13 **a** AB: solid. BC: solid + liquid. CD: liquid.

b Internal energy increases in all three sections.

c The specific heat capacity is greater when it is a solid. The gradient is greater when it is a liquid, so it takes less time to raise the temperature by the same amount, and therefore less energy.

14 Energy needed to change ice into water is $E = mL = 0.0020 \times 330\,000 = 660$ J. When a solid melts, only about one bond per atom or molecule is broken. On boiling, several remaining bonds are broken, requiring more energy.

15 Rate of loss of mass = 2.25 g per minute

$$= \frac{0.002\,25}{60} = 3.75 \times 10^{-5} \text{ kg s}^{-1}$$

Rate of supply of energy to alcohol = 40 × 80%
= 32 W
So specific latent heat of vaporisation

$$= \frac{\text{rate of supply of energy}}{\text{rate of loss of mass}} = \frac{32}{3.75 \times 10^{-5}}$$

$$= 853\,000 \text{ J kg}^{-1} \approx 850 \text{ kJ kg}^{-1}$$

Answers to SAQs

Chapter 22

1 Time taken to cross diameter of container once $= \frac{\text{distance}}{\text{speed}} = \frac{0.1}{400} = 2.5 \times 10^{-4}$ s. So, assuming the molecule's speed is constant, that each collision is perfectly elastic and that the molecule always crosses the diameter of the container, the number of collisions $= \frac{1}{t} = \frac{1}{2.5} \times 10^{-4} = 4000\ \text{s}^{-1}$

2 Different numbers of molecules hit one side than the other; the number of molecules striking any area per second will vary; the speed of the molecules hitting will vary; the angle that the molecules hit the smoke particle will vary.

3 The small dots of light (smoke particles) move faster and more erratically. At higher temperature the molecules of air are moving faster and on impact cause a greater change in momentum of the smoke particle.

4 Time taken to cross diameter of container once
$= \frac{\text{distance}}{\text{speed}} = \frac{0.1}{400} = 2.5 \times 10^{-4}$ s.
So, assuming molecule's speed is constant, each collision is perfectly elastic and molecule always crosses diameter of container,
number of collisions $= \frac{1}{t} = \frac{1}{2.5 \times 10^{-4}} = 4000\ \text{s}^{-1}$

5 Pressure increases because more molecules are hitting a unit area of the tyre per second. The molecules have the same speed and each collision causes the same impulse or change in momentum. Since more molecules hit per second, the rate of change of momentum, i.e. the force, increases.

6 At higher temperatures, the molecules have higher internal energy and move faster. The number of collisions per second increases and the change in momentum of each collision also increases. For both these reasons the rate of change of momentum of the molecules increases. The force on the wall is equal and opposite to the rate of change of momentum of the molecules. As the force increases, the pressure inside the can may cause it to explode.

7 **a** Number of atoms in one mole of carbon = Avogadro constant, $N_A = 6.02 \times 10^{23}$

b Number of moles $= \frac{\text{mass}}{\text{mass of one mole}} = \frac{54}{12} = 4.5$
Number of atoms = number of moles $\times N_A$
$= 4.5 \times 6.02 \times 10^{23} = 27.1 \times 10^{23}$

c Number of atoms $= \frac{1000}{12} \times 6.02 \times 10^{23}$
$= 5.0 \times 10^{25}$

8 **a** Single atom has mass $= \frac{\text{molar mass}}{N_A} = \frac{235}{6.02 \times 10^{23}}$
$= 3.90 \times 10^{-22}$ g

b **i** Number of moles $= \frac{\text{mass}}{\text{molar mass}} = \frac{0.020}{235}$
$= 8.51 \times 10^{-5} \approx 8.5 \times 10^{-5}$

ii Number of atoms $= 8.5 \times 10^{-5} \times 6.02 \times 10^{23}$
$= 5.12 \times 10^{19} \approx 5.1 \times 10^{19}$

9 Typical relative atomic mass = 10, so 1 kg contains 100 moles $= 6 \times 10^{25} \approx 10^{26}$ molecules. Note for heavier elements, e.g. iron, $A_r \approx 60$ and number of atoms $\approx 10^{25}$.

10 Boyle's law is $p_1V_1 = p_2V_2$. Rearranging this gives pressure required, $p_2 = \frac{p_1V_1}{V_2}$
$= \frac{120\,000 \times 0.04}{0.025} = 1.92 \times 10^5\ \text{Pa} \approx 1.9 \times 10^5\ \text{Pa}$

11 Rearrange ideal gas equation $pV = nRT$ to give temperature
$T = \frac{pV}{nR} = \frac{1.0 \times 10^4 \times 1.0}{1.0 \times 8.31} = 1200\ \text{K}$

12 **a** Number of moles $= \frac{\text{mass}}{\text{molar mass}} = \frac{100}{28} = 3.57$ mol

b Rearrange ideal gas equation $pV = nRT$ to give volume
$V = \frac{nRT}{p} = \frac{3.57 \times 8.31 \times 293.15}{1.01 \times 10^5} = 0.086\ \text{m}^3$ or $86\ \text{dm}^3$

13 Rearrange ideal gas equation $pV = nRT$ to give volume
$V = \frac{nRT}{p} = \frac{5.0 \times 8.31 \times 473.15}{1.01 \times 10^5} = 0.0195\ \text{m}^3 \approx 0.020\ \text{m}^3$

14 Number of moles of gas = $\frac{\text{number of atoms}}{N_A}$
$= \frac{3.0 \times 10^{24}}{6.02 \times 10^{23}} = 4.98\,\text{mol}$
Rearrange ideal gas equation $pV = nRT$ to give volume
$V = \frac{nRT}{p} = \frac{4.98 \times 8.31 \times 300}{1.20 \times 10^5} = 0.104\,\text{m}^3 \approx 0.1\,\text{m}^3$

15 Number of moles of gas = $\frac{\text{mass}}{\text{molar mass}} = \frac{1000}{32}$
$= 31.25\,\text{mol}$
Rearrange ideal gas equation $pV = nRT$ to give temperature
$T = \frac{pV}{nR} = \frac{1.0 \times 10^5 \times 1.0}{31.25 \times 8.31} = 385\,\text{K}$

16 **a** Rearrange ideal gas equation $pV = nRT$ to give number of moles of hydrogen
$n = \frac{pV}{RT} = \frac{20 \times 1.01 \times 10^5 \times 0.1}{8.31 \times 293.15} = 82.9\,\text{mol}$
So mass of hydrogen
= number of moles × molar mass
$= 82.9 \times 2.0 = 166\,\text{g} \approx 170\,\text{g}$

b If number of moles of oxygen is the same, 82.9 mol, then mass of oxygen
$= 82.9 \times 32 = 2.65\,\text{kg} \approx 2.7\,\text{kg}$

17 Base units of pressure $p = \frac{F}{A} = \frac{ma}{A}$
are $\frac{\text{kg m s}^{-2}}{\text{m}^2} = \text{kg m}^{-1}\,\text{s}^{-2}$
Base units of the right-hand side $\frac{Nm\langle c^2\rangle}{V}$
are $\frac{\text{kg}\,(\text{m s}^{-1})^2}{\text{m}^3} = \text{kg m}^{-1}\,\text{s}^{-2}$

18 **a** Rearrange $p = \frac{Nm\langle c^2\rangle}{3V}$ to give
$\langle c^2\rangle = \frac{3pV}{Nm} = \frac{3 \times 10^5 \times 1.0}{1.29} = 2.3 \times 10^5\,\text{m}^2\,\text{s}^{-2}$

b $\sqrt{\langle c^2\rangle} = \sqrt{2.3 \times 10^5} = 480\,\text{m s}^{-1}$, which is 50% greater than the speed of sound in air.

19 **a** $R = 8.31\,\text{J K}^{-1}$, $N_A = 6.02 \times 10^{23}$
So $k = \frac{R}{N_A} = \frac{8.31\,\text{J K}^{-1}}{6.02 \times 10^{23}} = 1.38 \times 10^{-23}\,\text{J K}^{-1}$

b $E = \frac{3}{2}kT$

20 Mean translational kinetic energy = $\frac{3RT}{2N_A}$
$= \frac{3 \times 8.31 \times 300.15}{2 \times 6.02 \times 10^{23}} = 6.2 \times 10^{-21}\,\text{J}$

21 Rearrange $E_{\text{mean}} = \frac{1}{2}m\langle c^2\rangle = \frac{3RT}{2N_A}$ to give temperature
$T = \frac{2N_A E_{\text{mean}}}{3R} = \frac{2 \times 6.02 \times 10^{23} \times 5.0 \times 10^{-21}}{3 \times 8.31}$
$= 242\,\text{K}$ (or $-31\,^\circ\text{C}$)

22 Temperature is proportional to (average speed)^2. So, if the speed doubles the temperature increases by a factor of $2^2 = 4$.

23 **a** Halved

b Remains the same

24 Mean k.e. = $\frac{3RT}{2N_A} = \frac{3 \times 8.31 \times 293.15}{2 \times 6.02 \times 10^{23}} = 6.1 \times 10^{-21}\,\text{J}$
Mean k.e. = $\frac{1}{2}m\langle c^2\rangle$ where
m (mass of one molecule) = $\frac{\text{molar mass}}{\text{Avogadro number}}$
so average speed of oxygen molecule
$= \sqrt{\frac{2 \times \text{mean k.e.}}{m}} = \sqrt{\frac{2 \times 6.1 \times 10^{-21} \times 6.02 \times 10^{23}}{0.032}}$
$= 480\,\text{m s}^{-1}$
average speed of nitrogen molecule
$= \sqrt{\frac{2 \times 6.1 \times 10^{-21} \times 6.02 \times 10^{23}}{0.028}} = 510\,\text{m s}^{-1}$

25 Internal energy = $E = N_A\left(\frac{3}{2}kT\right) = \frac{3}{2}RT$
Change per kelvin = $\frac{3}{2}R$

Answers to SAQs

Chapter 23

1 a Electric field strength due to first sphere,

$$E_1 = \frac{Q_1}{4\pi\varepsilon_0 r^2} = \frac{2.0\times10^{-6}}{4\times\pi\times8.85\times10^{-12}\times(0.25)^2}$$

$$= 2.87\times10^5 \approx 2.9\times10^5\,\text{V m}^{-1}\ (\text{or N C}^{-1})$$

b Centre-to-centre distance of spheres
= 20 + 20 + 10 = 50 cm
Force between two point charges

$$F = \frac{Q_1 Q_2}{4\pi\varepsilon_0 r^2} = \frac{-1.0\times10^{-6}\times2.0\times10^{-6}}{4\times\pi\times8.85\times10^{-12}\times(0.50)^2}$$

$$= -0.072\,\text{N}$$

The minus sign indicates a force of attraction.

c In part a, we calculated the electric field due to the first sphere at 25 cm from its centre (which is also the midpoint between the two spheres).
Electric field strength due to second sphere

$$E_2 = \frac{Q_2}{4\pi\varepsilon_0 r^2} = \frac{-1.0\times10^{-6}}{4\times\pi\times8.85\times10^{-12}\times(0.25)^2}$$

$$= -1.44\times10^5 \approx -1.4\times10^5\,\text{V m}^{-1}\ (\text{or N C}^{-1})$$

The minus sign indicates the field is attractive to a sample positive charge towards the second sphere, whereas the field of the first sphere is repulsive to a positive charge away from the first sphere (and towards the second sphere). So the total field strength due to both spheres = $1.4\times10^5 + 2.9\times10^5 = 4.3\times10^5$ V m^{-1} (or N C^{-1}) (towards the second, negative sphere).

2 Diameter of sphere 40 cm means distance of surface from centre = 20 cm.

Rearrange $E = \frac{Q}{4\pi\varepsilon_0 r^2}$ to find charge $Q = 4\pi\varepsilon_0 r^2 E$

$$= 4\times\pi\times8.85\times10^{-12}\times(0.20)^2\times4.0\times10^6$$

$$= 1.8\times10^{-5}\,\text{C}$$

3 a Electrical potential energy
$W = QV$ = +1 × potential difference.
So at different positions in the uniform field, electrical potential energy is
A, 0 J; B, 2 kJ;
C, 1 kJ
(C is halfway between A, 0 V and B, +2 kV);
D, 1.5 kJ (D is halfway between C, +1 kV and B, +2 kV).

b Electrical potential energy
$W = QV$ = +2 × potential difference
= +2 × answers in **a**.
A, 0 J; B, 4 kJ; C, 2 kJ; D, 3 kJ

4 Rearrange $V = \frac{Q_2}{4\pi\varepsilon_0 r^2}$ to give charge

$$Q = 4\pi\varepsilon_0 r^2 V$$

$$= 4\times\pi\times8.85\times10^{-12}\times(0.10)^2\times100\,000$$

$$= 1.1\times10^{-6}\,\text{C}$$

Treating charge on the sphere as concentrated at a point in the centre of the sphere, potential

$$V = \frac{Q_2}{4\pi\varepsilon_0 r^2} = \frac{-1.1\times10^{-6}}{4\times\pi\times8.85\times10^{-12}\times(0.20)^2} = 25\,\text{kV}$$

5 a Work done in uniform field = QV so
E → H, 5 kJ; E → F, 2.5 kJ; F → G, 0 J; H → E, −5 kJ

b i E → H, −5 kJ; E → F, −2.5 kJ; F → G, 0 J; H → E, +5 kJ

ii E → H, 10 kJ; E → F, 5 kJ; F → G, 0 J; H → E, −10 kJ

6 Electrostatic force $F_E = \frac{Q_1 Q_2}{4\pi\varepsilon_0 r^2}$

$$= \frac{(+1.6\times10_{-19})^2}{4\times\pi\times8.85\times10^{-12}\times(1.0\times10^{-1}5)^2} = 230\,\text{N}$$

Gravitational force $F_G = -\frac{GMm}{r^2}$

$$= \frac{6.67\times10^{-11}\times(1.67\times10^{-27})^2}{(1.0\times10^{-15})^2} = 1.9\times10^{-34}\,\text{N}$$

This answer tells us that the gravitational attraction is nowhere near enough to balance the electric repulsion. Some other force must hold the protons together. (In fact, it is the strong nuclear force.)

Answers to SAQs

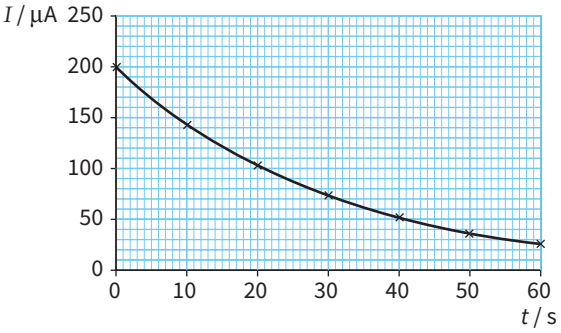
Chapter 24

1 Rearrange capacitance $C = \frac{Q}{V}$ to give charge stored
$Q = CV = 220 \times 10^{-6} \times 15$
$= 3300\ \mu C$ (or 3.3×10^{-3} C)

2 Capacitance $C = \frac{Q}{V} = \frac{1.0 \times 10^{-3}}{500}$
$= 2.0 \times 10^{-6}$ F (or 2.0 μF, 2.0×10^{6} pF)

3 Current $I = \frac{Q}{t}$ so charge $Q = It$
Substitute in $C = \frac{Q}{V} = \frac{It}{V}$
Rearrange so $I = \frac{CV}{t} = \frac{50 \times 10^{-6} \times 10}{0.01}$
$= 0.050$ A (50 mA)

4



Charge = area under graph ≈ 5.1 mC
Capacitance ≈ 6.0×10^{-4} F (600 μF)

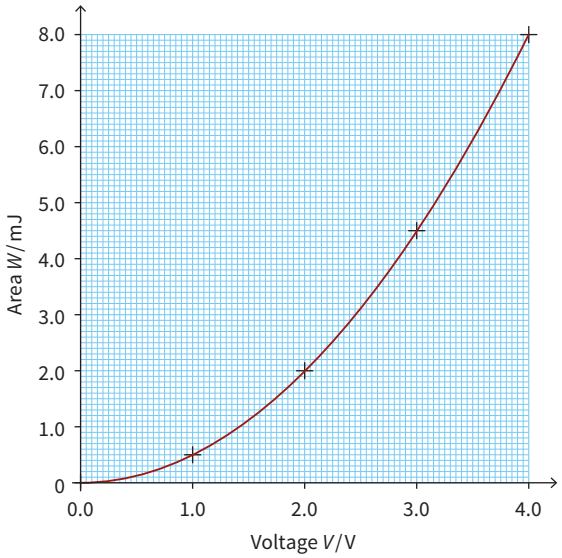
5 Gradient $= \frac{V}{Q} = \frac{1}{C}$

6 **a** Capacitance $= \frac{Q}{V}$
$= \frac{1}{\text{gradient}} = \frac{4 \times 10^{-3}}{4} = 1.0 \times 10^{-3}$ F (1 mF)

b

Q / mC	*V* / V	Area of strip Δ*W* / mJ	Sum of areas *W* / mJ
1.0	1.0	0.5	0.5
2.0	2.0	1.5	2.0
3.0	3.0	2.5	4.5
4.0	4.0	3.5	8.0

c



The graph is a parabola.

7 **a** Energy stored, $E = \frac{1}{2}CV^2$
$E = \frac{1}{2} \times 5 \times 10^{-3} \times (5.0)^2$
$= 6.25 \times 10^{-2}$ J ≈ 6.3×10^{-2} J

b $E = \frac{1}{2} \times 5 \times 10^{-9} \times (5.0)^2 = 6.3 \times 10^{-8}$ J

c $E = \frac{1}{2} \times 0.2 \times 10^{-3} \times (230)^2 = 5.29$ J ≈ 5.3 J

8 Charge is the same for both capacitors,
$Q = CV = 2.0 \times 10^{-2}$ C
Energy stored, $E = \frac{1}{2}CV^2$
For 100 μF capacitor,
$E = \frac{1}{2} \times 100 \times 10^{-6} \times (200)^2 = 2.0$ J
For 200 μF capacitor,
$E = \frac{1}{2} \times 200 \times 10^{-6} \times (100)^2 = 1.0$ J

9 **a** Energy stored
$E = \frac{1}{2}CV^2 = \frac{1}{2} \times 0.01 \times (12)^2 = 0.72$ J

b Rearrange power $P = \frac{\text{work done}}{\text{time taken}}$ to give
time $= \frac{\text{work done}}{\text{power}} = \frac{0.72}{36} = 0.02$ s

10 **a** Charge $Q = CV = 0.20 \times 9.0 = 1.8$ C
Energy $E = \frac{1}{2}CV^2 = \frac{1}{2} \times 0.2 \times (9.0)^2 = 8.1$ J

b Power $= \frac{\text{energy}}{\text{time taken}} = \frac{8.1}{0.01} = 810$ W

c Current $= \frac{\text{charge}}{\text{time}} = \frac{1.8}{0.01} = 180$ A

d Rearrange power $P = I^2R$ to give resistance
$R = \frac{P}{I^2} = \frac{810}{(180)^2} = 0.025\ \Omega$

Cambridge International AS and A Level Physics © Cambridge University Press 2014

11 **a** $C_{total} = C_1 + C_2 = 100 + 100 = 200\ \mu F$

b Charge stored, $Q = CV = 200 \times 10^{-6} \times 20$
$= 4.0 \times 10^{-3}\ C\ (4000\ \mu C)$

12 Two 20 μF and one 10 μF connected in parallel; or five 10 μF connected in parallel

13 Total capacitance of capacitors in series is given by
$\frac{1}{C_{total}} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} = \frac{1}{200} + \frac{1}{300} + \frac{1}{600} = \frac{6}{600} = \frac{1}{100}$
so $C_{total} = 100\ \mu F$

14 **a** $\frac{1}{C_{total}} = \frac{1}{C} + \frac{1}{C} = \frac{2}{C} = \frac{1}{0.5C}$
so $C_{total} = 0.5C$

b $\frac{1}{C_{total}} = \frac{n}{C}$ so $C_{total} = \frac{C}{n}$

c $C_{total} = 2C$

d $C_{total} = nC$

15 **a** Total resistance is given by $R_{total} = R_1 + R_2$
meaning $\frac{1}{G_{total}} = \frac{1}{G_1} + \frac{1}{G_2}$

b Total resistance is given by $\frac{1}{R_{total}} = \frac{1}{R_1} + \frac{1}{R_2}$
so $G_{total} = G_1 + G_2$

16 **a** $\frac{1}{C_{total}} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} = \frac{1}{100} + \frac{1}{100} + \frac{1}{100} = \frac{3}{100}$
so $C_{total} = \frac{100}{3} = 33.3\ \mu F$

b In parallel,
$C_{total} = C_1 + C_2 + C_3 = 100 + 100 + 100 = 300\ \mu F$

c Capacitance of the two in parallel = 200 μF, therefore
$\frac{1}{C_{total}} = \frac{1}{100} + \frac{1}{200} = \frac{3}{200}$
so $C_{total} = \frac{200}{3} = 66.7\ \mu F$

d Capacitance of the two in series is given by
$\frac{1}{C_{series}} = \frac{1}{100} + \frac{1}{100} = \frac{2}{100} = \frac{1}{50}$
so $C_{series} = 50\ \mu F$
Therefore $C_{total} = C_{series} + 100 = 50 + 100 = 150\ \mu F$

17 **a** Four in parallel

b Four in series

c Two in series with two in parallel

18 Maximum: in parallel, 900 pF
Minimum: in series, where $\frac{1}{C_{total}} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}$
$= \frac{1}{100} + \frac{1}{200} + \frac{1}{600} = \frac{10}{600} = \frac{1}{60}$ so $C_{total} = 60\ pF$.

19 Total capacitance of the two in parallel is
$C_{parallel} = 10 + 10 = 20\ \mu F$
So total network capacitance is given by
$\frac{1}{C_{total}} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} = \frac{1}{10} + \frac{1}{20} + \frac{1}{10} = \frac{5}{20} = \frac{1}{4}$
so $C_{total} = 4.0\ \mu F$

20 **a** $\frac{1}{C_{total}} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3} = \frac{1}{120} + \frac{1}{120} + \frac{1}{120} = \frac{3}{120} = \frac{1}{40}$
so $C_{total} = 40\ \mu F$

b Charge stored
$Q = CV = 40 \times 10^{-6} \times 10\,000 = 0.40\ C$

c Energy stored, $E = \frac{1}{2}CV^2$
$= \frac{1}{2} \times 40 \times 10^{-6} \times (10\,000)^2 = 2.0 \times 10^3\ J$

21 **a** Total capacitance in parallel = 20 + 5.0 = 25 μF

b Charge stored across first capacitor when connected to the power supply, $Q = CV_1$
$= 20 \times 10^{-6} \times 200 = 4.0 \times 10^{-3}\ C\ (4000\ \mu C)$

c Potential difference across the combination is given by the charge stored (which comes from the first capacitor) and the total capacitance of the combination.
So $V_{combination} = \frac{Q}{C} = \frac{4.0 \times 10^{-3}}{25 \times 10^{-6}} = 160\ V$

d Energy stored by first capacitor when charged $= \frac{1}{2}C(V_1)^2 = \frac{1}{2} \times 20 \times 10^{-6} \times (200)^2 = 0.4\ J$
Energy stored by combination when connected $= \frac{1}{2}C(V_{combination})^2$
$= \frac{1}{2} \times 25 \times 10^{-6} \times (160)^2 = 0.32\ J$
So energy dissipated when the capacitors are connected = 0.4 – 0.32 = 0.08 J (80 mJ)

22 Capacitance of the Earth (assuming Earth is a perfect sphere)
$= 4\pi\varepsilon_0 r = 4 \times \pi \times 8.85 \times 10^{-12} \times 6.4 \times 10^6$
$= 0.000\,71\ F \approx 700\ \mu F$

Answers to SAQs

Chapter 25

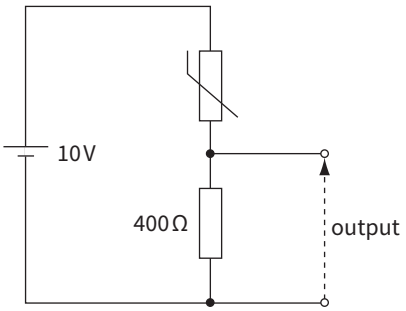
1 $\frac{3}{103} \times 10 = 0.29\,\text{V}$
There is 9.7 V across the LDR.

2 The easiest way is to swap the 3 kΩ resistor and the LDR resistor. Alternatively you can take the output as the voltage across the 3 kΩ resistor. You can also use an inverting amplifier, which is covered later in the chapter.

3 If the thermistor is in series with a fixed resistor and a battery, a changing temperature will cause a changing voltage across the thermistor.

4 Both are made from semiconductor material. Both have a decreasing resistance, for an LDR when the light intensity increases and for a thermistor when the temperature increases. Both have a non-linear change in resistance with light intensity or temperature.

5



The output voltage is shown across the 400 Ω resistor. When the temperature rises, the resistance of the thermistor decreases and so the p.d. across the thermistor decreases and the p.d. across the 400 Ω resistor increases. You can instead put the output voltage across the thermistor; then, when temperature rises, the output voltage falls.

6 **a** Using $V_{out} = \frac{R_2}{(R_1 + R_2)} \times V_{in}$ gives $5.1 = \frac{R}{120} \times 10$
So $R = 124.9\,\Omega$
Remember that Δl is proportional to ΔR.
A change of 0.10 cm gives a change in resistance of 2.4 Ω, and $\Delta R = 4.9\,\Omega$, so
$\Delta l = \frac{4.9}{2.4} \times 0.1 = 0.204 = 0.20\,\text{cm}$.

b Strain $= \frac{\Delta l}{l} = \frac{0.204}{10} = 0.0204 \approx 2.0\%$

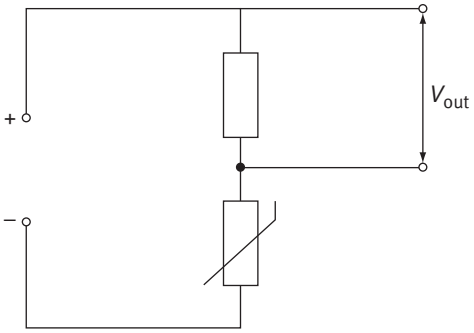
c Stretch of 0.05 cm produces change in R of $2.4 \times \frac{0.05}{0.1} = 1.2\,\Omega$
So R becomes 121.2 Ω
Use $V_{out} = \frac{R_2}{(R_1 + R_2)} \times V_{in} = \frac{121.2}{241.2} \times 10.0$
which gives $V = 5.0249 \approx 5.0\,\text{V}$.

7 When thermistor is at 20 °C,
$V_{out} = \frac{R_2}{R_1 + R_2} \times V_{in} = \frac{20\,000}{1000 + 20\,000} \times 10 = 9.5\,\text{V}$
When thermistor is at 60 °C,
$V_{out} = \frac{R_2}{R_1 + R_2} \times V_{in} = \frac{100}{1000 + 100} \times 10 = 0.91\,\text{V}$

8 The resistance of the LDR decreases so the output voltage decreases.

9 In full sunlight,
$V_{out} = \frac{R_2}{R_1 + R_2} \times V_{in} = \frac{300}{300 + 300} \times 12 = 6.0\,\text{V}$
In darkness,
$V_{out} = \frac{R_2}{R_1 + R_2} \times V_{in} = \frac{1000\,000}{300 + 1000\,000} \times 10 = 12\,\text{V}$
(just under) in darkness.

10



11 $V_{out} = G_0(V^+ - V^-)$
Saturation is reached when $V_{out} = 9\,\text{V}$,
so $V_{out} < 9\,\text{V}$ meaning $G_0(V^+ - V^-) < 9\,\text{V}$.
So difference between V^+ and V^-
$(V^+ - V^-) < \frac{9}{G_0} \approx \frac{9}{10^5} = 9 \times 10^{-5}\,\text{V}$

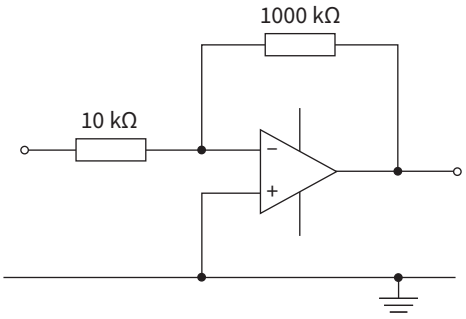
12 The op-amp switches at the same temperature. The output voltages become +15 V or −15 V.

13 Swap T and the 10 kΩ resistor; swap the connections to the V^+ and V^- terminals of the op-amp.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

14 $V^+ = \frac{10\,000}{(3000+10\,000)} \times 9\,\text{V} = +6.9\,\text{V}$, $V^- = +4.5\,\text{V}$ so, since $V^+ > V^-$, then $V_{out} = +9\,\text{V}$.

15



16 **a** Rearrange $G = \frac{V_{out}}{V_{in}}$ so

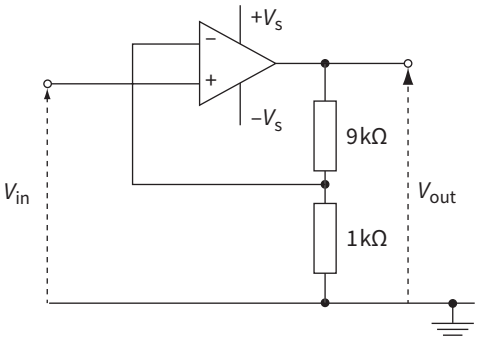
$V_{out} = GV_{in} = -20 \times (+20\,\text{mV}) = -0.40\,\text{V}$

b $V_{out} = GV_{in} = -20 \times (-400\,\text{mV}) = +8.0\,\text{V}$

c $V_{out} = GV_{in} = -20 \times (+1.0\,\text{V}) = -20\,\text{V}$, but this means the amplifier will be saturated, so the output voltage = supply voltage, i.e. −15 V.

17 $R_1 = 0$, $R_2 = \infty$, gain $= 1 + \frac{R_1}{R_2} = 1 + \frac{0}{\infty} = 1$

18



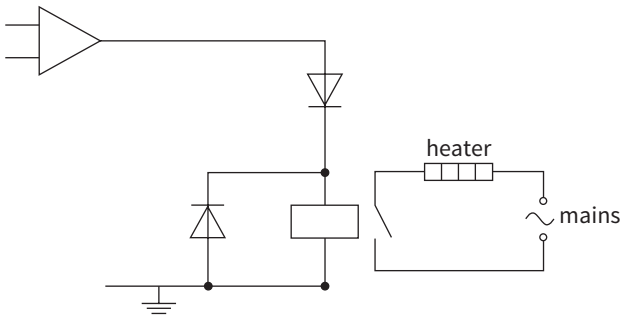
Any top resistor in the diagram must be nine times larger than the bottom resistor. You cannot choose very small resistors , e.g. 9 Ω and 1 Ω, because the op-amp cannot provide large enough currents through such small resistances. The input resistance is very high because of the very high input resistance to the op-amp and because there are no other resistors connected to the input.

19 **a** Gain $= 1 + \frac{R_1}{R_2} = 1 + \frac{50}{5} = 11$

b **i** $V_{out} = GV_{in} = 11 \times (-0.1\,\text{V}) = -1.1\,\text{V}$

ii $V_{out} = GV_{in} = 11 \times (+1.0\,\text{V}) = +11\,\text{V}$, but this means the amplifier will be saturated, so output voltage = supply voltage, i.e. +10 V

20

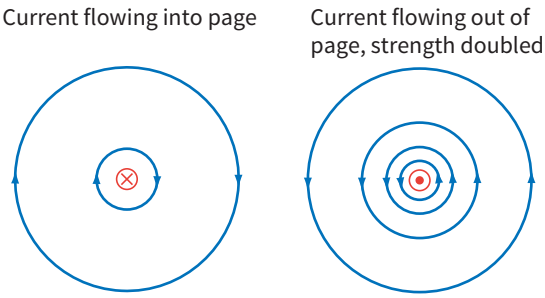


21 When the op-amp switches off the coil in the relay, there is a large induced e.m.f. because of the change in magnetic flux in the coil. The reverse-biased diode prevents any induced current reaching the op-amp. There is no induced e.m.f. in an LED and so the reverse-biased diode is not needed.

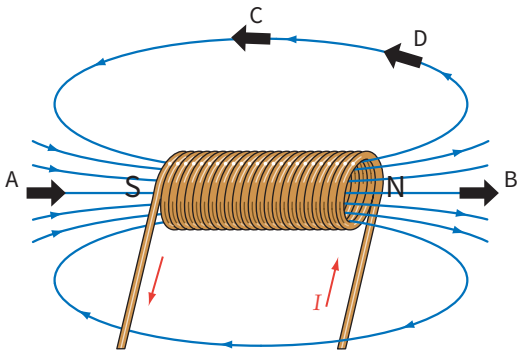
Answers to SAQs

Chapter 26

1



2



3 Pair **a** will repel; pair **b** will attract.

4 **a** No force

b Force into the plane of the paper

c Force down the page

5 Force $F = BIl = 0.06 \times 0.20 \times 2.50 = 0.030\,\text{N}$

6 Rearrange force $F = BIl$ to give flux density

$$B = \frac{F}{Il} = \frac{0.015}{(1.5 \times 0.20)} = 0.050\,\text{T}$$

7 **a** Current $I = \frac{Q}{t} = \frac{ne}{t} = 10^{18} \times \frac{1.60 \times 10^{-19}}{1} = 0.16\,\text{A}$

b Force $F = BIl = 0.005 \times 0.16 \times 0.50 = 4.0 \times 10^{-4}\,\text{N}$ (0.40 mN)

8 **a** Section of wire in field tilts down

b Tilts down

c Will try to move horizontally, into horseshoe

d No movement

9 Remember the principle of moments: sum of anticlockwise moments = sum of clockwise moments if lever is balanced. Remember also that force due to gravity on mass m is mg.

So for force F due to magnet on wire,
$F \times 0.20 = 0.02 \times 10^{-3} \times 9.81$, meaning
$F = 1.96 \times 10^{-4}\,\text{N}$
Rearrange force $F = BIl$ to give flux density

$$B = \frac{F}{Il} = \frac{1.96 \times 10^{-4}}{0.50 \times 0.050} = 7.8 \times 10^{-3}\,\text{T}$$

10 Force, $F = BIl$
$= 0.005 \times 2.4 \times 0.50 = 6.0 \times 10^{-3}\,\text{N}$ (6.0 mN)

11 **a** Force exerted by magnetic field is at a maximum when the section of wire in the field lies at right angles to the field. Remembering there are 200 turns,
Force $F = BIl = 0.05 \times 1.0 \times 200 \times 0.20 = 2.0\,\text{N}$

b Pivoted along a line parallel to one edge with the magnetic field in the same plane as the coil.

c Greater torque could be provided by: increasing current, increasing number of turns in coil, increasing length of side in field, pivoting by centre of coil and having magnets either side, having magnets all round the circle through which the coil turns, increasing field strength.

12 **a** Force $F = BIl \sin\theta = 0.25 \times 3.0 \times 0.50 \times \sin 90°$
$= 0.375\,\text{N} \approx 0.38\,\text{N}$

b Force $F = BIl \sin\theta = 0.25 \times 3.0 \times 0.50 \times \sin 45°$
$= 0.265\,\text{N} \approx 0.27\,\text{N}$

c Current is parallel to magnetic field so force $F = 0\,\text{N}$.
Both **a** and **b** are into the plane of the paper.

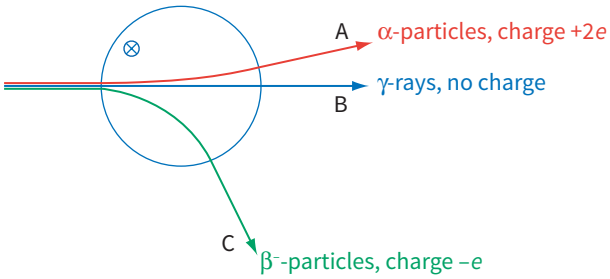
13 They attract. If you consider each flat coil as a small electromagnet, then unlike poles are facing one another. If you think of the currents in the wires, these are parallel (rather than anti-parallel) and the coils attract. When the current is reversed, it is reversed in both coils and they still attract.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Answers to SAQs

Chapter 27

1

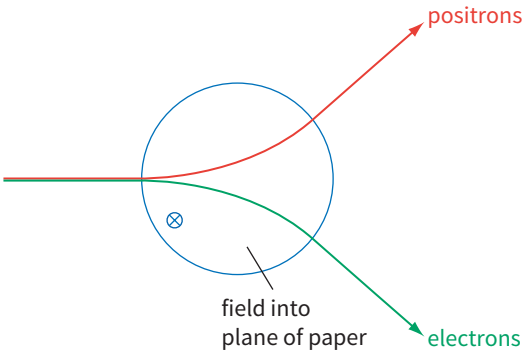


Track A is pushed upwards, meaning it is the equivalent of a positive current flowing to the right: A must be α-particles.
Track B is not deflected by the magnetic field: B must be γ-rays.
Track C is pushed downwards, meaning it is the equivalent of a negative current flowing to the right: C must be β^--particles.

2 **a** At right angles to the magnetic field, force
$F = BQv = 0.50 \times 1.6 \times 10^{-19} \times 1.0 \times 10^{6}$
$= 8.0 \times 10^{-14}$ N

b At an angle of 45° to the magnetic field, force
$F = BQv \sin\theta$
$= 0.50 \times 1.6 \times 10^{-19} \times 1.0 \times 10^{6} \times \sin 45°$
$= 5.66 \times 10^{-14}$ N $\approx 5.7 \times 10^{-14}$ N

3 Since the particles have opposite charges, when moving through a magnetic field at right angles to the direction of travel the positrons experience a force in the opposite direction to the force on the electrons. This will separate the particles into two beams.



4 Out of the plane of the photograph.

5 All have same mass, charge and speed. (There is more about this when you study how J.J. Thomson discovered the electron, later in this chapter of the coursebook.)

6 **a** Circular path will have smaller radius.

b Electrons will circle in the opposite direction.

c Circular path will have smaller radius.

d Electrons will spiral around field lines because they will have a constant component of velocity in the direction of the field lines.

7 **a** The magnetic force is upwards, towards the positive plate, and the electric force is downwards, towards the negative plate.

b Speed of ion, $v = \frac{E}{B} = \frac{1.5 \times 10^{3}}{0.30} = 5.0 \times 10^{3}$ m s^{-1}

c Magnetic force > electric force; the ion travels in an upward curved path (towards the positive plate) and hence misses the slit S.

8 **a** Hall voltage $V_H = \frac{BI}{nte}$

$= \frac{0.10 \times 0.020}{(1.5 \times 10^{23} \times 0.05 \times 10^{-3} \times 1.6 \times 10^{-19})}$

$= 1.67 \times 10^{-3}$ V

≈ 1.7 mV

b The current in a Hall probe must be maintained at a constant value because the force that acts on the electrons to generate a voltage is proportional to both the magnetic flux density and the current: $F = BIl$. The current must be kept constant so that the size of the force, and therefore the voltage measured, varies only as the magnetic flux density changes.

9 Remember that the Hall voltage $V_H = \frac{BI}{nte}$. Using a single sample of a semiconducting material of measured thickness, and using a constant magnetic field of known flux density, the Hall voltage could be measured for different values of current. By plotting a graph of Hall voltage against current, the graph should form a straight line with gradient = $\frac{BI}{nte}$, from which the number density of the semiconductor can be calculated.

10 Charge-to-mass ratio = $\frac{e}{m}$, so mass of electron, $m = \frac{1.60 \times 10^{-19}}{1.76 \times 10^{11}} = 9.1 \times 10^{-31}$ kg

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Answers to SAQs

Chapter 28

1 The rotating magnet changes the magnetic flux linking the coil, hence an e.m.f. is induced. The induced current in the coil will light the lamp connected to the generator.

2 Current flows from A to B and from C to D. Y is positive so that current flows from Y to X in the external circuit to make it flow from A to B inside the coil.

3 Left wingtip will be positive. It is negative in the southern hemisphere because the field direction is reversed.

4 Magnetic flux = BA. A stronger magnet means greater flux linking the coil and hence a greater induced e.m.f. Faster movement means more flux cut or linked per second and more current generated, or larger e.m.f.

5 The wire is moved parallel to the field, hence only small components are cut due to slight curvature at edges of field.

6 Frequency is determined by speed of rotation (so for this to be kept constant, the generator must be geared). E.m.f. is affected by magnet strength, number of turns in coil, size of coil. The e.m.f. is normally affected by the speed of rotation, but in this case this has to be fixed as the frequency is fixed.

7 Magnetic flux $\Phi = BA = 0.15 \times 0.01 \times 0.015$
$= 2.25 \times 10^{-5}\,\text{Wb} \approx 2.3 \times 10^{-5}\,\text{Wb}$

8 Magnetic flux linkage $= N\Phi = NBA = \frac{NB\pi d^2}{4}$
$= \frac{200 \times 2.0 \times 10^{-5} \times \pi \times (0.05)^2}{4} = 7.9 \times 10^{-6}\,\text{Wb}$

9 Magnetic flux linkage $= N\Phi = NBA$
$= 120 \times 1.2 \times 0.05 \times 0.075 = 0.54\,\text{Wb}$

10 Rate of change in area $= lv$
Rate of change of flux $= B \times (lv) = Blv$

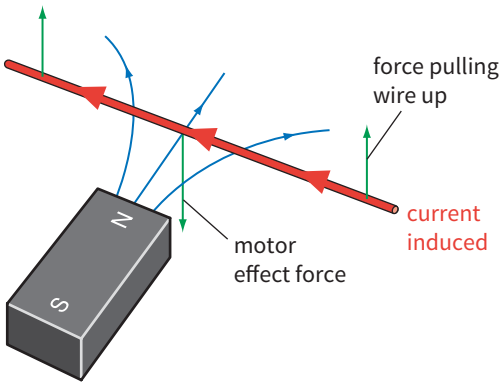
11 Change in magnetic flux $= B \times$ change in area
$= 1.5 \times (0.10 \times 0.02) = 3.0 \times 10^{-3}\,\text{Wb} = \Delta(N\Phi)$
Use Faraday's law to determine e.m.f.:
$E = \frac{\Delta(N\Phi)}{\Delta t} = \frac{3.0 \times 10^{-3}}{0.50} = 6.0\,\text{mV}$

12 Faraday's law gives e.m.f. $E = \frac{\Delta(N\Phi)}{\Delta t} = \frac{BAN}{\Delta t}$
Rearrange to give $B = \frac{E\Delta t}{AN}$
$= \frac{0.40 \times 0.20}{1.2 \times 10^{-4} \times 2000} = 0.33\,\text{T}$

13 **a** Stop pushing implies no change in flux linkage, so no current is generated. Therefore, no magnetic poles are formed and no work is done; there is no movement.

b Pull away implies that flux is decreased in the flux linkage, but the end of solenoid near the magnet becomes a south pole, so the poles attract each other, and work has to be done to pull magnet and coil apart.

14



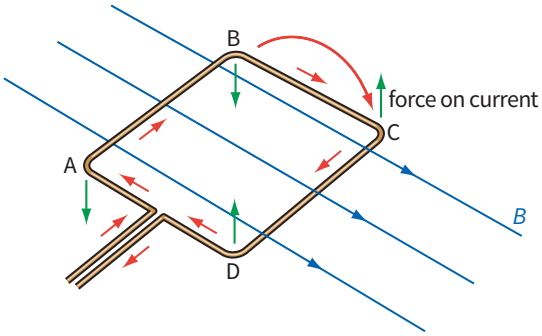
15 **a** There is a sudden increase in the flux linkage for the coil, so there is an induced e.m.f.

b There is no change in the flux linking the coil. The motion is parallel to the field.

c Magnet leaves coil, there is a decrease in the flux linking the coil and hence the e.m.f. is in the reverse (negative) direction. The induced current is in the opposite direction (Lenz's law). Peak e.m.f. is greater because the magnet is moving faster (acceleration due to gravity); the rate of change of flux linkage is greater.

16 You have to do work against the motor effect force from an induced current when the lights are on. When the lights are off, there is no induced current (although there is still an induced e.m.f.) and so no motor effect force.

17



18 Alternating current. Usually, a bar magnet rotates inside a fixed coil. As the north pole passes one side of the coil, the current flows one way. Then the south pole passes, and the current reverses.

19 B greater means greater flux, hence $\frac{\Delta(N\Phi)}{\Delta t}$ is greater; therefore $E \propto B$.
A greater means greater flux, hence $\frac{\Delta(N\Phi)}{\Delta t}$ is greater; therefore $E \propto A$.
N greater means greater flux linkage, hence $\frac{\Delta(N\Phi)}{\Delta t}$ is greater; therefore $E \propto N$.
f greater means rate of change of flux linkage is greater, hence $\frac{\Delta(N\Phi)}{\Delta t}$ is greater; therefore $E \propto f$.

20 For a d.c. supply, the flux linkage is constant – there is no change in the flux, and hence no induced e.m.f.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Answers to SAQs

Chapter 29

1 **a** 2 A, positive

b 15 ms

c 20 ms

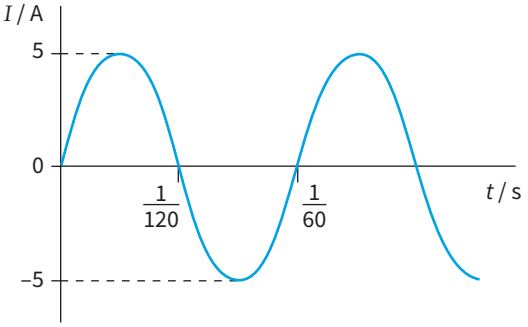
d Frequency $= \frac{1}{T} = \frac{1}{0.020} = 50$ Hz

2 **a** I_0 = peak value of current = 2 A
ω = angular frequency
$= 2\pi f = 2\pi \times 50 = 100\pi$ rad s^{-1}

b $I = 2\sin(100\pi t)$

3 **a** $I_0 = 5$ A
$\omega = 120\pi$ rad s^{-1}
Frequency $f = \frac{\omega}{2\pi} = 60$ Hz
Period $T = \frac{1}{f} = \frac{1}{60} = 17$ ms

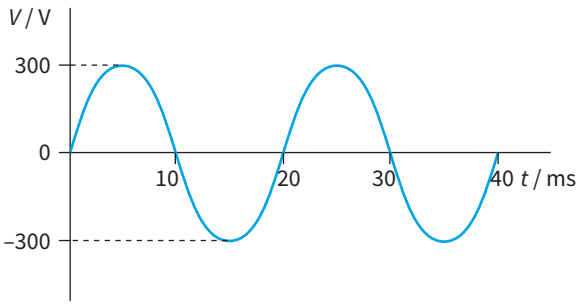
b



4 **a** V_0 = peak value of voltage = 300 V
$\omega = 100\pi$ rad s^{-1}
frequency $f = \frac{\omega}{2\pi} = 50$ Hz

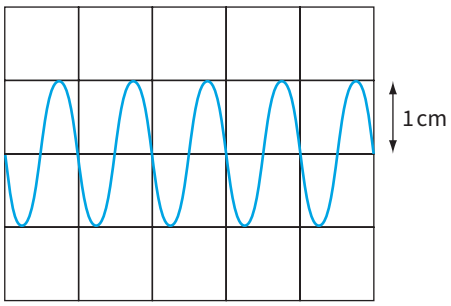
b At $t = 0.002$ s, $V = 300\sin(100\pi \times 0.002)$
$= 300 \times 0.588 = 176$ V

c



5 Amplitude $= 2 \times 5 = 10$ V
Period $T = 4 \times 10 = 40$ ms
Frequency $= \frac{1}{T} = \frac{1}{0.040} = 25$ Hz

6



The exact starting point on the graph may be different, but the trace has an amplitude of 1 cm and one complete wave is 1 cm horizontally.

7 $I_{rms} = \frac{I_0}{\sqrt{2}} = \frac{2.5}{\sqrt{2}} = 1.8$ A

8 Rearrange $V_{rms} = \frac{V_0}{\sqrt{2}}$ to give
$V_0 = V_{rms} \times \sqrt{2} = 230 \times \sqrt{2} = 325$ V

9 Power dissipated, $P = (I_{rms})^2 R = \left(\frac{I_0}{\sqrt{2}}\right)^2 R = \frac{I_0^2 R}{2}$
$= (3.0)^2 \times \frac{100}{2} = 450$ W

10 **a** $V_{rms} = \frac{V_0}{\sqrt{2}} = \frac{325}{\sqrt{2}} = 230$ V

b $I_{rms} = \frac{V_{rms}}{R} = \frac{230}{1000} = 0.23$ A

c Power dissipated, $P = (I_{rms})^2 R$
$= (0.23)^2 \times 1000 = 53$ W

d Peak power dissipated, $P = \frac{V_0^2}{R} = \frac{325^2}{1000} = 105.6$
≈ 110 W

11 Alternative answers are given here because there are two ways of interpreting the diagrams and the numbers of turns.

a Step-up: 15 : 5 = 3, **or** 16 : 6 = 2.7

b Step-down: 4 : 8 = 0.5, **or** 5 : 9 = 0.56

c Step-up: 30 V **or** 27 V; step-down: 5.0 V **or** 5.6 V

12 Step-up ratio required $= \frac{400\text{ kV}}{25\text{ kV}} = 16$
So number of turns $= 16 \times 2000 = 32\,000$ turns

13 **a** $V_s = V_p \times$ turns ratio $= 20 \times 10 = 200$ V
$I_s = \frac{I_p}{\text{turns ratio}} = \frac{50}{10} = 5$ mA

b Power $P = VI$
If no power loss occurred,
$P = 200 \times 0.005 = 1.0$ W
Actual power $= 180 \times 0.0045 = 0.81$ W
So percentage of power that is wasted
$= \frac{(1.0 - 0.81)}{1.0} = 19\%$

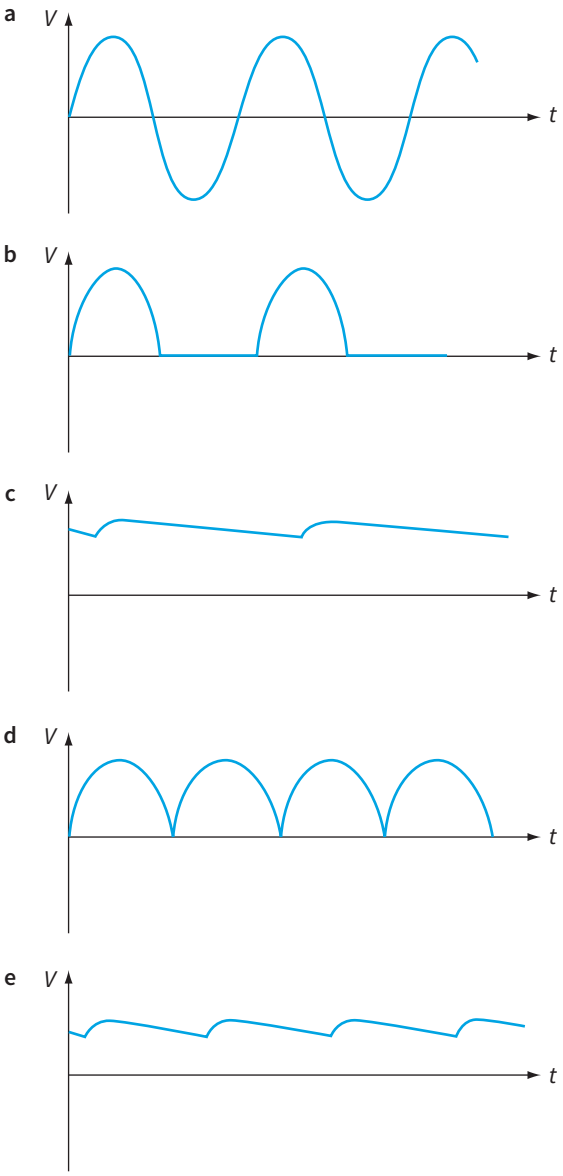
Cambridge International AS and A Level Physics © Cambridge University Press 2014

14

Transformer	N_p	N_s	V_p / V	V_s / V	I_p / A	I_s / A	P / W
A	100	500	230	1150	1.0	0.2	230
B	500	100	230	46	1.0	5.0	230
C	100	2000	12	240	0.2	0.01	2.4

15 Diode 3 is pointing the wrong way, so the current flows through diode 4, through *R* and then through diode 1.

16



17 The voltage will be half-wave rectified. Current can flow through diode 2 when terminal A is positive. When terminal B is positive, current cannot flow because there is no complete path to terminal A.

18 **a** Less pronounced ripple

b More pronounced ripple

Answers to SAQs

Chapter 30

1 Energy $E = hf = 6.63\times10^{-34}\times1\times10^{26}$
$= 6.63\times10^{-8}\,\text{J} \approx 6.6\times10^{-8}\,\text{J}$

2 Energy $E = \frac{hc}{\lambda}$
For red light, $E = \frac{6.63\times10^{-34}\times3.0\times10^{8}}{700\times10^{-9}}$
$= 2.8\times10^{-19}\,\text{J}$
For violet light, $E = \frac{6.63\times10^{-34}\times3.0\times10^{8}}{400\times10^{-9}}$
$= 5.0\times10^{-19}\,\text{J}$

3 Rearrange energy $E = \frac{hc}{\lambda}$ to give wavelength $\lambda = \frac{hc}{E}$

a $\lambda = \frac{6.63\times10^{-34}\times3.0\times10^{8}}{10^{-12}} \approx 2.0\times10^{-13}\,\text{m}$, γ-ray

b $\lambda = \frac{6.63\times10^{-34}\times3.0\times10^{8}}{2.0\times10^{-10}} \approx 2.0\times10^{-10}\,\text{m}$, X-ray

c $\lambda = \frac{6.63\times10^{-34}\times3.0\times10^{8}}{10^{-18}} \approx 2.0\times10^{-7}\,\text{m}$, ultraviolet

d $\lambda = \frac{6.63\times10^{-34}\times3.0\times10^{8}}{10^{-20}} \approx 2.0\times10^{-5}\,\text{m}$, infrared

e $\lambda = \frac{6.63\times10^{-34}\times3.0\times10^{8}}{10^{-25}} \approx 2.0\,\text{m}$, radio wave

4 Energy of one photon, $E = \frac{hc}{\lambda}$
$= \frac{6.63\times10^{-34}\times3.0\times10^{8}}{6.48\times10^{-7}} = 3.07\times10^{-19}\,\text{J}$
$\text{Number of photons per second} = \frac{\text{power of laser}}{\text{energy per photon}}$
$= \frac{0.0010}{3.07\times10^{-19}} = 3.26\times10^{15}\,\text{s}^{-1} \approx 3.3\times10^{15}\,\text{s}^{-1}$

5 Energy change, $W = 1.2\,\text{eV} = QV$
$= 1.6\times10^{-19}\times1.2 = 1.92\times10^{-19}\,\text{J} \approx 1.9\times10^{-19}\,\text{J}$

6 Energy in joules, $E = hf$
Therefore energy in electronvolts, $E = \frac{hf}{e}$
where e is the electric charge of the electron.
$E = \frac{6.63\times10^{-34}\times3.0\times10^{8}}{1.6\times10^{-19}} = 12\,400\,\text{eV} \approx 12\,\text{keV}$

7 Rearrange energy $E = \frac{hc}{\lambda}$ to give wavelength
$\lambda = \frac{hc}{E} = \frac{6.63\times10^{-34}\times3.0\times10^{8}}{10\times1.6\times10^{-19}}$
$\approx 1.24\times10^{-7}\,\text{m}$, ultraviolet

8 **a** Note that these calculations assume the proton is accelerated from rest.
Energy gained
= work done on proton by potential difference
$= QV = 1.6\times10^{-19}\times1500 = 2.4\times10^{-16}\,\text{J}$

b Kinetic energy $E_k = \frac{1}{2}mv^2$ so speed $v = \sqrt{\frac{2E_k}{m}}$
$= \sqrt{\frac{2\times2.4\times10^{-16}}{1.7\times10^{-27}}} = 5.31\times10^{5}\,\text{m s}^{-1} \approx 5.3\times10^{5}\,\text{m s}^{-1}$

9 Rearrange $eV = \frac{hc}{\lambda}$ to give Planck's constant, $h = \frac{eV\lambda}{c}$
For infrared LED,
$h_i = \frac{1.6\times10^{-19}\times1.35\times910\times10^{-9}}{3.0\times10^{8}} = 6.55\times10^{-34}\,\text{J s}$
For red LED,
$h_r = \frac{1.6\times10^{-19}\times1.70\times670\times10^{-9}}{3.0\times10^{8}} = 6.07\times10^{-34}\,\text{J s}$
For amber LED,
$h_a = \frac{1.6\times10^{-19}\times2.00\times610\times10^{-9}}{3.0\times10^{8}} = 6.51\times10^{-34}\,\text{J s}$
For green LED,
$h_g = \frac{1.6\times10^{-19}\times2.30\times560\times10^{-9}}{3.0\times10^{8}} = 6.87\times10^{-34}\,\text{J s}$
Average value
$= \frac{1}{4}(6.55\times10^{-34} + 6.07\times10^{-34} + 6.51\times10^{-34} + 6.87\times10^{-34})$
$\approx 6.5\times10^{-34}\,\text{J s}$

10 **a** Energies that exceed the work function can cause the release of an electron from the metal: 2.0 eV, 3.0 eV.

b Rearrange $hf = \varphi + \text{k.e.}_{\text{max}}$ to give
$\text{k.e.}_{\text{max}} = hf - \varphi$
= photon energy – work function
For 2.0 eV photon, $\text{k.e.}_{\text{max}} = 2.0\,\text{eV} - 1.8\,\text{eV}$
$= 0.2\,\text{eV}$
in joules, $0.2\,\text{eV} = 0.2\times1.6\times10^{-19}\,\text{J}$
$= 3.2\times10^{-20}\,\text{J}$
For 3.0 eV photon, $\text{k.e.}_{\text{max}} = 3.0\,\text{eV} - 1.8\,\text{eV}$
$= 1.2\,\text{eV}$
in joules, $1.2\,\text{eV} = 1.2\times1.6\times10^{-19}\,\text{J}$
$= 1.9\times10^{-19}\,\text{J}$

11 **a** Gold

b Caesium

c For zinc, threshold frequency
$f_0 = \frac{\varphi}{h} = \frac{6.9\times10^{-19}}{6.63\times10^{-34}} = 1.04\times10^{15} \approx 1.0\times10^{15}\,\text{Hz}$

d For potassium, threshold frequency
$f_0 = \frac{\varphi}{h} = \frac{3.2\times10^{-19}}{6.63\times10^{-34}} = 4.83\times10^{14}\,\text{Hz}$
So wavelength $\lambda = \frac{c}{f_0} = \frac{3.0\times10^{8}}{4.83\times10^{14}} \approx 620\,\text{nm}$

Cambridge International AS and A Level Physics © Cambridge University Press 2014

12 **a** Energy $E = \frac{hc}{\lambda} = \frac{6.63\times10^{-34}\times3.0\times10^{8}}{2.4\times10^{-7}} = 8.3\times10^{-19}$ J

b $\text{k.e.}_{max} = E - \varphi = 8.3\times10^{-19} - 2.8\times10^{-19} = 5.5\times10^{-19}$ J

c Rearrange $\text{k.e.}_{max} = \frac{1}{2}m(v_{max})^2$ to give

$v_{max} = \sqrt{\frac{2E_k}{m}} = \sqrt{\frac{2\times5.5\times10^{-19}}{9.1\times10^{-31}}} = 1.1\times10^{6}\,\text{m s}^{-1}$

13 Rearrange $\frac{hc}{\lambda} = \varphi + \text{k.e.}_{max}$ to give

work function

$\varphi = \frac{hc}{\lambda} - \text{k.e.}_{max} = \frac{6.63\times10^{-34}\times3.0\times10^{8}}{2000\times10^{-9} - 4.0\times10^{-20}} = 5.9\times10^{-20}$ J

14 **a** Energy $\Delta E = E_1 - E_2 = (7.8 - 2.2)\times10^{-18} = 5.6\times10^{-18}$ J

Frequency $f = \frac{\Delta E}{h}$

$= \frac{5.6\times10^{-18}}{6.63\times10^{-34}} = 8.44\times10^{15} \approx 8.4\times10^{15}$ Hz

Wavelength $\lambda = \frac{c}{f}$

$= \frac{3.0\times10^{8}}{8.44\times10^{15}} = 3.6\times10^{-8}$ m (emission)

b Energy $\Delta E = E_1 - E_2$

$= (2.2 - 1.7)\times10^{-18} = 5.0\times10^{-19}$ J

Frequency $f = \frac{\Delta E}{h}$

$= \frac{5.0\times10^{-19}}{6.63\times10^{-34}} = 7.54\times10^{14} \approx 7.5\times10^{14}$ Hz

Wavelength $\lambda = \frac{c}{f}$

$= \frac{3.0\times10^{8}}{7.54\times10^{14}} = 4.0\times10^{-7}$ m (emission)

c Energy $\Delta E = E_1 - E_2$

$= (3.9 - 1.7)\times10^{-18} = 2.2\times10^{-18}$ J

Frequency $f = \frac{\Delta E}{h}$

$= \frac{2.2\times10^{-18}}{6.63\times10^{-34}} = 3.32\times10^{15} \approx 3.3\times10^{15}$ Hz

Wavelength $\lambda = \frac{c}{f}$

$= \frac{3.0\times10^{8}}{3.32\times10^{15}} = 9.0\times10^{-8}$ m (absorption)

15 9.0 eV, 11 eV, 25 eV, 34 eV and 45 eV correspond to differences between energy levels, so they can all be absorbed; 6.0 eV and 20 eV do not correspond to differences between energy levels and so cannot be absorbed.

16 **a** Energy $E = \frac{hc}{\lambda} = \frac{6.63\times10^{-34}\times3.0\times10^{8}}{83\times10^{-9}}$

$= 2.40\times10^{-18}\,\text{J} = \frac{2.4\times10^{-18}}{1.6\times10^{-19}} \approx 15$ eV

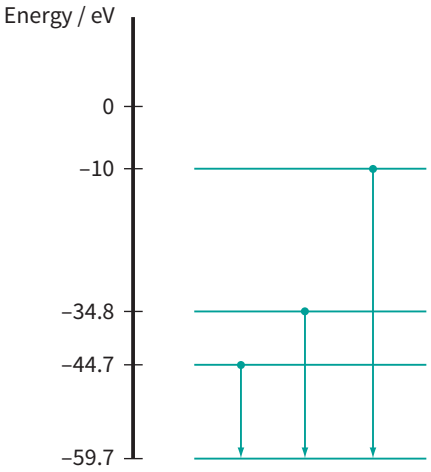
Energy $E = \frac{hc}{\lambda} = \frac{6.63\times10^{-34}\times3.0\times10^{8}}{50\times10^{-9}}$

$= 3.98\times10^{-18}\,\text{J} = \frac{3.98\times10^{-18}}{1.6\times10^{-19}} \approx 25$ eV

Energy $E = \frac{hc}{\lambda} = \frac{6.63\times10^{-34}\times3.0\times10^{8}}{25\times10^{-9}}$

$= 7.96\times10^{-18}\,\text{J} = \frac{7.96\times10^{-18}}{1.6\times10^{-19}} \approx 50$ eV

b See figure for one possible solution.



17 **a** Insulator

b Metal

c Semiconductor

18 In a semiconductor, the valence band is fully occupied by electrons but the conduction band is empty. The bands are close in energy, so some electrons gain enough energy to jump across into the conduction band. As temperature decreases, the number of electrons that can jump into the conduction band also decreases. This means conductivity decreases, and resistivity increases.

19 **a** Electrons can behave as waves so they can be diffracted by spaces between atoms.

b Each metal has a different lattice structure, so each will produce a different diffraction pattern.

20 **a** 1.0 keV

b Rearrange kinetic energy $E_k = \frac{1}{2}mv^2$ to give

speed $v = \sqrt{\frac{2E_k}{m}} = \sqrt{\frac{2\times1000\times1.6\times10^{-19}}{9.1\times10^{-31}}}$

$= 1.9\times10^{7}\,\text{m s}^{-1}$

Momentum $mv = 9.1\times10^{-31}\times1.9\times10^{7}$

$= 1.7\times10^{-23}\,\text{kg m s}^{-1}$

c De Broglie wavelength

$\lambda = \frac{h}{mv} = \frac{6.63\times10^{-34}}{1.7\times10^{-23}} = 3.9\times10^{-11}$ m

d The wavelength is much smaller than the spacing, so there will only be a small amount of diffraction.

Answers to SAQs

Chapter 31

1 **a** $^{220}_{86}\text{Rn} \rightarrow {}^{216}_{84}\text{Po} + {}^{4}_{2}\text{He}$

b $^{25}_{11}\text{Na} \rightarrow {}^{25}_{12}\text{Mg} + {}^{0}_{-1}\text{e}$

2 $^{41}_{18}\text{Ar} \rightarrow {}^{41}_{19}\text{K} + {}^{0}_{-1}\text{e}$

3 Rearrange $E = mc^2$ to give mass decrease per second, $m = \frac{E}{c^2} = \frac{4.0 \times 10^{26}}{(3.0 \times 10^8)^2} \approx 4.4 \times 10^9\ \text{kg}$

4 **a** Helium nucleus is formed from two protons and two neutrons. Energy released is given by the difference in mass between the four separate nucleons and the helium nucleus:
$\Delta m = 2m_p + 2m_n - m_{He}$
$= (2 \times 1.672\,623 + 2 \times 1.674\,929 - 6.644\,661) \times 10^{-27}$
$= 5.04 \times 10^{-29}\ \text{kg}$
Energy released, $\Delta E = \Delta m c^2$
$= 5.04 \times 10^{-29} \times (3.00 \times 10^8)^2 = 4.5 \times 10^{-12}\ \text{J}$

b $1.1 \times 10^{-12}\ \text{J}$

5 Kinetic energy of golf ball
$= \frac{1}{2}mv^2 = \frac{1}{2} \times 0.150 \times (50)^2 = 187.5\ \text{J}$
Mass equivalent of energy
$= \frac{E}{c^2} = \frac{187.5}{(3.0 \times 10^8)^2} = 2.1 \times 10^{-15}\ \text{kg}$
Percentage of rest mass $= \frac{2.1 \times 10^{-15}}{0.150} = 1.4 \times 10^{-12}\ \%$

6 **a** Mass of atom $^{56}_{26}\text{Fe} = 55.934\,937 \times 1.6605 \times 10^{-27}$
$= 9.2880 \times 10^{-26}\ \text{kg}$

b Mass of atom $^{16}_{8}\text{O} = \frac{2.656\,015 \times 10^{-26}}{1.6605 \times 10^{-27}} = 15.995\ \text{u}$

7 **a** Mass excess of helium-4 $= 4.002\,602 - 4$
$= 0.002\,602$

b Mass excess of potassium-40 $= 39.963\,998 - 40$
$= -0.036\,002$

8 **a** $^{10}_{4}\text{Be} \rightarrow {}^{10}_{5}\text{B} + {}^{0}_{-1}\text{e}$

b Mass difference
$\Delta m = 1.662\,38 \times 10^{-26} - 1.662\,19 \times 10^{-26} - 9.109\,56 \times 10^{-31}$
$= 9.890\,44 \times 10^{-31}\ \text{kg}$
Energy released
$\Delta m c^2 = 9.890\,44 \times 10^{-31} \times (2.997\,92 \times 10^8)^2$
$\approx 8.9 \times 10^{-14}\ \text{J}$
The energy is released as kinetic energy of the products.

9 It is a single nucleon and hence does not have binding energy.

10 **a** Mass defect $\Delta m = 4 \times m_p + 4 \times m_n - m_{Be}$
$\approx (8 \times 1.67 - 13.3) \times 10^{-27} = 6.00 \times 10^{-29}\ \text{kg}$

b Binding energy $E_B = \Delta m c^2$
$= 6.00 \times 10^{-29} \times (3.00 \times 10^8)^2$
$= 5.4 \times 10^{-12}\ \text{J} = \frac{5.4 \times 10^{-12}}{1.6 \times 10^{-19}}\ \text{eV} = 3.4 \times 10^7\ \text{eV}$

c Binding energy per nucleon $= \frac{E_B}{n}$
$= \frac{5.4 \times 10^{-12}}{8} = 6.8 \times 10^{-13}\ \text{J} = 4.2 \times 10^6\ \text{eV}$

11 Fission for $A < 20$ is unlikely because the products would have a smaller binding energy per nucleon. The reaction would require an input of external energy. Fusion for $A > 40$ is unlikely for the same reason.

12 Activity $A = \lambda N = 0.30 \times 500\,000$
$= 150\,000\ \text{s}^{-1}$, or $150\,000\ \text{Bq}$

13 Activity $A = 10 \times$ count rate of detector
$= 10 \times 20\ \text{min}^{-1} = 200\ \text{min}^{-1} = \frac{200}{60}\ \text{s}^{-1} = \frac{10}{3}\ \text{s}^{-1}$
Rearrange activity $A = \lambda N$ to give decay constant
$\lambda = \frac{A}{N} = \frac{10}{(3 \times 1.5 \times 10^9)} = 2.2 \times 10^{-9}\ \text{s}^{-1}$

14 Count rate might be less than activity because:

i Gamma-rays are not always detected (weakly ionising).

ii The counter is inefficient.

iii Some radiation is absorbed within the sample before reaching the detector.

iv The detector is directional, so some radiation will move away from the detector rather than towards it.

15 **a** $N = N_0 e^{-\lambda t}$

b After 10 minutes, half the nuclei will decay, so half will remain undecayed, or 4.0×10^{10}
After a further 10 minutes, half the remaining number will decay, leaving one-quarter of the original number, or 2.0×10^{10}

c After 30 minutes, three half-lives have elapsed. Number of original nuclei decaying in this time $= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8}$ of total $= 7.0 \times 10^{10}$

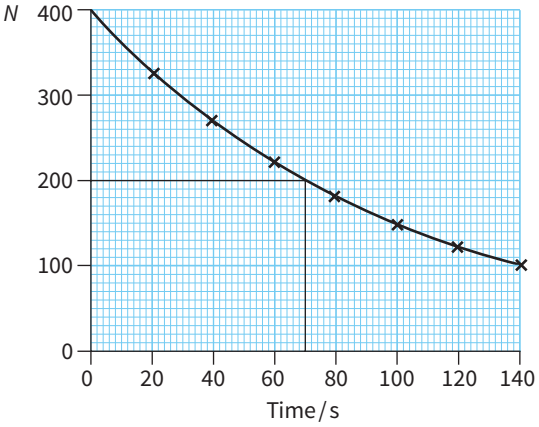
Cambridge International AS and A Level Physics © Cambridge University Press 2014

16 **a** $N = N_0 e^{-\lambda t} = 5.0 \times 10^9 \times e^{(-0.10 \times 50)}$
$= 3.37 \times 10^7 \approx 3.4 \times 10^7$

b Activity $A = \lambda N = 0.10 \times 3.37 \times 10^7 \approx 3.4 \times 10^6$ Bq

17

t / s	0	20	40	60	80	100	120	140
N	400	330	272	224	185	153	126	104



Half-life is about 70 s.

18 Half-life is 2.4 years.
Decay constant $\lambda = \frac{0.693}{2.4} = 0.29\,\text{year}^{-1}$

19 Time taken for activity to decrease to $\frac{1}{8}$ of initial value is three half-lives (sequence $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}$).
Therefore time taken $= 3t_{1/2} = 3 \times \frac{0.693}{\lambda}$
$= \frac{3 \times 0.693}{3.0 \times 10^{-4}} = 6.93 \times 10^3 \approx 6900$ s

20 **a** Decay constant $\lambda = \frac{0.693}{7.4} = 0.094\,\text{s}^{-1}$

b **i** $N = N_0 e^{-\lambda t} = 5.0 \times 10^3 \times e^{(-0.094 \times 14.8)} = 1250$

ii $N = N_0 e^{-\lambda t} = 5.0 \times 10^3 \times e^{(-0.094 \times 20.0)} = 760$ (approximately)

21 **a** We need to find an expression for the decay constant λ so that we can substitute it into the decay equation.
The quantity f is the ratio of atoms remaining to decay to the original number of atoms in the sample: $f = \frac{N}{N_0}$
When $t = t_{1/2}$, $f = \frac{1}{2} = e^{-\lambda t_{1/2}}$
Taking logarithms of both sides, $\ln\left(\frac{1}{2}\right) = -\lambda t_{1/2}$
Therefore $\lambda = -\frac{\ln\left(\frac{1}{2}\right)}{t_{1/2}}$
Then the equation $f = e^{-\lambda t}$ becomes $f = e^{\ln\left(\frac{1}{2}\right)\frac{t}{t_{1/2}}}$
Remember that $e^{\ln\left(\frac{1}{2}\right)} = \frac{1}{2}$
So $f = \frac{1}{2}^{\frac{t}{t_{1/2}}}$

b **i** 0.50

ii 0.25

iii 0.177 ≈ 0.18

iv 0.0032

Answers to SAQs

Chapter 32

1 a Electrical energy from power supply is transferred to energy of electron beam (100%). Energy of electron beam is transferred to internal energy of the anode (~99%) and energy of X-ray photons (~1%).

b Energy $= qV = 80\,\text{keV} = 80\,000 \times 1.6 \times 10^{-19}$
$= 1.28 \times 10^{-14}\,\text{J} \approx 1.3 \times 10^{-14}\,\text{J}$
Impact speed is given by rearranging kinetic energy $E_k = \frac{1}{2}mv^2$ to give
$v = \sqrt{\frac{2E_k}{m}} = \sqrt{\frac{2 \times 1.3 \times 10^{-14}}{9.1 \times 10^{-31}}} = 1.69 \times 10^8\,\text{m s}^{-1}$
$\approx 1.7 \times 10^8\,\text{m s}^{-1}$

2 Photon energy $E = \frac{hc}{\lambda} = V \times e$
Wavelength $\lambda = \frac{hc}{Ve} = \frac{6.63 \times 10^{-34} \times 3.0 \times 10^8}{120 \times 10^3 \times 1.6 \times 10^{-19}}$
$= 1.04 \times 10^{-11}\,\text{m} \approx 1.0 \times 10^{-11}\,\text{m}$

3 When $x = x_{1/2}$ the intensity has dropped to half its initial value. Hence $I = I_0 e^{-\mu x}$ becomes $\frac{I}{2} = I_0 e^{-\mu x_{1/2}}$. Taking logs of both sides gives $\ln(\frac{1}{2}) = -\mu x_{1/2}$ or $\ln 2 = \mu x_{1/2}$. Rearranging gives $x_{1/2} = \ln(2/\mu)$.

4 Intensity $= \frac{P}{A} = \frac{400}{0.0005} = 8.0 \times 10^5\,\text{W m}^{-2}$

5 Intensity $I = I_0 e^{-\mu x} = 50 \times e^{(-1.2 \times 5.0)} = 0.12\,\text{W m}^{-2}$

6 The grid absorbs scattered X-rays, which would otherwise cloud the image.

7 Consider the ratio of attenuation coefficients bone : muscle. This is approximately 6 for 50 keV X-rays, so bone is a much better absorber at this energy than is muscle. At 4.0 MeV, the ratio is less than 2, so bone and muscle will not appear very different on the image. (You could also calculate the fraction of X-rays absorbed by, say, 1 cm of tissue. At 4.0 MeV, only a small fraction is absorbed, so the X-ray image will be flooded with unabsorbed X-rays.)

8 The ratio $\frac{Z_{\text{bone}}}{Z_{\text{soft tissue}}} \approx 2$. Since attenuation coefficient $\mu \propto Z^3$, the ratio $\frac{m_{\text{bone}}}{m_{\text{soft tissue}}} = 2^3 = 8$.

9 Breathing causes movement of the body so that organs or bones of interest may move in the X-ray beam as the image is processed.

10 The skull has bone all round. In a conventional X-ray, the beam must pass through both sides of the skull and this makes it difficult to see the inner tissue. In a CAT scan, the inner tissue shows up more clearly and any damage to the skull bones can be pinpointed accurately.

11 a Rearrange speed $v = f\lambda$ to give wavelength $\lambda = \frac{v}{f} = \frac{5700}{2.1 \times 10^6} = 2.7 \times 10^{-3}\,\text{m}$ (2.7 mm)

b $1.35 \times 10^{-3}\,\text{m} \approx 1.4\,\text{mm}$

12 a Mechanical to electrical
b Electrical to mechanical
c Mechanical to electrical
d Electrical to mechanical

13 $Z = \rho c = 1075 \times 1590 = 1.71 \times 10^6\,\text{kg m}^2\,\text{s}^{-1}$

14 $\frac{I_r}{I_0} = \frac{(Z_2 - Z_1)^2}{(Z_2 + Z_1)^2} = \frac{(1.34 - 1.50)^2}{(1.34 + 1.50)^2} = 3.2 \times 10^{-3}$ (0.32%)

15 There is a big change in acoustic impedance when ultrasound passes from fluid into skin, and from tissue into bone. These surfaces therefore give strong reflections. Other soft tissues have similar values of acoustic impedance and so reflections are very weak.

16 The brain is surrounded by solid bone which reflects ultrasound. Little penetrates the brain and hence the signal is very weak. An alternative is a CT scan.

17 Distance $= \frac{c\Delta t}{2} = \frac{1540 \times 0.034 \times 10^{-3}}{2}$
$= 0.026\,\text{m}$ (26 mm)

18 X-rays are ionising radiation and hence are damaging to the fetus. Ultrasound carries very little risk because it is not a form of ionising radiation. (The intensity used must not cause heating of the baby's tissues.)

Cambridge International AS and A Level Physics © Cambridge University Press 2014

19 **a** Rearrange $f_0 = \frac{\gamma B_0}{2\pi}$ to give gyromagnetic ratio

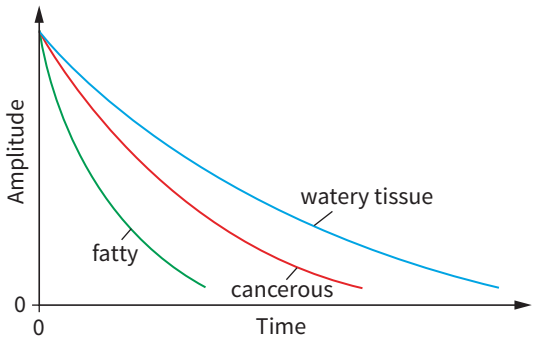
$\gamma = \frac{2\pi f_0}{B_0} = \frac{2 \times \pi \times 42.6 \times 10^6}{1.0} = 2.68 \times 10^8 \text{ Hz T}^{-1}$

For a field of 2.5 T, $f_0 = \frac{\gamma B_0}{2\pi}$

$= \frac{2.68 \times 10^8 \times 2.5}{2 \times \pi} = 106.5 \text{ MHz}$

b This is also their resonant frequency (106.5 MHz).

20



21 **a** MRI uses non-ionising RF electromagnetic radiation, whereas CAT scanning uses X-rays, which are a form of ionising radiation.

b CAT scans show up bone, which is poorly imaged by MRI.

c The patient's body does not have to be cut open, nor do any instruments have to be inserted into the body.

Answers to EOC questions

Chapter 1

1 a distance = speed × time [1]
$= \frac{120 \times 2}{60}$ [1]
= 4.0 km [1]

b The car's direction of motion keeps changing. Hence its velocity keeps changing. In the course of one lap, its displacement is zero, so its average velocity is zero. [2]

c distance travelled in 1 minute
= 0.5 × circumference [1]
but displacement = diameter of track [1]
$= \frac{\text{circumference}}{\pi}$ [1]
$= \frac{4000\text{ m}}{\pi} = 1270\text{ m}$ [1]

2 a By Pythagoras' theorem,
$\text{distance}^2 = 600^2 + 800^2\text{ m}^2$ [1]
$\text{distance} = \sqrt{1\,000\,000} = 1000\text{ m}$ [1]

b angle at B = $\tan^{-1}\left(\frac{800}{600}\right)$ [1]
displacement = 1000 m at an angle 53° E of N [1]

c velocity = $\frac{1000}{60}$ [1]
$= 16.7\text{ m s}^{-1}$ [1]
at an angle 53° E of N [1]

3 a distance in car = 0.25 × 60 = 15 km [1]
total distance = 2.2 + 15 = 17.2 km [1]

b By Pythagoras' theorem, displacement
$= \sqrt{2.2^2 + 15^2}$ [1]
= 15 200 m [1]
at an angle = $\tan^{-1}\left(\frac{2.2}{15}\right)$ [1]
= 8° E of N [1]

c time for 2.2 km at $2.0\text{ m s}^{-1} = \frac{2200}{2} = 1100\text{ s}$ [1]
total time = 1100 + 900 = 2000 s [1]

d average speed = $\frac{\text{distance}}{\text{time}}$ [1]
$= \frac{17\,200}{2000}$ [1]
$= 8.6\text{ m s}^{-1}$ [1]

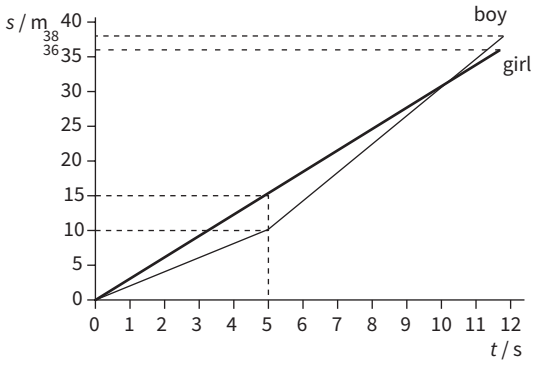
e average velocity = $\frac{\text{displacement}}{\text{time}}$ [1]
$= \frac{15\,200}{2000}$ [1]
$= 7.6\text{ m s}^{-1}$ [1]

4 resultant velocity = $\sqrt{1.0^2 + 2.40^2}$ [1]
$= 2.6\text{ m s}^{-1}$ [1]
at an angle of $\tan^{-1}\left(\frac{1.0}{2.4}\right)$ [1]
= 23° E of N [1]

5 a distance in a (particular) direction [1]

b when athlete returns to his original position **or** the start [1]
(direct) distance from original position zero [1]

6



a straight line from $t = 0$, $s = 0$ to $t = 12$, $s = 36$ [1]

b straight line from $t = 0$, $s = 0$ to $t = 5$, $s = 10$ [1]
straight line from $t = 5$, $s = 10$ to $t = 12$, $s = 38$ [1]

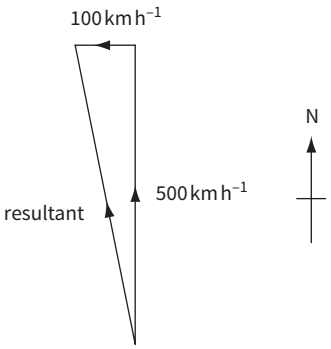
c 10 s where the graphs cross [1]

7 a Each second, it travels a constant distance. [1]
At least two examples:
108 – 84 = 24, 84 – 60 = 24, 60 – 36 = 24 cm [1]

b $s = \frac{d}{t} = \frac{24}{0.1}$ [1]
240 cm s^{-1} [1]

c 108 + 2 × 24 [1]
156 cm [1]

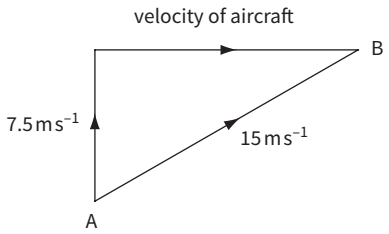
8 a Vector quantities have direction, and scalar quantities do not. [1]
One example of a vector, e.g. velocity, acceleration, displacement, force [1]
One example of a vector, e.g. speed, time, mass, pressure [1]



b Correct vectors drawn and labelled [1]
Scale stated and diagram of sufficient size [1]
Resultant velocity 510 (±10) km h^{-1} [1]
11° W of N **or** a bearing of 349° (±3°) [1]

c 0.25 × 510 = 128 ≈ 130 km 11° W of N

9 a



Correct vector diagram [1]
Velocity of aircraft in still air in easterly direction **or** calculation [1]

b $t = \frac{5000}{15} = 333$ s or $\frac{5000}{13.5} = 370$ s [1]
total time = 703 or 704 s or 703.7 s [1]
average speed = $\frac{10\,000}{703.7} = 14.2$ m s^{-1} [1]

Answers to EOC questions

Chapter 2

1 Using $v^2 = u^2 + 2as$ [1]
distance $s = \frac{(v^2 - u^2)}{2a}$ [1]
$= \frac{(30^2 - 10^2)}{2 \times 4.0}$ [1]
= 100 m [1]

2 Using $v = u + at$ [1]
final velocity = 50 – 0.50 × 100 [1]
= 0 [1]
Using $s = ut + \frac{1}{2}at^2$ [1]
distance travelled = $50 \times 100 - 0.5 \times 0.50 \times 100^2$ [1]
= 2500 m [1]
Train slows to rest and covers a distance of 2500 m [1]

3 **a** Using $s = ut + \frac{1}{2}at^2$ [1]
$s = 20t + 0.5 \times 9.8t^2 = 20t + 4.9t^2$ [1]

b Substituting values of t in the equation gives [1]
after 2.0 s, displacement = 20.4 m ≈ 20 m [1]
after 6.0 s, displacement = –56.4 m ≈ –56 m [1]

c Substituting $s = 0$ gives [1]
$0 = 20t - 4.9t^2$ [1]
$t = \frac{20}{4.9}$ [1]
= 4.08 s ≈ 4.1 s [1]

4 **a** distance travelled at constant speed = speed × time [1]
= 40 × 20 = 800 m [1]

b acceleration = $\frac{\text{change in speed}}{\text{time taken}}$ [1]
$= \frac{(50 - 25)}{20} = 2.5\ \text{m s}^{-2}$ [1]
distance travelled = average speed × time taken [1]
average speed = $\frac{(25 + 50)}{2} = 37.5\ \text{m s}^{-1}$ [1]
distance travelled = 37.5 × 20 = 750 m [1]

c B must travel an extra 50 m; its additional speed is 10 m s^{-1} [1]
so time required = $\frac{50}{10}$ = 5 s [1]

d Consider car A: it travels at 40 m s^{-1} for 25 s [1]
total distance travelled 40 × 2.5 = 1000 m [1]

5 **a** vertical component of velocity = $v \sin 30°$ [1]
$= 5.6 \sin 30° = 2.8\ \text{m s}^{-1}$ [1]
Using $s = ut + \frac{1}{2}at^2$ with $a = g$ [1]
$t = \frac{2.8}{4.9} = 0.57$ s [1]

b horizontal component of velocity = $v \cos 30°$ [1]
$= 5.8 \cos 30° = 4.85\ \text{m s}^{-1} \approx 4.9\ \text{m s}^{-1}$ [1]
horizontal distance = speed × time [1]
$= 4.85 \times 0.57 = 2.77\ \text{m} \approx 2.8\ \text{m}$ [1]

6 **a** $0.2 = \frac{1}{2} \times 9.81 \times t^2$ [1]
= 0.202 s ≈ 0.20 s [1]

b i $v^2 = u^2 + 2as$; $2.91^2 = 1.92^2 + 2a \times 0.25$ [1]
$a = 9.56\ \text{m s}^{-2} \approx 9.6\ \text{m s}^{-2}$ [1]

ii Air resistance [1]
Acts in the opposite direction to the velocity and so reduces the acceleration. [1]

7 **a i** Ball travels upwards (or reverses direction) on bouncing. [1]

ii In both cases, the ball is accelerating due to gravity only. [1]

iii Initial height of the ball above the ground. [1]

iv Ball does not bounce as high as initial position. **or**
(Kinetic) energy is lost (as heat / internal energy) during the bounce. [1]

b i $v^2 = u^2 + 2as$ leading to $v^2 = 2 \times 9.81 \times 1.2$ [1]
$v = 4.85\ \text{m s}^{-1} \approx 4.9\ \text{m s}^{-1}$ [1]

ii $v^2 = 2 \times 9.81 \times 0.8$ [1]
$v = 3.96\ \text{m s}^{-1} \approx 4.0\ \text{m s}^{-1}$ [1]

iii $v = u + at$ leading to $4.85 = -3.96 + a \times 0.16$ [1]
$a = 55.1 \approx 55\ \text{m s}^{-2}$ [1]
upwards direction [1]

8 **a** Tangent drawn at $t = 0.7$ s and gradient of graph determined [1]
$a = 0.8\ (\pm 0.2)\ \text{m s}^{-2}$ [1]

b Acceleration is constant from $t = 0$ to about $t = 0.5$ s [1]
Acceleration decreases from $t = 0.5$ s [1]
Gradient constant from $t = 0$ to $t = 0.5$ s and decreases from $t = 0.5$ s [1]

c Area under the graph used [1]
Correct method, e.g. trapezium rule or squares counted [1]
Distance = 0.20 ± 0.01 m [1]

d Random errors – the points are either side of the line [1]
Systematic errors – the whole line is shifted up or down [1]

9 **a** $v^2 = u^2 + 2as$; $0 = v^2 - 2 \times 2 \times 140$ [1]
$23.7 \approx 24\ \text{m s}^{-1}$ [1]

b $t = \frac{v}{s} = \frac{23.7}{60} = 0.39\ \text{s}$ [1]
The reaction time is approximately 0.3 s, so the driver was alert. [1]

c $100\ \text{km h}^{-1} = \frac{100\,000}{60 \times 60} = 27.8\ \text{m s}^{-1} \approx 28\ \text{m s}^{-1}$ [1]
The driver was not speeding, as the speed of $24\ \text{m s}^{-1}$ is less than the speed limit. [1]

10 **a** constant gradient [1]

b **i** 1.55 (±0.05) s [1]

ii Area under graph calculated between $t = 0$ and $t = 1.55$ s [1]
$= 15 \times \frac{1.55}{2} = 11.6 \approx 12$ m [1]

iii Area between $t = 1.55$ s and $t = 4.1$ s [1]
31.8 ≈ 32 m; accept error carried forward from time in **i** [1]

c **i** The initial speed of the ball or the hot-air balloon is $15\ \text{m s}^{-1}$ [1]

ii The acceleration is in the opposite direction to the initial speed of the ball.
or The acceleration due to gravity is downwards and the ball initially rises. [1]

11 **a** $v^2 = u^2 + 2as$; $20^2 = 0 + 2 \times 9.81 \times s$ [1]
$s = 20.4 \approx 20$ m [1]

b $v = u + at$; $20 = 0 + 9.81 \times t$ [1]
$t = 2.04 \approx 2.0$ s [1]

c distance = $80 \times 2.04 = 163$ m ≈ 160 m [1]

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Answers to EOC questions

Chapter 3

1 a average acceleration = $\frac{\text{change in velocity}}{\text{time}}$ [1]
$= \frac{70}{0.0005} = 140 \times 10^3\ \text{m s}^{-2}$ [1]
average force = mass × acceleration [1]
$= 14\,000 \times 0.046 = 6440\ \text{N}$
or 6.4 kN to 2 sig. figs [1]

b mass = $\frac{\text{weight}}{g}$ [1]
$= \frac{6440}{9.8} \approx 660\ \text{kg}$ [1]

2 a weight = mass × g [1]
$= 70 \times 1.6 = 112\ \text{N}$ [1]

b resultant force = force up – force down
$= 500 - 112 = 388\ \text{N}$ upwards [1]

c acceleration = $\frac{\text{resultant force}}{\text{mass}}$ [1]
$= \frac{388}{70} = 5.54\ \text{m s}^{-2}$ upwards [1]

3 a At first the only force is the weight [1]
but as its speed increases viscous drag increases. [1]
When viscous drag equals weight, the acceleration is zero and the speed is constant. [1]

b Put rubber bands around the cylinder the same vertical distance apart along the cylinder. [1]
Time the ball between the bands. [1]
When terminal velocity is reached, the time taken between successive bands will be constant. [1]

4 a speed = $\frac{3 \times 10^{-6}\ \text{m}}{0.005\ \text{s}}$ [1]
$= 6 \times 10^{-4}\ \text{m s}^{-1}$ [1]

b speed = $\frac{6000\ \text{m}}{3 \times 10^6\ \text{s}}$ [1]
$= 2 \times 10^{-3}\ \text{m s}^{-1}$ [1]

c speed = $\frac{8 \times 10^{-12}\ \text{m}}{4 \times 10^{-9}\ \text{s}}$ [1]
$= 2 \times 10^{-3}\ \text{m s}^{-1}$ [1]

5 a the Earth [1]
upwards [1]
gravitational force [1]

b the Earth or the ground under the man [1]
downwards [1]
contact force [1]

6 a i $F = ma = 1100 \times 1.5 = 1650\ \text{N}$ [1]

ii 1650 + 600 = 2250 N (so that resultant force is still 1650 N) [1]

b $s = ut + \frac{1}{2}at^2 = \frac{1}{2} \times 1.5 \times 10^2$ [1]
$s = 75\ \text{m}$ [1]

7 a $1.5\ \text{m s}^{-1}$ [1]

b Constant velocity is reached when weight = upward force due to air resistance. [1]
Air resistance increases with speed. [1]
Air resistance is less than the weight of the metal ball even at 2.5 to 3.0 m s^{-1}. [1]

c Initial acceleration is acceleration due to gravity **or** $9.81\ \text{m s}^{-2}$ [1]
Initially neither ball has any air resistance. [1]

8 a $F = ma = 1200 \times \frac{8}{2}$ [1]
$F = 4800\ \text{N}$ [1]

b i kg m s^{-2} [1]

ii kg m^{-1} [1]

iii $4800 = b \times 50^2$ [1]
$b = 1.92$ (kg m^{-1} **or** $\text{N s}^2\ \text{m}^{-2}$) [1]

iv Sketch graph showing increasing gradient and force values marked at speeds of 0 and $50\ \text{m s}^{-1}$ [1]
Resistive force increases with speed, so resultant force and acceleration decrease. [1]

9 a Mass is the amount of matter in a body. [1]
Weight is a force [1]
due to gravity acting on the body. [1]

b For example, body moves to the Moon **or** rises above the Earth. [1]
Amount of matter is constant, but force due to gravity is less on the Moon **or** at altitude. [1]

c mass: kg [1]
weight: kg m s^{-2} [1]

10 a For a body of constant mass, the acceleration is directly proportional to the resultant or net force applied to it. [1]
Directions of the acceleration and the resultant force are the same. [1]

b i It increases the time. [1]

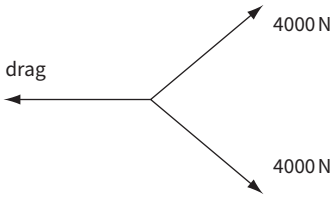
ii If time increases then acceleration decreases. [1]
Since $F = ma$, when acceleration is less, the net force is less, and there is less force between the ground and the legs. [1]

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Answers to EOC questions

Chapter 4

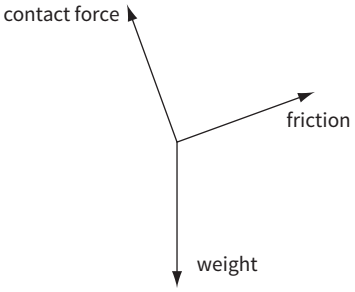
1 a



[2]

b correct diagram to scale [1]
= 6130 N + 70 N [1]
(allow also calculation from components)

2 a



[1]

b component of weight down slope
= weight × cos 70° **or** weight × sin 20° [1]
= 1.5 × 9.8 × cos 70° = 5.03 N
= 5.0 N to 2 sig. figs [1]

c Friction balances the component of weight down slope (as contact force is at 90° to slope), [1]
so friction = 5.03 N up the slope
= 5.0 N to 2 sig. figs [1]

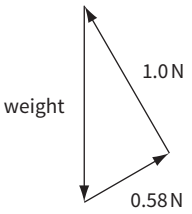
d The contact force balances the component of weight at 90° to slope, [1]
so contact force = weight × cos 20° [1]
= 1.5 × 9.8 cos 20° = 13.8 N
= 14 N to 2 sig. figs [1]

3 a Horizontal component of tension in string 1
= tension × cos 60° [1]
= 1.0 × cos 60° = 0.50 N to the left [1]
Horizontal component of tension in string 2
= tension × cos 30° [1]
= 0.58 × cos 30° = 0.50 N to the right [1]
These components cancel, as there is no resultant horizontal force. [1]

b Vertical component of tension in string 1
= tension × cos 30° [1]
= 1.0 × cos 30° = 0.87 N upwards [1]
Vertical component of tension in string 2
= tension × cos 60° [1]
= 0.58 × cos 60° = 0.29 N upwards [1]

c Weight is balanced by upward components of tensions [1]
= 0.87 + 0.29 = 1.16 N = 1.2 N to 2 sig. figs [1]

d



[1]

e By Pythagoras' theorem, [1]
weight = 1.16 N = 1.2 N to 2 sig. figs [1]

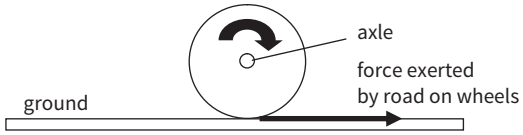
4 Component of F at 90° to lever = $F \sin 45°$ [1]
Moment of F about pivot = force × distance [1]
= $2.0 \times F \sin 45° = 40$ N m (from the question) [1]
so $F = \frac{40}{2 \sin 45°} = 28.3$ N = 28 N to 2 sig. figs [1]

5 Taking moments about the pivot: [1]
$3.3 \times 0.100 + 7.6 \times 0.060 = P \times 0.080$ [1]
$P = \frac{0.786}{0.080}$ [1]
= 9.83 N = 9.8 N to 2 sig. figs [1]

6 a i A couple is a pair of equal and opposite forces that act at different points. [1]

ii The torque of a force about a point is the product of the force and distance. [1]
The distance is the perpendicular distance of the line of action of the force to the point. [1]

b i rotation shown clockwise [1]
force drawn forwards [1]



ii zero [1]
The car moves at constant speed. **or**
The wheel turns at a constant rate. [1]

iii force = $\frac{\text{torque}}{\text{radius}} = \frac{200}{0.29} = 690$ N [1]

7 a The centre of gravity of an object is the point where all the weight of the object [1]
may be considered to act. [1]

b i Taking moments about the fixed end of the flagpole:

Sum of clockwise moments = Sum of anticlockwise moments

$(25 \times 9.81) \times 1.5 = Tx$ [1]

where x = perpendicular distance of the line of action of the tension from the fixed end of the flagpole, given by

$x = 2.5 \sin 30° = 1.25$ m [1]

$(25 \times 9.81) \times 1.5 = T \times 1.25$ [1]

$T = \frac{25 \times 9.81 \times 1.5}{1.25} = 294 \approx 290$ N [1]

ii The net vertical force = 0

Vertical component of force at fixed end + vertical component of T = weight

Vertical component of force at fixed end + $294 \sin 30° = 25 \times 9.81$ [1]

Vertical component of force ≈ 98 N [1]

8 a No net / resultant force [1]

No net / resultant moment [1]

b i Torque of the couple about the centre $= 30 \times 90 = 2700$ N cm [1]

ii Moment of force about the centre $= (T \times 24)$ N cm [1]

For equilibrium: $24T = 2700$ [1]

$T = 113 \approx 110$ N [1]

9 a The torque of a force about a point is the product of the force and distance. [1]

The distance is the perpendicular distance of the force to the point. [1]

b Tension in string B = vertical component of A = $8.0 \sin 50°$ [1]

$T_B = 6.13 \approx 6.1$ N [1]

Tension in string C = horizontal component of A = $8.0 \cos 50°$ [1]

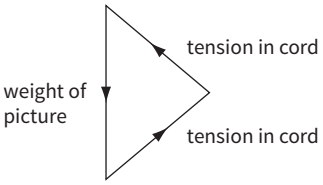
$T_C = 5.14 \approx 5.1$ N [1]

10 a No rotation [1]

No movement in any direction [1]

b Two tension forces in the cord [1]

Closed triangle of forces or parallelogram of forces, including weight [1]



c i Vertical component = $45 \sin 50°$
$= 34.5 \approx 35$ N [1]

ii Weight = sum of vertical components
$= 2 \times 34.5 = 69$ N [1]

Answers to EOC questions

Chapter 5

1 a Loss of gravitational potential energy → gain in kinetic energy [1]

b Kinetic energy → heat (in the brakes) [1]

c Loss of gravitational potential energy → gain in kinetic energy [1]

2 a i Horizontal distance moved = $0.5 \times 16 = 8.0$ m [1]
Horizontal component of force = $200 \cos 30° = 173$ N [1]
Work done = $173 \times 8.0 = 1.39 \times 10^3$ J ≈ 1.4 kJ [1]

ii Weight acts at 90° to displacement [1]
so work done = 0 J [1]

iii Contact force acts at 90° to displacement [1]
so work done = 0 J [1]

b $86.6 \approx 87$ W

3 Truck: k.e. = $\frac{1}{2}mv^2 = 9$ MJ [1]
Dust particle: k.e. = $\frac{1}{2}mv^2 = 14$ MJ [1]
The dust particle has greater kinetic energy than the truck. [1]

4 a Gain in g.p.e. = mgh [1]
$= 950 \times 9.8 \times 50 = 4.66 \times 10^5$ J $\approx 4.7 \times 10^5$ J [1]

b time = $\frac{\text{work done}}{\text{power}}$ [1]
$t = \frac{4.66 \times 10^5}{4000} = 116.5 \approx 120$ s [1]

c Wasted power = 2.9 kW [1]
Wasted energy = power × time [1]
$= 2900 \times 116.5 = 3.4 \times 10^5$ J [1]

5 a Rate at which work is done; W [2]

b kinetic energy = $\frac{1}{2} \times \text{mass} \times \text{velocity}^2$ [1]

c power = $\frac{\text{gain in k.e.}}{\text{time taken}}$ [1]
$= \frac{\frac{1}{2} \times 1100 \times 18^2}{25} = 7130 \approx 7.1$ kW [1]

6 a i Vertical distance = $40 \sin 5° = 3.49$ m [1]
p.e. lost = $mgh = 90 \times 9.81 \times 3.49$ [1]
$= 3078 \approx 3100$ J [1]

ii k.e. increase = $\frac{1}{2}mv^2 = \frac{1}{2} \times 90 \times 12^2$ [1]
$= 6480$ J [1]

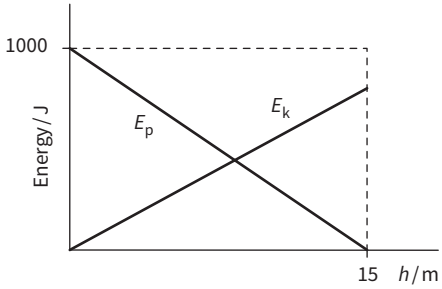
b i Energy produced by cyclist = $6480 - 3078 = 3402$ J [1]
useful power output = $\frac{\text{energy}}{\text{time}} = \frac{3402}{67}$ [1]
power = $50.8 \approx 51$ W [1]

ii Energy is wasted [1]
as work done against friction in the axle / chain or against air resistance. [1]

7 a Work is the product of force and distance moved. [1]
The distance moved is in the direction of the force. [1]

b i As he falls, his potential energy decreases, his kinetic energy increases and internal energy (thermal energy / heat) is produced as work is done against friction. [1]
The decrease in p.e. = increase in k.e. + internal energy produced. [1]

ii Graph with axes labelled and E_p decreasing linearly from 1000 J to 0 [1]
E_k increasing from 0 when $h = 0$ [1]
E_k increases as a straight line to a value below 1000 J at $h = 15$ m [1]



8 a k.e. = work done on body to increase speed from 0 to $v = F \times s$ [1]
since $F = ma$ and $v^2 = u^2 + 2as$ and $u = 0$, $s = \frac{v^2}{2a}$
k.e. = $ma \times \frac{v^2}{2a} = \frac{1}{2}mv^2$ [1]

b i k.e. = $\frac{1}{2}mv^2 = \frac{1}{2} \times 800 \times 20^2 = 160\,000$ J [1]
power = $\frac{\text{energy}}{\text{time}} = \frac{160\,000}{6} = 2.67 \times 10^4$
$\approx 2.7 \times 10^4$ W [1]

ii Air resistance increases (with speed). [1]
Net driving force less or more energy (per second) wasted, so less available to increase k.e. [1]

Cambridge International AS and A Level Physics © Cambridge University Press 2014

9 a i The potential energy of a body is the energy stored in the body by reason of its position or shape. [1]

ii Gravitational p.e. is energy due to position in a gravitational field. [1]
Elastic p.e. is energy contained in a stretched or squashed object. [1]
or
When an object is raised above the Earth's surface its gravitational p.e. increases. [1]
When a positively charged object is brought near another positively charged object its electric p.e. increases. [1]
(maximum [2])

b i mass = density × volume
$= 1030 \times 1.4 \times 10^6 \times 10.0 = 1.442 \times 10^{10}$
$\approx 1.4 \times 10^{10}$ J [1]

ii Water falls an average distance of 5.0 m when released through turbines.
p.e. lost = $mgh = 1.442 \times 10^{10} \times 9.81 \times 5.0$
(accept also use of $h = 10$ m) [1]
p.e. lost = $7.07 \times 10^{11} \approx 7.1 \times 10^{11}$ J
(**or** 1.4×10^{12} J if $h = 10$ used) [1]

iii Energy output = 0.5 × (value in **ii**)
$= 3.5 \times 10^{11}$ J (**or** 7.1×10^{11} J if $h = 10$ used) [1]
$\text{power} = \frac{\text{energy}}{\text{time}} = \frac{3.5 \times 10^{11}\text{ J}}{6 \times 60 \times 60}$
(**or** $\frac{7.1 \times 10^{11}\text{ J}}{6 \times 60 \times 60}$ if $h = 10$ used) [1]
power = $1.64 \times 10^7 \approx 1.6 \times 10^7$ W
(**or** 3.3×10^7 W if $h = 10$ used) [1]

Answers to EOC questions

Chapter 6

1 The law of conservation of momentum applies if the Earth is considered to rise as the ball falls. The momentum of the Earth upwards equals the momentum of the ball downwards. [1]
The weight of the ball has an equal and upwards force on the Earth due to Newton's third law. [1]

2 a Momentum of ball before striking wall
= mass × velocity
= 2 × 3.0 = 6.0 kg m s^{-1} towards the ball [1]
Momentum after striking the wall
= 6.0 kg m s^{-1} away from the wall [1]
Change in momentum of ball = 12 kg m s^{-1} away from the wall. [1]

b There is no change in kinetic energy as the ball's speed and mass are unchanged. [1]

3 a Linear momentum = mass × velocity [1]

b [units of mass] × [units of velocity] = kg m s^{-1} [1]

c Using $v^2 = 2as$,
$v = \sqrt{(2 \times 3.5 \times 40)} = \sqrt{280} = 16.7$ m s^{-1} [1]
so momentum = mass × speed = 900 × 16.7
$= 1.5 \times 10^4$ kg m s^{-1} [1]

d Combined momentum to left
= 3.0 × 4.0 – 2.0 × 4.0 = 4.0 kg m s^{-1} [1]
Combined mass = 8.0 kg [1]
So velocity after collision
$= \frac{4.0}{8.0} = 0.50$ m s^{-1} to the left [1]

4 a i In an elastic collision, both momentum and kinetic energy are conserved. [1]

ii In an inelastic collision, momentum is conserved but not kinetic energy. [1]

b Change in momentum
= momentum after – momentum before [1]
= 0.35 × 2.5 – 0.35 × (–2.8)
= 1.855 kg m s^{-1} ≈ 1.9 kg m s^{-1} [1]

c When the table (plus the Earth) is also considered, then the initial momentum of the ball is equal to the final momentum of the ball added to the momentum of the snooker table, and so momentum is conserved. [1]

5 a Change in momentum
= mass × change in velocity [1]
= 1100 × (–24) = –26 400 N s ≈ –26 000 N s [1]

b Force = $\frac{\text{change in momentum}}{\text{time}}$ [1]
$= \frac{26\,400}{20} = 1320$ N ≈ 1300 N [1]

c Average speed during braking = 12 m s^{-1} [1]
so distance travelled in 20 s = 12 × 20 = 240 m [1]

6 a Momentum = mass × velocity [1]
= 0.10 × 0.40 = 0.040 kg m s^{-1} [1]

b For each marble, component of momentum in x-direction = half of original momentum
= 0.020 kg m s^{-1} [1]
so momentum of one marble = $\frac{0.020}{\sin 45°}$
= 0.0283 kg m s^{-1} [1]
and velocity = $\frac{\text{momentum}}{\text{mass}} = \frac{0.0283}{0.10} = 0.283$ m s^{-1}
≈ 0.28 m s^{-1} [1]

c k.e. before = $\frac{1}{2}mv^2 = \frac{1}{2} \times 0.10 \times 0.40^2 = 0.0080$ J [1]
k.e. after = $2 \times \frac{1}{2} \times 0.10 \times 0.283^2 = 0.0080$ J [1]

7 a Initial momentum of ball = 0.16 × 25
= 4.0 kg m s^{-1} [1]
Change in momentum = 4.0 – (–4.0)
= 8.0 kg m s^{-1} [1]

b force = $\frac{\text{change in momentum}}{\text{time}} = \frac{8}{0.003}$ [1]
= 2667 N ≈ 2700 N [1]

c The bat slows down. [1]
The law of conservation of momentum requires that the change in momentum of the ball and of the bat are equal but in opposite directions. [1]
Energy is neither created nor destroyed, but thermal energy (heat / internal energy) and sound are created from the drop in k.e. (of the bat). [1]
The impact is non-elastic. [1]

8 a The total momentum before the collision is equal to the total momentum after the interaction. [1]

The system is closed.
or There are no external forces acting. [1]

b i final momentum = initial momentum
$0.35v = 0.25 \times 30$ [1]
$v = 21.4 \approx 21\ \text{m s}^{-1}$ [1]

ii Change in momentum
$= 0.25 \times 30 - 0.25 \times 21.4$ [1]
change in momentum $= 2.14 \approx 2.1\ \text{kg m s}^{-1}$
or $2.15 \approx 2.2\ \text{kg m s}^{-1}$ [1]

iii Change in total kinetic energy
$= \frac{1}{2} \times 0.25 \times 30^2 - \frac{1}{2} \times 0.35 \times 21.4^2$ [1]
change in total k.e. $= 32.4 \approx 32\ \text{J}$ [1]

iv The arrow stops and the ball moves off with a speed of $30\ \text{m s}^{-1}$ [1]
Relative speed remains unaltered in an elastic collision, $30\ \text{m s}^{-1}$ [1]

9 a i The total kinetic energy before the collision is equal to the total kinetic energy after the collision. [1]

ii In a completely inelastic collision, the maximum amount of kinetic energy is lost (subject to the law of conservation of momentum, which must be obeyed). [1]

b i Momentum is conserved, as there are no external forces / the system is closed. [1]
Momentum of alpha-particle in one direction must equal that of uranium nucleus in the exactly opposite direction for the change to be zero. [1]

ii $6.65 \times 10^{-27} \times v_\alpha + 3.89 \times 10^{-25} \times v_x = 0$ [1]

iii $\frac{v_\alpha}{v_x} = -58.5 \approx -58$ **or** -59
[1]

10 a Momentum and kinetic energy [1]

b i Momentum $= 0.014 \times 640 = 8.96$
$\approx 9.0\ \text{kg m s}^{-1}$ [1]

ii Bullets leave with momentum forwards and gun has equal momentum backwards. [1]
To stop motion / momentum of the gun, the soldier must provide a force. [1]

iii $F = \frac{\Delta p}{\Delta t}$; $140 = n \times 8.96$ [1]
Number of bullets per second = 15.6 **or** 15 **or** 16. [1]

11 a

	Change in momentum / kg m s^{-1}	Initial kinetic energy / J	Final kinetic energy / J
Truck X	6.0×10^4	2.5×10^5	4.0×10^4
Truck Y	6.0×10^4	1.5×10^4	1.35×10^5

One mark for each correct change in momentum. [2]
One mark for correct kinetic energy values for X. [2]
One mark for correct kinetic energy values for Y. [2]

b Total initial k.e. $= 2.65 \times 10^5$ J
and total final k.e. $= 1.75 \times 10^5$ J [1]
Collision is not elastic, because the total k.e. has decreased in the collision. [1]

c Force $= \frac{\Delta p}{\Delta t} = \frac{6.0 \times 10^4}{1.6}$ [1]
$3.75 \times 10^4 \approx 3.7$ **or** 3.8×10^4 N [1]

Answers to EOC questions

Chapter 7

1 a i Density is the mass per unit volume of a substance. [1]

ii Base units kg m^{-3}. [1]

b i Pressure is the normal force acting per unit cross-sectional area. [1]

ii Base units given by $\text{N m}^{-2} = \text{kg m s}^{-2}\,\text{m}^{-2}$
$= \text{kg m}^{-1}\,\text{s}^{-2}$. [1]

2 The *y*-axis should be labelled *F* / N and the *x*-axis labelled extension / m. [1]
Your graph should have a straight line from the origin to the point where force = 5 N and extension = 0.25 m. [1]
After the straight line, the graph continues with a positive gradient but the gradient decreases. [1]

3 Your diagram should show the two springs connected one below the other. [1]
Extension of one spring = $\frac{\text{load}}{\text{spring constant}}$ [1]
$= \frac{2.0}{20} = 0.10\,\text{m}$ [1]
Each spring has the same force, so the same extension. [1]
So total extension = 0.20 m [1]

4 a Cross-sectional area = $\pi r^2 = 1.96 \times 10^{-7}\,\text{m}^2$
$\approx 2.0 \times 10^{-7}\,\text{m}^2$ [1]

b Weight = stress × area [1]
$= 2.0 \times 10^8 \times 1.96 \times 10^{-7} = 39.3\,\text{N} \approx 39\,\text{N}$ [1]

c Strain = $\frac{\text{extension}}{\text{length}}$ [1]
$= \frac{0.050}{1.0} = 0.050$ **or** 5.0% [1]

d Young modulus = $\frac{\text{stress}}{\text{strain}}$ [1]
$= \frac{2.0 \times 10^8}{0.050} = 4.0 \times 10^9\,\text{Pa}$ [1]

5 a For the point at the top of the straight line:
stress = $\frac{\text{force}}{\text{area}} = \frac{10.0}{1.5 \times 10^{-7}} = 6.67 \times 10^{-7}\,\text{Pa}$ [1]
strain = $\frac{\text{extension}}{\text{original length}} = \frac{0.8 \times 10^{-3}}{2.0} = 0.4 \times 10^{-3}$ [1]
Young modulus = $\frac{\text{stress}}{\text{strain}} = 1.67 \times 10^{11}\,\text{Pa}$
$\approx 1.7 \times 10^{11}\,\text{Pa}$ [1]

b Energy stored = ½ × force × extension [1]
$= 0.5 \times 10.0 \times 0.8 \times 10^{-3} = 4.0 \times 10^{-3}\,\text{J}$ [1]

c Work done = average force × extension [1]
$= 7.5\,\text{N} \times 0.4 \times 10^{-3}\,\text{m} = 3.0 \times 10^{-3}\,\text{J}$ [1]

6 a Use a liquid with a high density, e.g. mercury [1]
and have long tubes in the manometer. [1]

b If the area of the tube increases, then the weight of liquid above any point increases in proportion. Because pressure = $\frac{\text{force}}{\text{area}}$, increasing both the force and the area by the same factor leaves pressure unchanged. [2]

7 a P has largest Young modulus. [1]
P has steepest gradient **or** largest value of stress / strain (before breaking). [1]

b Labelled diagram showing wire and weights used to stretch wire. [1]
Method of viewing extension, e.g. travelling microscope and marker. [1]
Series of weights and extensions noted. [1]
Diameter of wire measured. [1]
Initial length of wire measured to 'marker'. [1]
Graph of force (*y*-axis) against extension (*x*-axis) drawn. [1]
Young modulus $E = \frac{FL}{Ax}$
= gradient of force–extension graph × $\frac{\text{length}}{\text{area}}$ [1]

8 a Tensile stress is the ratio of the force to cross-sectional area in a wire. [1]
Tensile strain is the ratio of extension to initial length. [1]

b i stress = $\frac{60}{1.3 \times 10^{-6}}$ [1]
$= 4.62 \times 10^7 \approx 4.6 \times 10^7\,\text{Pa}$ [1]

ii strain = $\frac{\text{stress}}{E} = \frac{4.62 \times 10^7}{2.1 \times 10^{11}}$ [1]
$= 2.2 \times 10^{-4}$ [1]

iii extension = strain × original length
$= 2.2 \times 10^{-4} \times 1.6$ [1]
extension = $3.52 \times 10^{-4} \approx 3.5 \times 10^{-4}\,\text{m}$ [1]

9 a Tensile [1]
Railway line expands on heating, reducing stress.

or Railway line becomes unstressed when line expands with temperature and stresses are less than for higher temperatures. [1]

b i stress = $\frac{F}{A} = \frac{2.6 \times 10^5}{5.0 \times 10^{-3}} = 5.2 \times 10^7$ [1]

strain = $\frac{\text{stress}}{E} = \frac{5.2 \times 10^7}{2.1 \times 10^{11}}$ [1]

$= 2.48 \times 10^{-4} \approx 2.5 \times 10^{-4}$ m (0.025 %) [1]

ii temperature rise = $\frac{2.48 \times 10^{-4}}{1.4 \times 10^{-5}}$ [1]

$= 17.7 \approx 18$ °C [1]

10 a E = gradient of the graph [1]

$= \frac{1.2 \times 10^9}{8.0 \times 10^{-3}}$ [1]

$= 1.5 \times 10^{11}$ Pa [1]

b From the graph, strain = 4.0×10^{-3} [1]

strain = $\frac{x}{L}$

$x = 4.0 \times 10^{-3} \times 3.5 = 1.4 \times 10^{-2}$ m (1.4 cm) [1]

c Breaking stress = $\frac{\text{breaking force}}{\text{cross-sectional area}}$ [1]

breaking force = $1.52 \times 10^9 \times [\pi \times (0.42 \times 10^{-3})^2]$ [1]

$= 842 \approx 840$ N [1]

d At 0.6 GPa, force

$= 0.6 \times 10^9 \times [\pi \times (0.42 \times 10^{-3})^2] = 333$ N [1]

Energy stored = $\frac{1}{2}$ Fx = $\frac{1}{2} \times 333 \times 1.4 \times 10^{-2}$ [1]

$= 2.33 \approx 2.3$ J [1]

11 a i The spring constant **or** the stiffness of the spring [1]

ii The energy stored in the spring [1]

b i energy stored = $\frac{1}{2}kx^2 = \frac{1}{2} \times 80 \times 0.06^2$ [1]

$= 0.144 \approx 0.14$ J [1]

ii The law of conservation of momentum must be obeyed. [1]

The momentum of each trolley must be equal (but opposite) and, as momentum = mv and the masses are equal, then speeds must be the same. [1]

iii $2 \times (\frac{1}{2}mv^2) = 0.144$ [1]

$v = 0.60$ m s^{-1} [1]

12 a i mass = density × volume = ρAh [1]

ii pressure = $\frac{\text{force or weight of liquid}}{\text{area}}$ [1]

pressure = $\frac{\rho Ahg}{A} = \rho hg$ [1]

b i height difference in manometer

= 60 cm = 0.60 m [1]

pressure difference = ρhg

$= 1000 \times 0.6 \times 9.81$

$= 5886 \approx 5.9 \times 10^3$ Pa [1]

ii force or weight = pressure × area

$= 5886 \times 0.05$ [1]

$= 294 \approx 290$ N [1]

13 a $P = \rho hg = 1000 \times 0.5 \times 9.81$ [1]

$= 4905 \approx 4900$ Pa [1]

b At the same depth in the same liquid. **or** If pressure is different, liquid would flow from X to Y (and there is no depth difference). [1]

c Force down on base = pressure at X or Y multiplied by area of base [1]

The weight of the liquid is the downwards force on the base minus the upwards force on the top horizontal surface in the liquid. **or** Force down on base assumes liquid is all of height 0.5 m above base but container has some parts with water of depth only 0.3 m. [1]

14 a The weight of the air above 1 m^2 of area is different between the bottom of the mountain and the top. [1]

b i Liquid Y has a smaller height difference (50 cm) to the boundary than does liquid X (60 cm). [1]

Pressure is the same at the 20 cm mark on both sides, so the smaller volume of Y has the same weight (and hence mass) as liquid X. [1]

ii $(\rho hg)_Y = (\rho hg)_X$ [1]

$800 \times 0.5 \times 9.81 = \rho X \times 0.6 \times 9.81$ [1]

$\rho X = 666 \approx 670$ kg m^{-3} [1]

iii Weight due to Y and 10 cm of X = weight due to 70 cm of X. [1]

iv $\rho hg = 666 \times 9.81 \times 0.7$ [1]

pressure = 4580 ≈ 4600 Pa [1]

Answers to EOC questions

Chapter 8

1 **a** The foil is positively charged and experiences a force in the same direction as the electric field. [1]

b The foil will become negatively charged [1] and will experience a force in the opposite direction to the field. [1]

2 Field strength $E = \frac{F}{Q}$ [1]

$= \frac{4.4 \times 10^{-13}\,\text{V}}{8.8 \times 10^{-17}\,\text{C}} = 5000\,\text{N C}^{-1}$ [1]

3 p.d. $V = E \times d$ [1]

$= 4000 \times 0.04 = 160\,\text{V}$ [1]

4 **a** Separation $d = \frac{V}{E}$ [1]

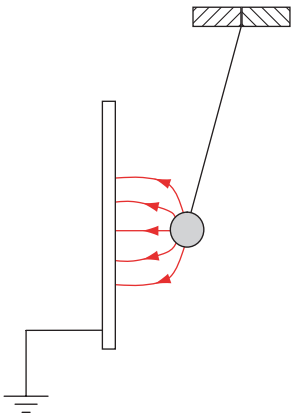
$= \frac{2400}{3.0 \times 10^{4}} = 0.08\,\text{m} = 8.0\,\text{cm}$ [1]

b Field strength $E = \frac{V}{d}$ [1]

$= \frac{2400}{0.02} = 1.2 \times 10^{5}\,\text{V m}^{-1}$ [1]

5 The field is directly proportional to the p.d., so doubling the p.d. doubles the field strength. [1]
The field strength is inversely proportional to the plate separation, so reducing the separation by a factor of 3 trebles the field strength. [1]
Therefore the electric field strength is increased by a factor of 6. [1]

6 **a**



Five good lines, not touching, good shape [2]
(three good, not touching and good shape [1])
Arrows in correct direction [1]

b The positive charge on the sphere induces negative charges on the plate. [1]
The opposite charges attract. [1]

c **i** The sphere would still be attracted to the plate. [1]
The negative charge on the sphere now induces positive charges on the plate. [1]

ii The field direction would reverse. [1]
(**but** any indication that the shape changes [0])

7 **a** A series of parallel lines between the plates [1]
Arrows vertically downwards [1]

b Vertically downwards [1]

c $6.4 \times 10^{-14}\,\text{N}$ [1]

d $E = \frac{F}{q} = \frac{6.4 \times 10^{-14}}{1.6 \times 10^{-19}}$ [1]
$E = 400\,000\,\text{V}$ [1]

e $E = \frac{V}{d}$ leading to
$V = dE = 400\,000 \times 2.5 \times 10^{-2}$ [1]
$= 10\,000\,\text{V}$ [1]

8 **a** Electric field strength is
force per unit charge [1]
per unit **positive** charge [1]

b $E = \frac{V}{d} = \frac{5.0 \times 10^{6}}{8.0 \times 10^{-2}}$ [1]
$F = EQ = \left(\frac{5.0 \times 10^{6}}{8.0 \times 10^{-2}}\right) \times 1.6 \times 10^{-19}$ [1]
$F = 1.0 \times 10^{-11}\,\text{N}$ [1]

c $W = F\Delta x$
$= 1.0 \times 10^{-11} \times 8.0 \times 10^{-2}$ [1]
work done $= 8.0 \times 10^{-13}\,\text{J}$ [1]

d $8.0 \times 10^{-13}\,\text{J}$ [1]

e $E_k = \frac{1}{2}mv^2 = 8.0 \times 10^{-13}$ [1]
$v^2 = \frac{2 \times 8.0 \times 10^{-13}}{1.7 \times 10^{-27}}$ [1]
$v = 3.1 \times 10^{7}\,\text{m s}^{-1}$ [1]

9 **a** **i** Arrows from the inner electrode to the outer electrode. [1]

ii Lines are closer together. [1]

b $\Delta V = E\Delta x = 5.0 \times 10^{6} \times 1.25 \times 10^{-3}$ [1]
$\Delta V = 6250\,\text{V}$ [1]

c Given that $E = 5.0 \times 10^{6}\,\text{N C}^{-1}$ (**or** $5.0 \times 10^{6}\,\text{V m}^{-1}$), in a distance of 4.0 μm the potential drops
$5.0 \times 10^{6} \times 4.0 \times 10^{-6}\,\text{V}$ [1]
potential drop = 20 V [1]
(Other routes are possible)

Answers to EOC questions

Chapter 9

1 $Q = It$ [1]
$= 0.150 \times 40 \times 60$ [1]
$= 360\,\text{C}$ [1]

2 $t = \frac{Q}{I}$ [1]
$= \frac{2000}{40} = 50\,\text{s}$ [1]

3 $Q = It$ [1]
$= 30 \times 10^{3} \times 2000 \times 10^{-6}$ [1]
$= 60\,\text{C}$ [1]

4 **a** $I = \frac{V}{R} = \frac{4.5}{15}$ [1]
$= 0.30\,\text{A}$ [1]
b $R = \frac{V}{I} = \frac{230}{6.5}$ [1]
$= 35\,\Omega$ [1]
c $V = IR = 2.4 \times 3.5$ [1]
$= 8.4\,\text{V}$ [1]

5 **a** $Q = It = 2.4 \times 10 \times 60$ [1]
$= 1440\,\text{C}$ [1]
b $W = QV = 1440 \times 6.0$ [1]
$= 8640\,\text{J}$ [1]

6 $W = QV = 1.6 \times 10^{-19} \times 50 \times 10^{-3}$ [1]
$= 8.0 \times 10^{-15}\,\text{J}$ [1]

7 Current taken by the hairdryer $= \frac{P}{V} = \frac{450}{120}$ [1]
$= 3.75\,\text{A}$ [1]
So fuse taken is 5 A, the nearest above the supplied current. [1]

8 **a** **i** Arrow from left to right [1]
ii Arrow (and label) from left-hand electrode to positive terminal **or** from negative terminal to right-hand electrode [1]
b **i** Charge = charge per ion × number of ions
$= 1.6 \times 10^{-19} \times 3.6 \times 10^{16}$ [1]
charge $= 5.76 \times 10^{-3} \approx 5.8 \times 10^{-3}\,\text{C}$ [1]
ii $I = \frac{\Delta Q}{\Delta t} = \frac{5.8 \times 10^{-3}}{8.0 \times 60}$ [1]
$I = 0.000\,012\,\text{A} = 12\,\mu\text{A}$ [1]

9 **a** Recognition that current through milliammeter = current through tube [1]
$Q = It = 4.5 \times 10^{-3} \times 3 \times 60$ [1]
$Q = 0.81\,\text{C}$ [1]
b $n = \frac{\text{total charge}}{\text{charge per electron}} = \frac{8.1 \times 10^{-1}}{1.6 \times 10^{-19}}$ [1]
$n = 5.06 \times 10^{18} \approx 5.1 \times 10^{18}$ [1]
Correct powers of 10 throughout [1]
c $W = VQ = 75 \times 1.6 \times 10^{-19}$ [1]
$W = 1.2 \times 10^{-17}\,\text{J}$ [1]

10 **a** Electron number density n = number of atoms in $1\,\text{m}^3$ = number of atoms in 8900 kg [1]
$= \frac{8900 \times 6.0 \times 10^{26}}{54} = 9.89 \times 10^{28}\,\text{m}^{-3} \approx 10^{29}\,\text{m}^{-3}$ [1]
b Drift velocity $v = \frac{I}{nAe}$ [1]
$= \frac{3.5 \times 10^{-3}}{(9.9 \times 10^{28} \times 5.0 \times 10^{-8} \times 1.6 \times 10^{-19})}$ [1]
$= 4.4\,\text{mm s}^{-1}$ [1]

11 **a** The potential difference across the terminals of a cell is the work done per coulomb of charge in the external circuit. [1]
The e.m.f. of a cell is the work done per coulomb of charge in the complete circuit. [1]
b **i** $Q = It = 100 \times 3600$ [1]
$Q = 360\,000\,\text{C}$ [1]
ii $W = VQ = 12 \times 360\,000$ [1]
$W = 4.3 \times 10^{6}\,\text{J}$ [1]
c $R = \frac{V^2}{W}$ [1]
$R = \frac{12^2}{27}$ [1]
$R = 5.33 \approx 5.3\,\Omega$ [1]

12 **a** $W = VQ = 1000 \times 3600$ [1]
$W = 3600\,000\,\text{J} = 3.6\,\text{MJ}$ [1]
b **i** $I = \frac{P}{V} = \frac{9500}{230}$ [1]
$I = 41.3 \approx 41\,\text{A}$ [1]
ii Very large current could cause heating of wires or voltage drop on sockets. [1]
iii 50 A (or any value up to 100 A) [1]
c $W = Pt = 9500 \times 5 \times 60$ [1]
$W = 2.85 \approx 2.9\,\text{MJ}$ [1]

Answers to EOC questions

Chapter 10

1 **a** $W = 3.6\,\text{A}$ to the right [1]

b $X = 4.3 - 2.4 = 1.9\,\text{A}$ downwards [1]

c $Y = 4.8 - 2.7 = 2.1\,\text{A}$ to the left [1]

d $Z = 0$ [1]

2 $X = 6.5 - 2.0 = 4.5\,\text{mA}$ to the right [1]

$Y = 4.5 - 4.2 = 0.3\,\text{mA}$ downwards [1]

3 **a** $X = 2.2 - 1.4 = 0.8\,\text{V}$ [1]

b $X = 6.3 + 2.4 - 5.0 = 3.7\,\text{V}$ [1]

c $X = 6.0 - 2.4 - 1.4 = 2.2\,\text{V}$ [1]

d $X = 4.3 + 4.7 = 9.0\,\text{V}$ [1]

$Y = X = 9.0\,\text{V}$ [1]

4 **a** Current in resistor $I = \frac{1.8}{220} = 8.2\,\text{mA}$ [1]

b p.d. *across* lamp $V = 6.0 - 1.8 = 4.2\,\text{V}$ [1]

c Resistance $= \frac{V}{I} = \frac{6.0}{0.0082} = 730\,\Omega$ [1]

d Charge $= It = 0.0082 \times 60$ [1]

$= 0.492\,\text{C}$ [1]

Number of electrons $= \frac{Q}{e} = \frac{0.492}{1.6 \times 10^{-19}}$ [1]

$= 3.1 \times 10^{18}$ [1]

5 **a** Current in Y $= 2.0 - 0.5 = 1.5\,\text{A}$ [1]

b p.d. across Y $= 0.5 \times 6.0 = 3.0\,\text{V}$ [1]

so resistance of Y $= \frac{3.0}{1.5} = 2.0\,\Omega$ [1]

c p.d across X $= 12 - 3.0 = 9.0\,\text{V}$ [1]

so resistance of X $= \frac{9.0}{2.0} = 4.5\,\Omega$ [1]

6 **a** The potential difference across the terminals of a battery is the work done per coulomb of charge in the external circuit. [1]

The e.m.f. of a battery is the work done per coulomb of charge in the complete circuit. [1]

b **i** 0.75 A [1]

ii $V = IR = 0.75 \times 12$ [1]

$= 9.0\,\text{V}$ [1]

iii Use the circuit loop including both batteries and the 3 Ω resistor:

$9 = E_2 + (1 \times 3)$ [1]

$E_2 = 6\,\text{V}$ [1]

iv $I = \frac{V}{R} = \frac{6}{12}$ [1]

$= 0.5\,\text{A}$ [1]

7 **a** The ammeter goes in the main circuit. [1]

It must have low resistance so little energy is lost / small p.d. across it. [1]

b **i** Resistance of combined voltmeter and 400 Ω resistor $= \left(\frac{1}{1200} + \frac{1}{400}\right)^{-1}$ [1]

resistance $= 300\,\Omega$ [1]

Current in the circuit $= \frac{9.0}{300} = 0.030\,\text{A}$ [1]

Potential drop across 100 Ω resistor $= 0.030 \times 100 = 3.0\,\text{V}$, therefore

e.m.f. $= 9.0 + 3.0 = 12.0\,\text{V}$ [1]

ii New resistance of combination $= \left(\frac{1}{12\,000} + \frac{1}{400}\right)^{-1} = 387\,\Omega$ [1]

New current $= \frac{12.0}{487} = 0.0246\,\text{A}$ [1]

Potential drop across combination $= 387 \times 0.0246 = 9.53 \approx 9.5\,\text{V}$ [1]

iii The voltmeter is in parallel with the main circuit [1]

so it reduces the resistance of any combination it is in, as shown in answers **b i** and **b ii** above. [1]

8 **a** Resistance is the potential difference across a component divided by the current through it. [1]

b Resistance of upper arm $= (40 + 20) = 60\,\Omega$ [1]

Resistance of lower arm $= \left(\frac{1}{60} + \frac{1}{40}\right)^{-1} + 96$

$= 24 + 96 = 120\,\Omega$ [1]

Resistance of network $= \left(\frac{1}{60} + \frac{1}{120}\right)^{-1} = 40\,\Omega$ [1]

c Total potential difference across whole lower arm $= 6.0\,\text{V}$ [1]

p.d. across parallel part $= \frac{24}{120} \times 6.0 = 1.2\,\text{V}$ [1]

Current through 60 Ω resistor $= \frac{1.2}{60} = 0.02\,\text{A}$ [1]

Answers to EOC questions

Chapter 11

1 a i When p.d. = 2.0 V, current $I = 0.25$ A [1]
so resistance $R = \frac{2.0}{0.25} = 8\,\Omega$ [1]

ii resistance = $\frac{5.0}{0.5} = 10\,\Omega$ [2]

b A filament lamp [1]

2 a Graph showing a current greater than zero at 0 °C, with a positive gradient; it may or may not be linear. [1]

b Use the graph as a calibration graph. Keeping the voltage across the thermistor constant, place the thermistor at the point where the temperature is to be measured. [1]
Read the current and convert to a temperature using the calibration graph. [1]

3 a The number density of free electrons is very high in copper. [1]
In silicon the number density of free electrons is very much less (a million times). [1]

b In a metallic conductor such as copper, the vibration of the ions increases their effective cross-section to the migrating electrons. [1]
The higher the temperature, the more vibration, hence the greater the effective cross-section and the more collisions there are between the electrons and the ions. This reduces the mean drift velocity. [1]
In semiconductors, thermal energy gives electrons sufficient energy to escape from their parent atoms. [1]
The greater the temperature, the greater the number of electrons that can escape, so the more charge carriers there are and the lower the resistance. [1]

4 a Resistance = $\frac{\rho l}{A} = \frac{1.3 \times 10^{-8} \times 1.5}{0.008 \times 10^{-6}}$ [1]
$= 2.4\,\Omega$ [1]

b $l = R\frac{A}{\rho} = \frac{30 \times 8.0 \times 10^{-9}}{1.3 \times 10^{-8}}$ [1]
$= 18.5 \approx 18$ m [1]

5 a $V = IR = 0.48 \times 5$ [1]
$= 2.4$ V [1]

b i Current = 0.72 – 0.48 = 0.24 A [1]

ii 0.24 A [1]

c Resistance of the thermistor decreases [1]
so circuit resistance decreases [1]
so ammeter reading increases. [1]

6 a The atoms vibrate more [1]
so their effective cross-sectional area increases and there are more collisions. [1]

b Cross-sectional area [1]
Material of which the wire is made [1]

c i $R = \frac{V}{I} = \frac{1.5}{0.24}$ [1]
$= 6.25\,\Omega$ [1]

ii $R = \frac{\rho l}{A}$ so $A = \frac{1.69 \times 10^{-8} \times 5}{6.25}$ [1]
$A = 1.35 \times 10^{-8}$ m^2 [1]
$d = \sqrt{\frac{4A}{\pi}} = 1.3 \times 10^{-4}$ m [1]

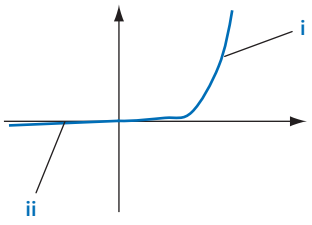
d Less area open to air [1]
Less cooling, therefore temperature rises. [1]
Current decreases. [1]

7 a $R = \frac{\rho l}{A}$ and A = thickness × width, so $t = \frac{\rho l}{Rw}$ [1]
$t = \frac{2.3 \times 10^{3} \times 36 \times 10^{-3}}{1.1 \times 10^{6} \times 32 \times 10^{-3}}$ [1]
$= 0.0023$ m (2.3 mm) [1]

b $I = \frac{V}{R} = \frac{12}{1.1 \times 10^{6}}$ [1]
$= 1.1 \times 10^{-5}$ A [1]

c Resistance would decrease. [1]
Current would increase. [1]
Silicon wafer would get even hotter, with avalanche effect creating more and more charge carriers. [1]

8 a



b Diode has very low resistance (in forward direction), so large current. [1]
Heating would lead to damage to diode / supply. [1]

c $R = \frac{V}{I} = \frac{1.4}{20 \times 10^{-3}}$ [1]
$= 70\,\Omega$ [1]

9 **a** The current through the resistor is proportional to the p.d. across it [1]
at constant temperature. [1]

b **i** Straight horizontal line (assuming V is on x-axis) [1]

ii As above, but at a higher resistance. [1]

iii Both at a (slightly) higher level. [1]

c $R \propto$ length:
$2 \times$ length $\rightarrow 2 \times$ resistance [1]
$R \propto \frac{1}{\text{cross-sectional area}}$:
$2 \times$ diameter $\rightarrow 4 \times$ area $\rightarrow \frac{1}{4} \times$ resistance [1]
new resistance $= 2 \times \frac{1}{4} = \frac{1}{2} \times$ old resistance [1]

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Answers to EOC questions

Chapter 12

1 a Terminal p.d. = $2.5 \times 0.30 = 0.75\,\text{V}$ [1]
There is work done inside the cell against the internal resistance **or** there is a voltage (lost volts) across the internal resistance. [1]

b $E = V + Ir$ [1]
$1.5 = 0.75 + (2.5 \times r)$ [1]
$2.5r = 0.75$, so $r = 0.30\,\Omega$ [1]

c i Power $P = I^2R = 2.5^2 \times 0.30 = 1.875 \approx 1.88\,\text{W}$ [1]

ii Power for 0.5 Ω:
total resistance $R + r = 0.80\,\Omega$;
current $= \frac{1.5}{0.80} = 1.875\,\text{A}$;
power $= 1.875^2 \times 0.50 = 1.76\,\text{W}$ [2]
Power for 0.2 W:
total resistance $R + r = 0.50\,\Omega$;
current $= \frac{1.5}{0.580} = 3.0\,\text{A}$;
power $= 3.0^2 \times 0.50 = 0.45\,\text{W}$
Both are less than 1.88 W. [2]

2 a i The test cell is the wrong way round [1]
so he must reverse it. [1]

ii At the balance point, the ammeter reading is zero. [1]

b $\frac{\text{e.m.f.}}{1.434} = \frac{22.5}{33.6}$ so e.m.f. $= 1.434 \times \frac{22.5}{33.6}$ [1]
$= 0.933\,\text{V}$ [1]

3 a The resistance due to the work done in driving current through the cell [1]
equals the 'lost volts' / current. [1]

b i $E = I(R + r)$ [1]
$= 0.625(2 + r) = 0.341(4 + r)$ [1]
$r = 0.40\,\Omega$ [1]

ii Substitution into $E = I(R + r)$ so $E = 1.50\,\text{V}$ [1]

c Internal resistance is too high. [1]
Maximum current < 4 A [1]

4 a The e.m.f. of a cell is the work done per coulomb of charge [1]
in the complete circuit. [1]

b When there is no / negligible current through the cell, there is no potential drop across the internal resistance. [1]
When there is a current, there is such a potential drop. [1]

c i $I = \frac{V}{R} = \frac{8.40}{12}$ [1]
$= 0.70\,\text{A}$ [1]

ii Lost volts = 0.54 V [1]
$R = \frac{V}{I} = \frac{0.54}{0.70} = 0.77\,\Omega$ [1]

iii The resistance of the voltmeter $\gg r$ or R. [1]

5 a In circuit 1, the p.d. across the bulb varies from 0 to 240 V. [1]
In circuit 2, it never falls to zero. [1]

b i $P = \frac{V^2}{R}$, leading to $R = \frac{240^2}{60}$ [1]
$R = 960\,\text{W}$ [1]

ii Resistance greater when switched on. **or** Resistance is lower at room temperature. [1]
Resistance of a metal increases with increasing temperature (**or** decreases with decreasing temperature). [1]

6 a Voltmeter reading will decrease [1]
because current through R_2 decreases. [1]

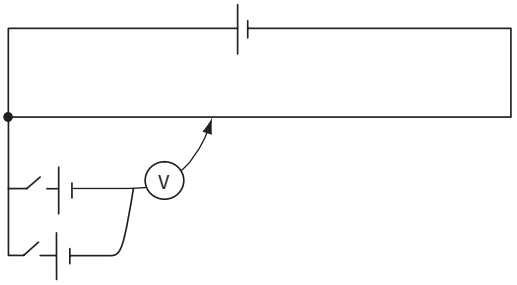
b $V_{\text{out}} = \left(\frac{R_2}{R_1 + R_2}\right) V_{\text{in}}$ [1]
$R_2 = 1645\,\Omega \approx 1600\,\Omega$ [1]

c Resistance of R_2 and voltmeter in parallel
$= \left(\frac{1}{470} + \frac{1}{2000}\right)^{-1} = 380\,\Omega$ [1]
$V_{\text{out}} = \left(\frac{380}{1645 + 380}\right) \times 9.0 = 1.69 \approx 1.7\,\text{V}$ [1]

7 a i Straight line through origin with positive gradient [1]
Graph axes labelled V (x-axis) and I (y-axis) [1]

ii A 0 V, B 2.2 V [1]

iii General diagram (with one or two cells) [1]



Two cells in correct polarity [1]
Switches, **or** suitable comment indicating that only one cell is used at a time [1]

b $E = I(R + r)$ [1]
$E = 0.60 \times (8 + r) = 1.50 \times (2 + r)$ [1]
$r = 2\,\Omega$ [1]
Substitution into either equation gives
$E = 6.0\,\text{V}$ [1]

Answers to EOC questions

Chapter 13

1 Two waves occupy 4.8 divisions,
so 1 wave = 2.4 div [1]
Time for 2.4 div = 2.4 × 500 μs = 1200 μs [1]
Frequency = $\frac{1}{\text{period}} = \frac{1}{1200 \times 10^{-6}}$
= 833 Hz
≈ 830 Hz [1]

2 **a** They travel through a vacuum. [1]
Their speed in vacuum is 3×10^8 m s^{-1}. [1]

b Microwaves have a shorter wavelength (**or** higher frequency) than radio waves. [1]

c **i** Between 10^{-8} and 10^{-13} m [1]

ii Using $c = f\lambda$, frequency lies between 10^{16} and 10^{21} Hz [1]

3 **a** Doppler effect – source moving towards / away from observer leads to decreased / increased wavelength. [1]

b Maximum frequency when boat's velocity is directed towards the observer. [1]
Observed frequency = $\frac{420 \times 330}{305}$ = 390 Hz [1]
Minimum frequency when boat's velocity is directed away from observer. [1]
Observed frequency = $\frac{420 \times 330}{355}$ = 390 Hz [1]

c When boat's velocity is directed at the student. [1]

4 **a**



i Any C correctly marked. [1]

ii Any R correctly marked. [1]

b Vibrates [1]
parallel to the direction of oscillation. [1]

c Moves from the equilibrium position to maximum displacement, back to equilibrium, then to maximum displacement in opposite direction and back to equilibrium. [1]
240 times per second. [1]

d $v = f\lambda$ leading to $\lambda = \frac{320}{240}$ [1]
λ = 1.3 m [1]

5 **a** Wave transmitted by vibration of electric and magnetic fields [1]
at right angles to the direction of propagation. [1]

b **i** Intensity of light from star A is $\frac{1}{4}$ that from B [1]
Intensity is proportional to $\frac{1}{r^2}$ [1]

ii Intensity ∝ amplitude^2 so amplitude ∝ $\sqrt{\text{intensity}}$ [1]
Amplitude of signal from A is $\sqrt{\frac{1}{4}} = \frac{1}{2}$ that of B. [1]

c $v = f\lambda$ leading to $f = \frac{3 \times 10^8}{7.5 \times 10^{-8}}$ [1]
$f = 4 \times 10^{15}$ Hz [1]

6 **a** Particles vibrate parallel to direction of propagation. [1]
Perpendicular displacement of beam on screen relates to [1]
parallel displacement of the particle. [1]

b 5 waves in $6 \times 5 \times 10^{-3}$ s
so period = 6×10^{-3} s [1]
$f = \frac{1}{T} = 167 \approx 170$ Hz [1]

c $v = f\lambda$ leading to $v = 167 \times 1.98$ [1]
v = 330 m s^{-1} [1]

7 **a** Equal [1]
Same time for one wave and thus same frequency [1]

b Waves where the peaks or troughs do not occur at the same time have a phase difference. [1]
One wave occurs in about 2.2 horizontal divisions and the waves are separated by 0.3 div. [1]
Phase difference = $\frac{0.3}{2.2 \times 360} = 50° \pm 20°$ [1]

c Amplitude of upper wave is $\frac{2}{3}$ of lower wave. [1]
Since intensity ∝ amplitude^2, ratio of intensity of upper wave to lower wave = $\frac{4}{9}$ [1]

8 **a** Change in frequency [1]
caused by a relative motion between source and observer. [1]

b Observed frequency = $\frac{fc}{(v+c)}$
$= \frac{6.5 \times 10^{14} \times 3.0 \times 10^8}{(3.0 \times 10^8 + 6.4 \times 10^5)}$ [1]
$= 6.486 \times 10^{14}$ (Hz) [1]

Cambridge International AS and A Level Physics © Cambridge University Press 2014

change in frequency = $(6.500 - 6.486) \times 10^{14}$
$= 1.4 \times 10^{12}$ Hz [1]

c i Away from Earth [1]
Red colour means longer wavelength and smaller frequency so Doppler effect indicates star is moving away. [1]

ii Further away the faster the stars are moving away [1]

9 a i Lower [1]

ii In the time between emitting one wave and the next, the source moves away. [1]
The wavelength is larger or waves appear further apart. [1]
Since $v = f\lambda$ a larger wavelength is a smaller frequency. [1]

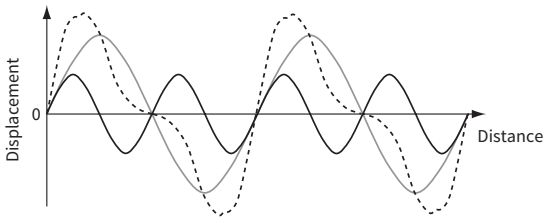
b Observed frequency = $\frac{fc}{(v+c)} = \frac{4.00 \times 1500}{(1500+30)}$ [1]
$= 3.9216$ (MHz) [1]
Shift in frequency = 0.078 MHz = 78 000 Hz [1]

c The Doppler effect occurs when the observer (the particles) moves away from the source (the transmitter). [1]

Answers to EOC questions

Chapter 14

1 a



The dashed line represents the resultant wave. [2]
(Your diagram should show a good attempt to sum the two waves.)

b Wavelength is the same as that of the longer wave. [1]

2 a More rounded [1]

b Even flatter [1]

3 Radio waves have a long enough wavelength, up to 1 km, that they can diffract round the hills. [1]
TV waves have very short wavelength (centimetres or millimetres), so cannot diffract round the hills. [1]

4 Damita is correct; the sound from the speakers will have many different frequencies and cannot be coherent. [2]

5 Using $ax = \lambda d$ [1]
wavelength $\lambda = \frac{ax}{d} = \frac{1.5 \times 1.2}{8.0} = 0.225\,\text{m} \approx 0.23\,\text{m}$ [1]

6 When the waves are in phase, they add up to give loud sound. [1]
They gradually go out of phase, and when they are in antiphase the sound is at its quietest. [1]
The waves gradually come back into phase and become loud again. [1]

7 Separation of slits $d = \frac{1}{5000} = 2.0 \times 10^{-4}\,\text{cm}$
$= 2.0 \times 10^{-6}\,\text{m}$ [1]
First maximum when $n = 1$:
$\sin q = \frac{\lambda}{d} = \frac{656 \times 10^{-9}}{2.0 \times 10^{-6}} = 0.328$ [1]
So $t = \sin^{-1} 0.328 = 19.1°$ [1]
Second maximum when $n = 2$:
$\sin q = \frac{2\lambda}{d} = \frac{2 \times 656 \times 10^{-9}}{2.0 \times 10^{-6}} = 0.656$ [1]
So $t = \sin^{-1} 0.656 = 41.0°$ [1]

8 a Superposition is the algebraic summing of the displacements [1]
of two (or more) waves. [1]

b $\lambda = \frac{ax}{D}$ leading to
$a = \frac{\lambda D}{x} = \frac{590 \times 10^{-9} \times 1.8 \times 12}{16.8 \times 10^{-3}}$ [1]
$a = 6.3 \times 10^{-4}\,\text{m}$ [1]

c i More fringes seen on screen **or** fringe brightness decreases less from middle to edge of screen [1]
less bright [1]

ii Fringes wider / farther apart [1]
same brightness [1]

9 a Coherent: constant phase difference [1]
Monochromatic: very small or no range of wavelengths / frequencies [1]

b i First order produced by waves with path difference of one wavelength [1]
Second order produced by waves with path difference of two wavelengths [1]

ii *Any two from:*
lines at end A are further apart **or** lines at end B are closer together [1]
lines at end A are thinner **or** lines at end B are wider [1]
lines at end A are brighter **or** lines at end B are dimmer [1]

iii $n\lambda = d \sin q$ [1]
$n = 1$, leading to $\lambda = \frac{\sin 19.5°}{5000 \times 10^{2}}$ [1]
$\lambda = 6.68 \times 10^{-7} \approx 6.7 \times 10^{-7}\,\text{m}$ [1]

iv $\sin q = \frac{n\lambda}{D}$ leading to
$\sin q = 2 \times 6.68 \times 10^{-7} \times 5000 \times 10^{2}$ [1]
$q = 41.8 \approx 42°$ [1]

10 a The superposition of two waves 180° out of phase [1]
to give (nearly) zero resultant. [1]

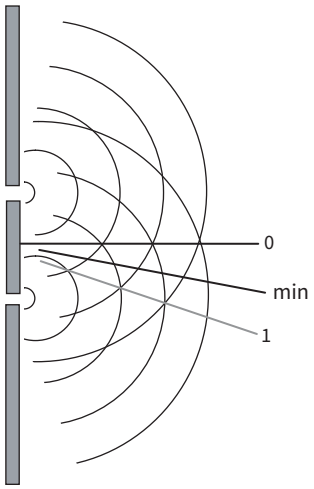
b $\lambda = \frac{ax}{D}$ leading to $x = \frac{D\lambda}{a} = \frac{1.2 \times 1.5 \times 10^{-2}}{12.5 \times 10^{-2}}$ [1]

Number of fringes in 45 cm distance
$= \frac{45 \times 10^{-2}}{x} = 3.125$ [1]
three maxima [1]

c $c = f\lambda$ leading to $f = \frac{3 \times 10^{8}}{1.5 \times 10^{-2}}$ [1]
$f = 2.0 \times 10^{10}$ Hz [1]

11 **a** Spreading out of a wave after passing through a gap in a barrier [1]
or around an object. [1]

b **i**, **ii** and **iii**



One mark for each line (the min line and the 1 line can be above the central line) [3]

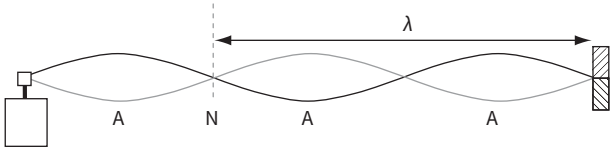
c $\lambda = \frac{ax}{D}$ leading to $= \frac{18 \times 10^{-2} \times 12 \times 10^{-2}}{60 \times 10^{-2}}$ [1]
$= 3.6 \times 10^{-2}$ m [1]

Answers to EOC questions

Chapter 15

1 a Node and antinode marked [1]

b Wavelength marked [1]



c There would be double the number of loops (6). [1]

2 a Resonance is when the frequency of one source of vibration coincides with the natural frequency of vibration of a body, causing the body to vibrate with a large amplitude. [1]

b $\frac{\lambda}{4} = 0.312\,\text{m}$ [1]

$v = f\lambda = 256 \times 4 \times 0.312 = 319 \approx 320\,\text{m s}^{-1}$ [1]

3 a Similarities:

Points in the waves vibrate. [1]

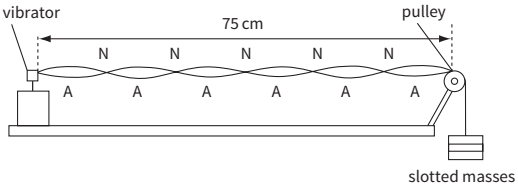
The wave speed = $f\lambda$ [1]

Differences:

A progressive wave transfers energy; a standing waves does not transfer energy. [1]

Points in a progressive wave have different amplitudes of vibration with respect to time; points of a standing wave have the same amplitude with respect to time. [1]

b i



Any one node and any one antinode shown. [2]

ii Wavelength = $\frac{75}{3}$ cm [1]

$c = f\lambda = \frac{120 \times 0.75}{3}$ (error in λ carry forward) [1]

$c = 30\,\text{m s}^{-1}$ [1]

c Speed of the waves down the string changes (with tension) [1]

so wavelength changes. [1]

4 a i Vibrating [1]

back and forth parallel to the tube. [1]

ii Stationary [1]

b i $\frac{3}{4}\lambda = 59 + e$ and $\frac{5}{4\lambda} = 99 + e$ [1]

leading to $\lambda = 80.8\,\text{cm}$ [1]

$v = f\lambda = 400 \times 80.8 \times 10^{-2}$ [1]

$v = 323 \approx 320\,\text{m s}^{-1}$ [1]

ii $e = (\frac{3}{4} \times 80.8) - 59 = 1.6\,\text{cm}$ [1]

5 a i The waves have a constant phase difference [1]

over a period of time. [1]

ii The amount by which one wave leads or lags another [1]

Expressed using phase angle [1]

b i Maxima occur when the reflected waves are exactly in phase with the incident wave. [1]

Minima occur when the reflected waves are 180° out of phase with the incident wave. [1]

ii Each time the wave amplitude falls to zero, the plate has moved through $\frac{1}{2}\lambda$ [1]

$\lambda = \frac{42.0}{2.5}$ [1]

$= 16.8\,\text{cm}$ [1]

$c = f\lambda$ leading to $f = \frac{3 \times 10^8}{16.8 \times 10^{-2}}$ [1]

$= 1.78 \times 10^9 \approx 1.8 \times 10^9\,\text{Hz}$ [1]

c Phase shift of 180° [1]

on reflection [1]

6 a $\frac{7}{4} \times \lambda = 90\,\text{cm}$ so $\lambda = 51.4\,\text{cm}$ [1]

$c = f\lambda = 512 \times 51.4 \times 10^{-2}$ [1]

$= 263\,\text{m s}^{-1}$ [1]

b Q no movement [1]

P, R, S and T vibrate from side to side parallel to the tube [1]

P and S have largest amplitude [1]

c Any two correct points [1]

Answers to EOC questions

Chapter 16

1 The plum pudding model predicts that there will be no large-angle scattering events. An α-particle is more massive than an electron and will not be deviated by an electron. [1]
The actual experiment shows a few α-particles scattered through more than 90°. [1]
These α-particles must have interacted with a charged object more massive than an electron. This object, containing most of the mass of the atom, was called the nucleus. [1]

2 The nucleus contains 38 protons and 52 neutrons. [1]

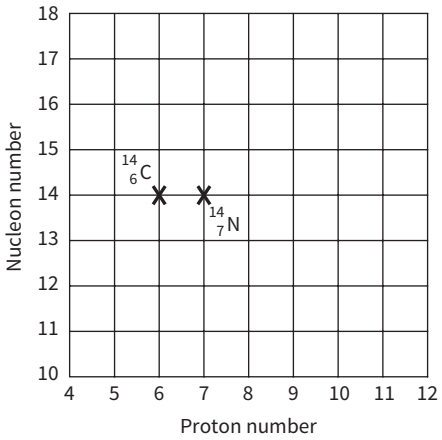
3 When the nucleus ejects an α-particle, it emits two protons and two neutrons, [1]
so the nucleon number decreases by 4 ($A-4$) and the proton number (atomic number) decreases by 2 ($Z-2$). [1]
The nucleus then ejects two β^--particles, formed by decay of two neutrons to protons. [1]
The nucleon number (mass number) remains at $A-4$, but the proton number now increases by 2, returning to Z. [2]

4 $^{131}_{53}\text{I} \rightarrow {}^{131}_{54}\text{X} + {}^{0}_{-1}\text{e} + \bar{\nu}$ [3]

5 **a** β^--particles are fast-moving electrons that come from the nucleus. [1]

b $^{14}_{6}\text{C} \rightarrow {}^{14}_{7}\text{N} + {}^{0}_{-1}\text{e} + \bar{\nu}$
Proton numbers correct [1]
Nucleon numbers and all symbols correct [1]

c **i** Graph and axes drawn correctly and labelled [1]
Carbon-14 marked correctly [1]



ii Nitrogen-14 marked correctly [1]

6 **a** An α-particle contains two protons and two neutrons. [1]
A β^--particle is a fast-moving electron. [1]
An α-particle is much heavier than a β^--particle (about 7400 **or** 8000 times larger). [1]
An α-particle is positively charged ($+2e$), whereas a β^--particle is negatively charged ($-e$). [1]

b Isotopes have the same number of protons in the nucleus **or** the same proton number. [1]
Isotopes have different numbers of neutrons in the nucleus **or** different nucleon numbers. [1]

c $^{236}_{92}\text{U} \rightarrow {}^{236}_{90}\text{X} + {}^{4}_{2}\text{He}$
or
$^{236}_{Z}\text{U} \rightarrow {}^{236}_{Z-2}\text{X} + {}^{4}_{2}\text{He}$
All three nucleon numbers correct [1]
All three proton numbers correct [1]

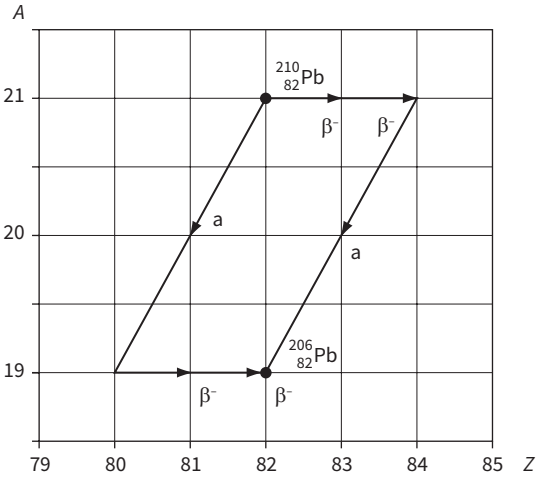
7 **a** Volume of a sphere = $\frac{4}{3}\pi r^3$;
ratio of volumes = $\frac{(10^{-10})^3}{(10^{-15})^3}$ [1]
ratio = 10^{15} [1]

b Assuming that most of the mass of the atom is concentrated in the nucleus, [1]
then 19 000 kg of gold atoms has a volume of $1\,m^3$; from **a** the volume of the nucleus with this mass is $1 \times 10^{-15}\,m^3$ [1]
Density of gold nucleus
$= \frac{1.9 \times 10^4}{10^{-15}\,m^3} = 1.9 \times 10^{19}\,kg\,m^{-3}$ [1]

8 **a** 82 protons [1]
128 neutrons [1]

b Because the particles are charged, when they collide with or pass close to atoms, [1]
they knock electrons from the atoms, leaving behind charged atoms or ions. [1]

c Two solutions, each with two β^--decays and one α-decay. [1]
Each β^--decay correct [1]
Each α-decay correct [1]



9 **a** Most of the mass is concentrated in a small space. [1]
This space / nucleus is positive and surrounded by negative charged electrons. [1]

b *Any three from:*
Most α-particles pass undeviated through the gold, showing that most of an atom is empty space [1]
Some α-particles are deflected though large angles / greater than 90°, showing that
Part of the atom is much heavier than an α-particle (and positive) [1]
Electrons (negative) were known to exist, so other particles in the nucleus must have been positive and heavier [1]
The relative number of large deflections showed the size of the nucleus [1]

10 **a** **i** $\rightarrow {}^{0}_{-1}e + {}^{46}_{21}Sc + \bar{\nu}$ [1]

ii $\rightarrow {}^{0}_{+1}e + {}^{23}_{11}Na + \nu$ [1]

b **i** A down quark turns into an up quark, an electron and an antineutrino [1]

ii An up quark turns into a down quark, a positron and a neutrino [1]

c the weak nuclear force [1]

11 **a** A neutron is made up of other particles but a quark is thought not to. [1]

b **i** proton charge +1(*e*), made up of two up quarks and one down quark [1]
neutron charge 0, made up of two down quarks and one up quark [1]

ii The proton is positive, so either one or both up and down quarks are positive. [1]
The neutron is neutral, so if one quark is positive, the other is negative. [1]

c strong interaction or strong nuclear force [1]

d ${}^{1}_{0}n \rightarrow {}^{1}_{1}p + {}^{0}_{-1}e + \bar{\nu}$ [2]

e electron – lepton, neutrino (or antielectron, positron, antineutrino) [1]
neutron – hadron or baryon, antineutron, antiproton or a variety of mesons [1]

Answers to EOC questions

Chapter P1

1 **a** 25.8 ± 3.3 cm **or** better 26 ± 3 cm [1]

b 2.80 ± 0.18 cm **or** better 2.8 ± 0.2 cm [1]

c 24.0 ± 0.1 g [1]

2 **a** 1.1 ± 0.1 V; 0.7 ± 0.1 V [1]

b If $R = kl$ then the two values of k are 0.0440 Ω cm^{-1} and 0.0467 Ω cm^{-1}. [1]
The criterion is that, if the resistance is proportional to length, then the percentage difference in the values of k is less than the percentage uncertainty in the value of V. [1]
Percentage difference in k values $= \frac{0.27}{0.044} = 0.6\%$ [1]
Percentage uncertainty in lowest value of V is 14%, so the data is consistent with R proportional to l. [1]

3 Apart from taking only two readings, which is not enough, it is difficult to:

- release the ball without giving it an initial velocity [1]
- start the stopwatch as the ball is released [1]
- stop the stopwatch exactly as the ball crosses the line [1]
- release the ball exactly on the line [1]
- measure small times that are not much larger than reaction times [1]
- avoid the balling rolling to the side of the track [1]
- measure the distance without introducing parallax error [1]

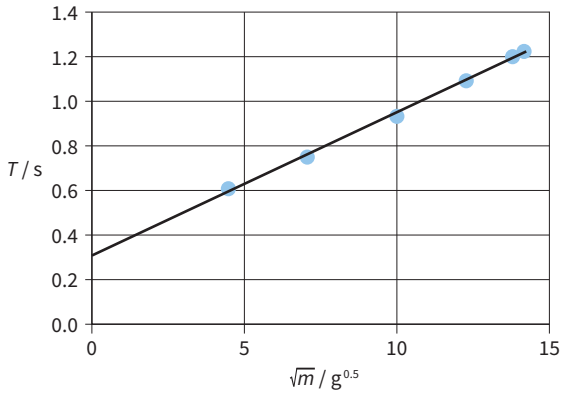
Suggested improvements

- take readings at about five values of s and
- plot a graph of T^2 against s [1]
- take a video of the ball rolling down the ramp with a timer in the background and play it back frame by frame [1]
- use a metal ball, held to an electromagnet at the same start position each time, turning off the magnet to start the ball [1]
- use longer distances [1]
- have a groove down the middle of the track [1]
- always view from directly above the mark on the track when using a rule [1]

(Maximum [8] marks: [4] for the problems, and the other [4] for solutions / improvements.)

4 **a** Values for $\sqrt{m}$ correct and to the same number of significant figures, or one more than data [1]
Values for T correct and to the same number of significant figures, or one more than data [1]

Mass / g	T_{20} / s	$\sqrt{m}$ / g$^{0.5}$	T / s
20	12.2	4.5	0.610
50	15.0	7.1	0.750
100	18.7	10.0	0.935
150	21.8	12.2	1.090
200	24.5	14.1	1.225
190	24.0	13.8	1.200



b Sensible axes used and labelled with quantity and unit [1]
All points plotted to within half a small square [1]
All points close to line [1]
Line of best fit correct [1]

c Gradient drawn with more than half length of drawn line; value between 0.062 and 0.064 [1]
y-intercept between 0.30 and 0.32 [1]

d C = value given for y-intercept with unit s, e.g. 0.31 s [1]
k = value given for gradient with unit, e.g. 0.063 s g$^{-0.5}$ [1]

5 a i 127 cm [1]

ii 3 cm [1]
2% [1]

iii 4.7 s [1]

iv 0.1 s [1]
2% [1]

b i When $l = 40$ cm, $v = 27$ cm s^{-1}; when $l = 60$ cm, $v = 38$ cm s^{-1} [1]

ii Two significant figures, as this was the lower of the number of significant figures for s (3 sig. figs) and t (2 sig. figs) [1]

c i, ii $k = \frac{s}{l} = 3.175$ and 3.167. The percentage difference in k values is 0.26%. [1]
This is less than the percentage uncertainty in s = 2%, so the data is consistent with s proportional to l. [1]

d, e See comments for experiment 2, plus the difficulty of measuring to the middle of the trolley for l and s (measure from end of trolley and add on half the length of a trolley found separately), difficulties that may occur as the toy car hits the straight section (have a slightly curved join), and difficulties with the car not running straight (have wheels in grooves) [8]
(Marks: [1] for each sensible source of uncertainty, to a maximum of [4];
and [1] for each sensible improvement, to a maximum of [4])

6 a Average value of voltage was 5.9 and uncertainty is ±0.2 **or** possibly 0.3 V. Percentage uncertainty is between 3% and 5%. [1]

b i 0.42 °C V^{-2} [1]
and 0.35 °C V^{-2} [1]

ii Two significant figures, as V was measured to 2 sig. figs and q was measured to 3 sig. figs. Choose the lower number. [1]

iii Percentage difference in values of k is 18%. This is larger than the 5% uncertainty in V (which is larger than the percentage uncertainty in the value of θ). Thus the data is not consistent with the suggested relationship. [1]

c Only two voltage readings is not enough. [1]
The voltmeter reading fluctuates because of contact resistance. [1]
The voltmeter reading fluctuates because of a change in mains voltage. [1]
The voltmeter reading falls because the resistance of the wire changes as it heats up. [1]
The temperature increases after the switch is turned off (while heat passes from the resistor to the water). [1]
(Maximum of [4] marks)

d Clean the wires first. [1]
Use a car battery. [1]
Use a metal whose resistance changes little with temperature. [1]
Use more than two voltages and plot a graph of θ against V^2 **or** calculate more than two values of k. [1]
Stir well and take the highest reading on the thermometer after switching off. [1]
(Maximum of [4] marks)

Answers to EOC questions

Chapter 17

1 a The angle subtended at the centre of a circle by an arc equal in length to the radius. [1]
(accept: a radian = $\frac{\text{length of arc}}{\text{radius}}$)

b angular velocity = $\frac{2\pi}{15}$ [1]
$= 0.42\ \text{rad s}^{-1}$ [1]

2 a At the position shown, the weight of the truck. [1]

b acceleration of truck = $g = \frac{v^2}{r}$ [1]
$v^2 = rg = 4.0 \times 9.8 = 39.2\ \text{m}^2\ \text{s}^{-2}$ [1]
$v = 6.3\ \text{m s}^{-1}$ [1]

3 a Centripetal force is the net force acting on an object describing a circle – it is directed towards the centre of the circle. [1]

b i speed = $\frac{\text{distance}}{\text{time}}$ so speed = $\frac{2 \times \pi \times 0.15}{3.0}$ [1]
$\frac{\text{distance}}{\text{time}} = 0.314\ \text{m s}^{-1}$ [1]
$F = \frac{mv^2}{r} = 0.060 \times \frac{0.314^2}{0.15}$ [1]
$= 0.0394 \approx 0.039\ \text{N}$ [1]

ii The centripetal force on the toy increases with its speed. [1]
The toy falls off because the frictional force between the turntable and the toy is not sufficient to provide the centripetal force. [1]

4 a change in potential energy = kinetic energy [1]
$mgh = \frac{1}{2}mv^2$ leading to $v = \sqrt{2gh}$
$= \sqrt{2 \times 9.81 \times 0.70}$ [1]

b centripetal force = $\frac{mv^2}{r} = \frac{0.050 \times 3.712^2}{1.50}$ [1]
$= 0.458\ \text{N}$ [1]
$T - mg = 0.458$ [1]
$T = 0.458 + (0.050 \times 9.81) = 0.95\ \text{N}$ [1]

c The weight is only equal to the tension when the ball is at rest in the vertical position. [1]
The ball is not in equilibrium in the vertical position because it has an upward (centripetal) acceleration. [1]

5 Friction between the tyres and the road provides the centripetal force. [1]
Oil reduces the frictional force, so the car carries on in a wider arc. (The frictional force is not enough to push the car around the desired curve.) [1]

6 a Weight acting vertically downwards, lift force perpendicular to the aeroplane wings. [2]

b (lift = L, angle with horizontal = θ)
vertically: $L \sin\theta = mg$ [1]
horizontally: $L \cos\theta = \frac{mv^2}{r}$ [1]
So $\tan\theta = \frac{(80 \times 9.8)}{75^2} = 0.139$ [1]
$\theta = 7.9°$ [1]

7 a Angle covered per unit time [1]
Angle in radians [1]

b i $T\cos\theta = mg$ leading to $T = \frac{0.200 \times 9.81}{\cos 56°}$ [1]
$= 3.5\ \text{N}$ [1]

ii Centripetal force = $T\sin\theta$
$= 3.5 \sin 56° = 2.9\ \text{N}$ [1]
Substituting in $F = m\omega^2 r$ leads to
$\omega^2 = \frac{2.9}{0.200 \times 0.40} = 36.25$ [1]
$\omega = 6.02 \approx 6.0\ \text{rad s}^{-1}$ [1]

iii Period = $\frac{2\pi}{\omega} = 1.04 \approx 1.0\ \text{s}$ [1]

8 a The force acting towards the centre of a circle [1]
on a body moving round that circle [1]

b i $\frac{mv^2}{r} = mg$ leading to $v^2 = rg = 0.9 \times 9.81$ [1]
$v^2 = 8.829$ [1]
$v = 2.97 \approx 3.0\ \text{m s}^{-1}$ [1]

ii $F = 2mg = 2 \times 5.4 \times 9.81$ [1]
$F = 105.9 \approx 106\ \text{N}$ [1]

9 a radian = $\frac{\text{length of arc}}{\text{radius}}$ [1]

b Pilot feels thrown out from the circle **or** pilot feels pressed outwards against seat. [1]
Inertia means that with no force pilot would travel in a straight line. [1]
Centripetal force pulls the pilot in towards the centre of the circle. [1]

c i $\frac{mv^2}{r} = 6mg$ leading to
$v^2 = 6rg = 6 \times 5.0 \times 9.81$ [1]
$v^2 = 294$ [1]
$v = 17.15 \approx 17.2\ \text{m s}^{-1}$ [1]

ii $\omega = \frac{v}{r} = \frac{17.15}{5}$ [1]
number of revs per second = $\frac{\omega}{2\pi} = 0.54\,s^{-1}$ [1]

d Sharp turns in combat at high speed [1]
mean the pilot will experience high *g* forces. [1]

10 **a** Angle in radians = $\frac{\text{length of arc}}{\text{radius}}$
For a complete circle,
arc length = circumference = $2\pi r$ [1]
angle in radians = $\frac{2\pi r}{r} = 2\pi$ [1]

b **i** $\omega = f \times 2\pi$ leading to 540 rpm = $\frac{540}{60} \times 2\pi$ [1]
$\omega = 56.5\,\text{rad}\,s^{-1}$ [1]

ii $F = mr\omega^2 = 20 \times 10^{-3} \times 0.1 \times 56.5^2$ [1]
$F = 6.4\,N$ [1]

c Gravitational force $\approx 20 \times 10^{-3} \times 10 = 0.2\,N$ [1]
Very much less than the required centripetal force [1]

Answers to EOC questions

Chapter 18

1 $F = \frac{GMm}{r^2}$ [1]

$= \frac{6.67\times10^{-11}\times0.020^2}{0.005^2}$ [1]

$= 1.1\times10^{-9}$ N [1]

2 **a** Arrow vertically downwards labelled 'weight' or 'Earth's gravitational pull' [1]

Arrow to right labelled 'pull of mountain' [1]

b $F = \frac{Gm_1m_2}{r^2} = \frac{6.67\times10^{-11}\times0.020\times3.8\times10^{12}}{1200^2}$ [1]

$= 3.5\times10^{-6}$ N [1]

c Earth's force $F = mg = 0.020\times9.8 = 0.196$ N [1]

which is 5.6×10^4 times as large [1]

3 **a** Arrows towards the Earth [1]

b For a rise of 10 000 m, the Earth's field may be considered to be uniform, [1]

but when something moves a significant distance away from the Earth we must recognise that there is a significant reduction in the field. [1]

4 **a** Mass of Mercury $= \frac{4}{3}\pi r^3 \times$ density [1]

$= 3.286\times10^{23}$ kg [1]

g at surface $= \frac{GM}{r^2}$ [1]

$= 3.68 \approx 3.7$ N kg^{-1} [1]

b weight $= mg = 900\times3.68$ [1]

$= 338 \approx 340$ N [1]

5 Distance from centre of Mars

$= 3.4\times10^6 + 20\times10^6 = 23.4\times10^6$ km [1]

Potential energy $= -\frac{GMm}{r}$ [1]

$= -\frac{6.67\times10^{-11}\times6.4\times10^{23}\times250}{23.4\times10^6}$ [1]

$= -4.6\times10^8$ J [1]

6 **a** $r^3 = \frac{GMT^2}{4\pi^2} = 9.54\times10^{22}$ m^3 [1]

$r = 4.6\times10^7$ m [1]

b The proximity of the very large planet, Jupiter, would disrupt the orbit, through its gravitational pull. [1]

7 **a** Circumference of orbit $= 2\pi r = 2\pi\times1.5\times10^{11}$

$= 9.42\times10^{11}$ m [1]

Speed $= \frac{\text{circumference}}{\text{length of year}} = \frac{9.42\times10^{11}}{365\times24\times3600}$ [1]

$= 2.99\times10^4 \approx 3.0\times10^4$ m s^{-1} [1]

b $a = \frac{v^2}{r} = \frac{(2.99\times10^4)^2}{1.5\times10^{11}}$ [1]

$= 5.96 \approx 6.0\times10^{-3}$ m s^{-2} [1]

c The Sun's gravitational pull provides the centripetal acceleration

so $g = 5.96 \approx 6.0\times10^{-3}$ N kg^{-1} [1]

8 **a** **i** $g = -\frac{GM}{r^2}$ [1]

$g = 3.70 \approx 3.7$ m s^{-2} [1]

ii $\phi = -\frac{GM}{r}$ [1]

$\phi = -1.26\times10^7 \approx 1.3\times10^7$ J kg^{-1} [1]

b 1.3×10^7 J [1]

c Use $v^2 = 2\phi = 2\times1.26\times10^7$ [1]

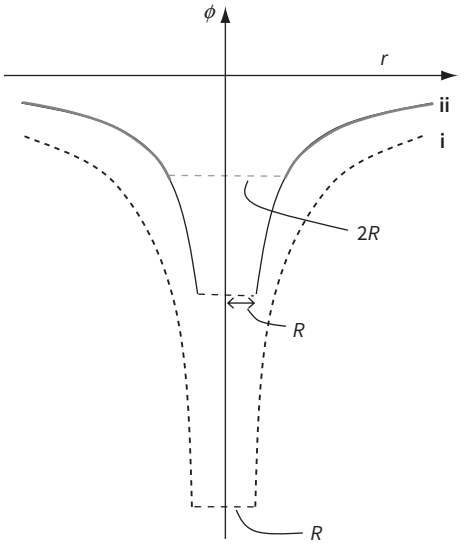
$v = 5019$ m s^{-1} linked to 5000 m s^{-1} [1]

d The work done in lifting the mass of a Mars probe from the Earth is very large [1]

so requires very powerful rocket to do it all in one go. [1]

9 **a** Work done to bring unit mass from infinity to that point [1]

[1]



b **i** Dashed line in diagram: same radius but depth 2 times original [1]

outside original 'well' [1]

ii Bold line on diagram: goes along the present lines [1]

but stops at twice the radius. [1]

c Curve **ii** [1]
Smaller hill to get up [1]

d Using $\phi = -\frac{GM}{r}$ once [1]
Using $\phi = -\frac{GM}{r}$ a second time [1]
(Using $\Delta\phi = -GM\left(\frac{1}{r_2} - \frac{1}{r_1}\right)$ scores both marks)
Correct substitution [1]
Energy required = 1.86×10^6 J [1]

10 **a** Force per unit mass [1]
at the point [1]

b $F = \frac{Gm_1m_2}{r^2}$ [1]
Correct substitution [1]
$F = 3.30 \times 10^{18}$ N [1]

c Development of $T = 2\pi\sqrt{\frac{mr}{F}}$ [1]
Correct substitution [1]
$T = 5.60 \times 10^5$ s [1]

d Otherwise they would lose stability in their orbits (or words to that effect). [1]

11 **a** 0.8 N kg^{-1} (accept range ±0.1) [1]

b 0.8 m s^{-2} (must be numerically equal to **a**) [1]

c $a = \frac{v^2}{r}$ leading to $v^2 = ar$ [1]
$v = 3920 \approx 3900$ m s^{-1} [1]

d **i** Path spiralling into the centre [1]

ii Work done against friction in the atmosphere causes heating [1]
causing satellite to 'burn up'. [1]

Answers to EOC questions

Chapter 19

1 **a** No, constant acceleration under gravity towards ground but not on hitting ground and when hit by player. [2]

b Yes, the restoring force is directed towards a point and is proportional to the displacement from the point. [2]

c No, it travels with constant acceleration towards each plate and then suddenly reverses its direction of velocity on hitting the plate. [1]

d Yes, the restoring force is directed towards a point and is proportional to the displacement from the point. [2]

2 **a** $x = 4\sin 2\pi t$ for x in cm or $x = 4 \times 10^{-2} \sin 2\pi t$ for x in m [2]

b **i** maximum velocity $= 2\pi Af = 2\pi \times 4.0$ [1]

$= 25.1\ \text{cm s}^{-1}$ [1]

ii $v = 21.8\ \text{cm s}^{-1}$ [1]

3 **a** $f = \frac{1}{T} = \frac{1}{0.84} = 1.19\ \text{Hz}$ [1]

b $v_{max} = 2\pi Af = 120\ \text{mm s}^{-1}$ [1]

c $\text{k.e.}_{max} = \frac{1}{2} m(v_{max})^2 = 3.6 \times 10^{-4}\ \text{J}$ [1]

as it passes through the equilibrium position [1]

d $\text{g.p.e.}_{max} = 3.6 \times 10^{-4}\ \text{J}$ (i.e. the same as maximum k.e.) [1]

4 **a** **i** $\frac{1}{2}$ cycle [1]

ii 180° [1]

iii π rad [1]

b **i** $\frac{1}{4}$ cycle [1]

ii 90° [1]

iii $\frac{\pi}{2}$ rad [1]

c **i** $\frac{3}{8}$ cycle [1]

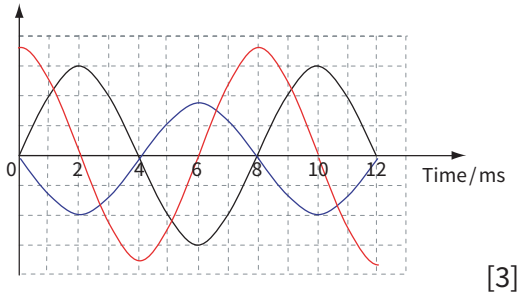
ii 135° [1]

iii $\frac{3\pi}{4}$ rad [1]

5 **a** period $T = 8.0\ \text{ms}$ [1]

frequency $= \frac{1}{T} = 125\ \text{Hz}$ [1]

b Displacement / Velocity / Acceleration



[3]

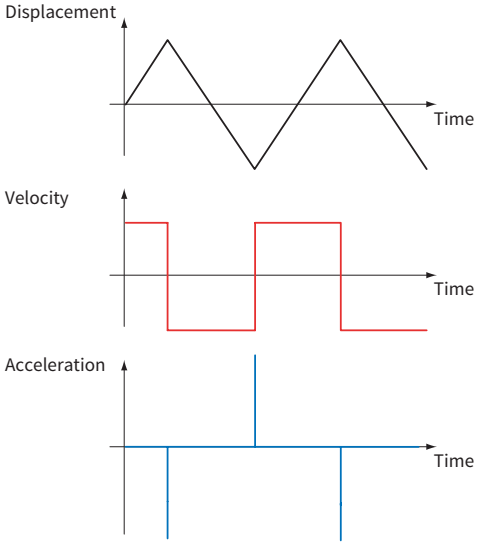
6 **a** No, it is not moving with simple harmonic motion because the displacement does not show a sine or cosine relationship with time.

or

Comment regarding gradient constant for time showing constant velocity, then a sudden change, not in keeping with a force proportional to displacement. [1]

b **i** [3]

ii



[3]

7 **a** Vibration / oscillation in which there is acceleration towards a fixed point. [1]

The magnitude of the acceleration is proportional to the displacement from the point. [1]

b $f = 70\ \text{Hz}$ [1]

c **i** $v = \omega r$ [1]

$= 55\ \text{m s}^{-1}$ [1]

ii $a = \omega^2 r$ [1]

$= 24\,000\ \text{m s}^{-2}$ [1]

iii Use of $F = ma = 5800\ \text{N}$ [1]

8 **a** Waves that are exactly in step with each other / each point on one wave moves in the same way as on the second wave. [1]

b $x = 15\sin(3\pi t)$ [1]

c **i** Angular displacement from equilibrium position = $\omega t = 30° = \frac{\pi}{6}$ rad [1]

$x = 15\sin\left(\frac{\pi}{6}\right)$ [1]

$= 7.5$ cm [1]

ii $v = v_0 \cos \omega t = \omega x_0 \cos \omega t = 15 \times 3\pi \cos\left(\frac{\pi}{6}\right)$ [1]

$v = 120\ \text{cm s}^{-1}$ [1]

iii 60° or $\frac{\pi}{3}$ rad [1]

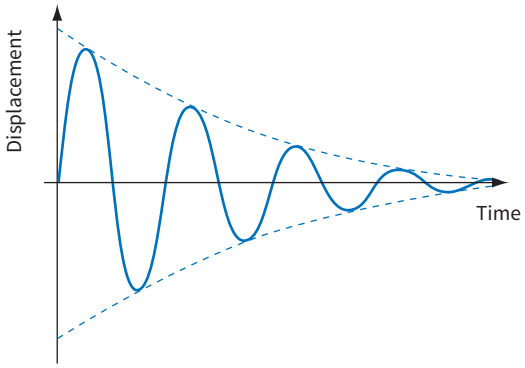
9 **a** Vibration / oscillation in which there is acceleration towards a fixed point. [1]
The magnitude of the acceleration is proportional to the displacement from the point. [1]

b $a = \omega^2 r$ [1]

$a = 400\ \text{m s}^{-2}$ [1]

c $F = ma = 190$ N [1]

d Graph showing sine-type wave, with amplitude getting smaller [1]
frequency remaining the same. [1]



10 **a** If driving frequency = natural frequency [1]
resonance occurs [1]
giving false reading for the shock wave strength. [1]

b Shows that the acceleration is proportional to the displacement [1]
and in the opposite direction to the displacement. [1]

c $\omega^2 = \frac{a}{r}$ [1]

$\omega^2 = 500$ [1]

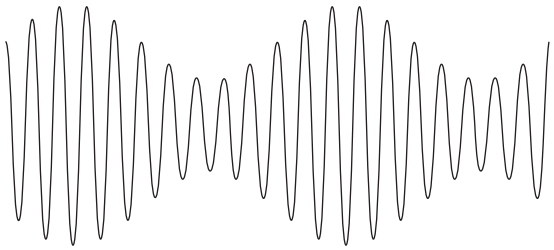
$\omega = 22.3$ [1]

$f = \frac{22.3}{2\pi} = 3.6$ Hz [1]

Answers to EOC questions

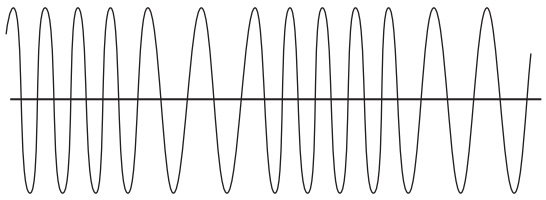
Chapter 20

1 a



The carrier wave has a higher frequency than the signal. [1]
The amplitude of the signal is used to change or modulate the amplitude of the carrier wave. [1]
The height of the carrier wave is made to fall and rise with the actual value of the signal. [1]

b



The carrier wave has a higher frequency than the signal. [1]
The amplitude of the signal is used to alter the frequency of the carrier wave. [1]
Thus the frequency of the transmitted wave is the frequency of the carrier wave, but it varies at any time according to the amplitude of the signal at that time. [1]

2 a i An analogue signal is any continuous signal where the quantity, for example, a voltage, varies in time with the signal. In principle, any small change in the quantity causes a small change in the signal, i.e. the signal can have any value [2]

ii The bandwidth of a signal is the range of frequencies present in a signal. [1]

b The bandwidth of the microphone is less than the range of audible frequencies, which is approximately 20 Hz to 20 kHz. This means that the microphone will not reproduce high frequencies and, for example, music will not be of high quality, although the microphone may be suitable for speech or for a telephone. [2]

3 a Attenuation $= 10 \lg\left(\frac{P_2}{P_1}\right)$ [1]
$= 10 \times \lg\left(\frac{1.26}{1.12}\right) = 0.51\,\text{dB}$ [1]

b Attenuation per unit length $= \frac{0.51}{60}$ [1]
$= 8.53 \times 10^{-3} \approx 8.5 \times 10^{-3}\,\text{dB m}^{-1}$ [1]

4 a Number of decibels $= 10 \lg\left(\frac{P_2}{P_1}\right)$ [1]
$100 = 10 \times \lg\left(\frac{P_2}{6.0 \times 10^{-12}}\right)$ [1]
$P_2 = 6.0 \times 10^{-2}\,\text{W}$ (60 mW) [1]

b Allowed attenuation of optic fibre
$= 100 - 30 = 70\,\text{dB}$ [1]
Length of fibre $= \frac{70}{0.3} = 233 \approx 230\,\text{km}$ [1]

5 a A satellite in geostationary orbit around the Earth takes 24 hours to make an orbit. It travels above the equator, travelling in the same direction as the rotation of the Earth and, from the Earth, appears to be above the same point on the equator. The satellite is at a height of 3.6×10^4 km above the Earth's surface. [2]

b Between 0.001 and 0.300 m [1]

c Advantage:
Satellite dishes do not need to be moved to track the satellite. [1]
Disadvantages:
Communication with some polar regions is not possible. [1]
The satellite is always above the same point on the equator and cannot 'see' other parts of the Earth. [1]
The satellite is higher above the equator and so the signal is delayed longer and is of lower intensity than a polar satellite, and the 'resolution' when viewing the Earth is lower. [1]

6 a The braid in a coaxial cable is the return path for the signal and is usually connected to earth (ground). [1]
By surrounding the central wire, it reduces the amount of external interference received by the inner conductor. [1]

b Increased bandwidth has meant that many telephone calls can be carried on the same line. [1]

c i Noise is the random variation in a received signal not present in the original signal. [1]

Cambridge International AS and A Level Physics © Cambridge University Press 2014

ii Examples of noise sources include: radio emissions from the spark plug of a nearby car, a nearby mobile phone and the random thermal motion of electrons in a wire. [2]

7 Surface waves, frequency below 3 MHz, range up to 1000 km [2]
Space waves, frequency greater than 30 MHz, range is line-of-sight only, except when using satellite and then range is worldwide [2]
Sky waves, frequency between 3 and 30 MHz, range is worldwide by reflection [2]

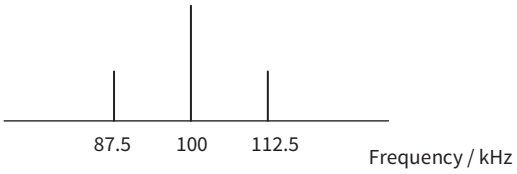
8 **a** Time for one oscillation of the carrier wave $= \frac{80}{8} = 10\,\mu s$ [1]

Frequency $= \frac{1}{T} = \frac{1}{(10 \times 10^{-6})}$
$= 1.0 \times 10^{5}\,Hz$ (100 kHz) [1]

b Time for one oscillation of the audio signal = 80 μs [1]

Frequency $= \frac{1}{T} = \frac{1}{(80 \times 10^{-6})}$
$= 12.5 \times 10^{4}\,Hz \approx 12$ or 13 kHz [1]

c



Frequency spectrum showing central line and two symmetrically placed sidebands [1]
central line at 100 kHz [1]
sidebands at 87.5 and 112.5 kHz [1]

9 **a** Analogue [1]

b

Time / hours	*V* / V decimal	*V* / V binary
0	5.4	1100
2	4.0	0101
4	3.8	0100

All values of *V* decimal correct [1]
At least two values of *V* binary correct [1]
All values of *V* binary correct [1]

c **i** Variations in the signal that occur between successive samples are not reproduced. [1]
The sample is only accurate at any instant to the nearest 0.2 V, so any variation between, for example, 3.1 and 3.3, will be taken as 3.2 V. [1]

ii The signal should have more bits, for example use an 8-bit binary number, so that the sample is accurate to a smaller value than 0.2 V. [1]
The samples should be taken as often as possible (for example, at twice the frequency of the highest signal present), so that all variations are detected. [1]

10 **a** **i** Attenuation from cable = 20 × 10
= 200 dB [1]

ii 4 × 40 = 160 dB [1]

b Total attenuation = 200 – 160 = 40 dB [1]

$40 = 10 \lg\left(\frac{200}{P_{out}}\right)$

$P_{out} = 2.0 \times 10^{-2}$ mW **or** 2.0×10^{-5} W [1]

c $200 = 10 \lg\left(\frac{200}{P_{out}}\right)$

$P_{out} = 2.0 \times 10 - 18$ mW **or** 2.0×10^{-21} W [1]

11 **a** *Any two from:*
Takes 24 hours to make an orbit of the Earth. [1]
It travels above the equator in the same direction as the rotation of the Earth. **or** It appears to be above the same point on the equator. [1]
It orbits at a height of 3.6×10^{4} km above the Earth's surface. [1]

b Some of the signal sent from Earth is picked up by the receiver that is near to the transmitter. Since the signal received back from the satellite is weaker, it will be swamped by the direct signal from the transmitter nearby unless different frequencies are used. [1]
(Using a different frequency, the tuning circuit only passes the weaker signal received from the satellite.)

c Less / little power is available in the satellite to generate radio waves. [1]
(The downlink signal is generated by the satellite and much less power is available to the satellite than is available on the ground.)

Answers to EOC questions

Chapter 21

1 **a** Kinetic energy remains constant, potential energy increases, internal energy increases. [3]

b Kinetic energy increases, potential energy remains constant, internal energy increases. [3]

2 Just before the stone hits the ground, it has kinetic energy. All the molecules are moving together in the same direction. [1]
When it hits the ground, this unidirectional movement of the molecules is converted into movement of individual molecules in random directions. [1]
The kinetic energy for movement in random directions is internal energy, and hence the temperature rises. [1]

3 The air is being compressed, so work is being done on it. [1]
From the first law of thermodynamics, the change in internal energy of a body is equal to the energy supplied by heating plus the energy supplied by doing work. [1]
In this case there is no energy supplied by heating but work is done in compressing the air. [1]

4 If two bodies are at the same temperature, no energy flows from one body to another. Therefore, if no energy flows from A to B and none flows from B to C, but energy flowed from C to B, it would mean that A and B were at the same temperature, and A and C are at the same temperature, but C is hotter than B. This is clearly a nonsense! We can see how fundamental this law is, and why it is called the zeroth law. [2]

5

Substance	Melting point		Boiling point	
	°C	K	°C	K
oxygen	−223	50	−183	90
hydrogen	−259	14	−253	20
lead	327	600	1750	2023
mercury	−39	234	357	630

6 **a** no temperature difference, so 0 V [1]

b temperature $= \frac{100 \times 49}{63}$ [1]
$= 77.8 \approx 78\,°C$ [1]

c **i** voltage $= \frac{63 \times 327}{100}\,\mu V$ [1]
$= 206\,\mu V$ [1]

ii That the variation between temperature difference and induced e.m.f. in the thermocouple remains linear beyond 100 °C [1]

7 **a** Liquid in glass – easy to use, quick, and great precision is not required. [2]

b Thermocouple – the smallness of the thermocouple means that the temperature can be measured at specific points on the cylinder head. [2]

c Thermocouple or thermistor – the operator can be remote from the device; she does not have to be inside the reactor vessel itself. [2]

8 **a** energy supplied $= mc\Delta\theta = 4180 \times 0.300 \times 80$ [1]
$= 100\,320\,J$ [1]
time $= \frac{\text{energy}}{\text{power}} = \frac{100\,320}{500}$ [1]
$= 201\,s$ [1]
No energy is needed to heat the element or the kettle. [1]

b energy supplied = power × time
$= 500 \times 120 = 60\,000\,J$ [1]
mass of water boiled away $= \frac{\text{energy}}{\text{s.l.h.}}$
$= \frac{60\,000}{2.26 \times 10^6} = 0.027\,kg$ [1]
mass remaining $= 300 - 27 = 273 \approx 270\,g$ [1]
No energy is lost to the surroundings; all the vapour escapes from the kettle. [1]

Cambridge International AS and A Level Physics © Cambridge University Press 2014

9 a The energy required to raise the temperature of unit mass of a substance [1]
by one degree Celsius. [1]

b i Small heat capacity [1]
Quick acting [1]

ii To allow for energy losses to the surroundings [1]
Comment about the same temperature rise therefore same energy losses **or** similar comment [1]

iii Energy input = $40 \times 190 = 7600$ J
(**or** $60 \times 114 = 6840$ J) [1]
Experiment 1:
energy input = $mc\Delta\theta$ + energy losses
$= 0.85 \times c \times 18 + Q_1$
$7600 = 15.3c + Q_1$ (1)
Experiment 2:
energy input = $mc\Delta\theta$ + energy losses
$= 0.85 \times c \times 18 + Q_2$
$6840 = 15.3c + Q_2$ (2)
Attempted use of either equation [1]
Energy loss in experiment 2 (Q_2)
$= \left(\frac{114}{190}\right) \times$ energy loss in experiment 1 (Q_1)
$Q_2 = 0.6Q_1$ (3) [1]
Subtracting equations (1) and (2):
$7600 - 6840 = Q_1 - Q_2$
Substituting equation (3) into this:
$7600 - 6840 = Q_1 - 0.6Q_1$
$Q_1 = 1900$ J [1]
Substituting for Q_1 in equation (1):
$7600 = 15.3c + 1900$
$c = 372$ J kg^{-1} °C^{-1} [1]

10 a The carbon dioxide expands, doing work against the atmosphere. [1]
It gives up internal energy to do this work. [1]
Link to supplying latent heat [1]

b Energy required to change state of unit mass (from solid to vapour) [1]
without change in temperature [1]

c i So allowance for energy gains from the atmosphere can be made [1]

ii Use of $E = ml$ [1]
m = mass decrease with heater on minus mass decrease with heater off [1]
$m = 20.9$ g
$l = \frac{12\,000}{20.9} = 574$ J g^{-1} [1]

d Total work done / energy required
= work done in separating molecules / breaking bonds from solid state (to liquid state) [1]
plus work done in further separating the molecules to gaseous state. [1]

11 a Energy goes to potential energy of the molecules [1]
as they are moved further apart (accept bonds broken) [1]

b To reduce the energy gained from the surroundings [1]

c Energy input = $40 \times 2 \times 60$ (= 4800 J) [1]
Use of $E = ml$ [1]
m = mass decrease with heater on minus half mass decrease with heater off [1]
$m = 23.8$ g
$l = \frac{4800}{23.8} = 202$ J g^{-1} [1]

12 a i Sum of the kinetic energy [1]
and potential energy of the molecules. [1]

ii Temperature at which energy cannot be removed from molecules [1]
Answer makes clear that the molecules still have (potential) energy. [1]
(Give [1] maximum for: all energy removed from molecules.)

b i Mass per second = ρAv
$= 1000 \times 4.8 \times 10^{-5} \times 1.2$ [1]
$= 0.058$ kg [1]

ii $E = mc\Delta\theta$ leading to $\Delta\theta = \frac{9000}{0.058 \times 4200}$ [1]
$= 37$ °C [1]
Final temperature of the water
$= 37 + 15 = 52$ °C [1]

iii The heater is 100% efficient. [1]

iv Decrease the rate of flow of water [1]

Answers to EOC questions

Chapter 22

1 a i 6.02×10^{23} [1]

ii Chlorine is diatomic, so 12.0×10^{23} [1]

iii $1000 \times N_A = 6.02 \times 10^{26}$ [1]

b i $4 \times N_A = 24.1 \times 10^{23}$ [1]

ii One C atom per molecule, so 24.1×10^{23} [1]

iii Two O atoms per molecule, so 48.2×10^{23} [1]

2 a 1 mol = 197 g [1]

so 1000 g = $\frac{1000}{197}$ mol = 5.08 mol ≈ 5.1 mol [1]

b number of atoms = $5.08 \times N_A$

$= 3.06 \times 10^{24} \approx 3.1 \times 10^{24}$ [1]

c mass of atom = $\frac{\text{mass of bar}}{\text{number of atoms in bar}}$

$= 3.27 \times 10^{-22}$ g ≈ 3.3×10^{-22} g [1]

3 a Pressure $p_2 = \frac{p_1 V_1}{V_2}$ [1]

$= \frac{10^5 \times 140}{42} = 3.4 \times 10^5$ Pa [1]

b The temperature of the gas would increase, causing pressure to be higher than in **a**. [1]

4 $V_2 = \frac{p_1 V_1}{p_2}$ [1]

$= \frac{35 \times 0.42}{10}$ [1]

$= 1.47 \approx 1.5$ cm^3 [1]

assumes that the temperature at 25 m depth is equal to the temperature at the surface of the water [1]

5 a Using $pV = nRT$, we have

$n = \frac{pV}{RT} = \frac{4.8 \times 10^5 \times 0.040}{8.31 \times 293}$ [1]

$= 7.89$ mol ≈ 7.9 mol [1]

b Using $pV = nRT$, we have

$n = \frac{pV}{RT} = \frac{4.8 \times 10^5 \times 0.040}{8.31 \times 293}$ [1]

$= 347$ g ≈ 350 g [1]

6 Using $pV = nRT$, we have [1]

$V = \frac{nRT}{p} = \frac{1 \times 8.31 \times 273}{1.01 \times 10^5} = 22.4$ dm^3 [1]

7 Using $pV = NkT$, we have $p = \frac{NkT}{V}$ so [1]

$p = \frac{3.0 \times 10^{26} \times 1.38 \times 10^{-23} \times 400}{0.200}$ [1]

$= 8.31 \times 10^6$ Pa [1]

8 a Using $p = \frac{1}{3}\rho\langle c^2\rangle$, we have $\langle c^2\rangle = \frac{3p}{\rho}$ [1]

$= \frac{3 \times 1.01 \times 10^5}{0.179} = 1.69 \times 10^6$ m^2 s^{-2} [1]

so average speed = $\sqrt{1.69 \times 10^6} = 1350$ m s^{-1} [1]

b This is considerably faster than air molecules at the same pressure because the He atoms have a much smaller mass, so greater speed needed for same energy. [1]

9 a i Average KE per molecule at 27 °C = $\frac{3}{2}kT$ [1]

$= 1.5 \times 1.38 \times 10^{-23} \times 300$ [1]

$= 6.2 \times 10^{-21}$ J [1]

ii Average KE at 243 °C = 1.07×10^{-20} J [1]

b Ratio of energies = $\frac{1.07}{0.62} = 1.72$ [1]

Ratio of speeds = $\sqrt{\text{(ratio of energies)}}$

$= 1.31 : 1$ [1]

or $\sqrt{\text{(ratio of temperatures)}} = 1.31 : 1$

10 a The molecules hit the wall / rebound from the wall. [1]

Each impact causes a small force / impulse on the wall. [1]

Molecules are in random motion / many impacts smear force over whole wall. [1]

b $pV = nRT$ [1]

$n = \frac{3.42 \times 10^5 \times 1.50}{8.31 \times 276}$ [1]

$= 224$ mol [1]

c i Using $pV = nRT$

$p = \frac{224 \times 8.31 \times 315}{1.50}$ [1]

$= 3.91 \times 10^5$ Pa [1]

ii $\Delta E = \frac{3}{2k\Delta T} = \frac{3}{2} \times 1.38 \times 10^{-23} \times (315 - 276)$ [1]

$\Delta E = 8.01 \times 10^{-22}$ J [1]

11 a Random zigzag movement of large particle (e.g. smoke, pollen, dust) [1]

Air / water molecules collide with bigger particles. [1]

Each collision gives the larger particle an impulse in a different direction. [1]

b $p = \frac{1}{3}\rho\langle c^2\rangle$ [1]

$\langle c^2\rangle = \frac{3 \times 1.03 \times 10^5}{1.21}$ [1]

$= 2.55 \times 10^5$ (m^2 s^{-2}) [1]

r.m.s. speed = 505 m s^{-1} [1]

c i Speed will increase. [1]

Temperature is a measure of the average kinetic energy of the molecules. [1]

ii Speed will be unchanged. [1]

12 **a** An ideal gas is a gas that obeys the ideal gas law
at all temperatures [1]
and pressures. [1]

b **i** 125 (mol) [1]

ii $125 \times 6.02 \times 10^{23} = 7.53 \times 10^{25}$ (molecules) [1]

c $pV = nRT$ [1]

$V = \frac{125 \times 8.31 \times 300}{5 \times 10^5}$ [1]

$= 0.62\,m^3$ [1]

d **i** The helium does work on the atmosphere. [1]

Energy is taken from the helium. [1]

ii The average speed drops. [1]

13 **a** 4410 N [1]

b **i** Mass = 3690 kg [1]

ii 1.27×10^5 mol [1]

c To achieve lift, a minimum of 450 kg of air must leave the balloon. [1]
Maximum mass of remaining air
= 3690 – 450 = 3240 kg [1]

d Remaining air = $\frac{3240}{29 \times 10^{-3}} = 1.12 \times 10^5$ mol [1]

$pV = nRT$ [1]

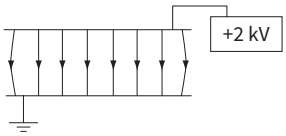
$T = \frac{1.03 \times 10^5 \times 3000}{1.12 \times 10^5 \times 8.31}$ [1]

$= 332$ K (59 °C) [1]

Answers to EOC questions

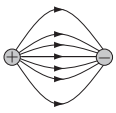
Chapter 23

1 a



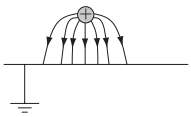
[1]

b



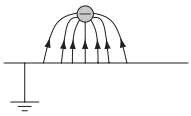
[1]

c



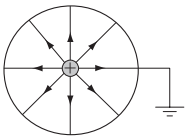
[1]

d



[1]

e



[1]

2 a $E = \frac{V}{d} = \frac{2500}{0.040}$ [1]

$= 6.25 \times 10^4\ \text{V m}^{-1}$ [1]

b i $F = EQ = 6.25 \times 10^4 \times 2.4 \times 10^{-9}$ [1]

$= 1.5 \times 10^{-4}\ \text{N}$ [1]

ii $a = \frac{F}{m} = \frac{1.5 \times 10^{-4}}{4.2 \times 10^{-6}}$ [1]

$= 3.6 \times 10^4\ \text{m s}^{-2}$ [1]

3 a $E = \frac{Q}{4\pi\varepsilon_0 r^2} = \frac{9 \times 10^9 \times 2.4 \times 10^{-9}}{0.02^2}$ [1]

$= 5.4 \times 10^4\ \text{V m}^{-1}$ [1]

b $E = \frac{Q}{4\pi\varepsilon_0 r^2} = \frac{9 \times 10^9 \times 2.4 \times 10^{-9}}{0.04^2}$ [1]

$= 1.35 \times 10^4\ \text{V m}^{-1}$ [1]

(**or** one-quarter of the answer to part **a** since distance has doubled)

4 a potential $V = \frac{Q}{4\pi\varepsilon_0 r}$ so

$Q = 4\pi\varepsilon_0 r \times V$ [1]

$= \frac{0.030 \times 20\,000}{9 \times 10^9}$ [1]

$= 6.7 \times 10^{-8}\ \text{C}$ [1]

b $E = \frac{V}{r} = \frac{20\,000}{0.030}$ [1]

$= 6.7 \times 10^5\ \text{V m}^{-1}$ **or** N C^{-1} [1]

5 $V = \frac{Q}{4\pi\varepsilon_0 r} = \frac{9 \times 10^9 \times 1.6 \times 10^{-19}}{1.05 \times 10^{-10}}$ [1]

$= 13.7\ \text{V}$ [1]

6 a The force at a point [1]

per unit positive charge at the point. [1]

b i $E = \frac{1}{4\pi\varepsilon_0} \times \frac{Q}{r^2} = \frac{9 \times 10^9 \times 2 \times 10^{-9}}{(5.0 \times 10^{-2})^2}$ [1]

$E = 7200\ \text{V m}^{-1}$ [1]

ii Field due to B = 7200 – 1800 = $5400\ \text{V m}^{-1}$ [1]

$E = \frac{1}{4\pi\varepsilon_0} \times \frac{Q}{r^2}$ leading to

$Q = E \times 4\pi\varepsilon_0 \times r^2$ [1]

$= 1.5 \times 10^{-9}\ \text{C}$ [1]

c i $V = \frac{1}{4\pi\varepsilon_0} \times \frac{Q}{r} = \frac{9 \times 10^9 \times 2 \times 10^{-19}}{1.0 \times 10^{-2}}$ [1]

$V = 1800\ \text{V}$ [1]

ii Greater before removing sphere B [1]

Potential is a scalar, thus two potentials tend to add. [1]

7 a i $W = VQ$ leading to

$8.0 \times 10^{-13} = V \times 2 \times 1.6 \times 10^{-19}$ [1]

$V = 2.5 \times 10^6\ \text{V}$ [1]

ii $E_k = \frac{1}{2}mv^2$ leading to

$8.0 \times 10^{-13} = \frac{1}{2} \times 6.65 \times 10^{-27} \times v^2$ [1]

$v^2 = 2.4 \times 10^{14}$ [1]

$v = 1.6 \times 10^7\ \text{m s}^{-1}$ [1]

b i Zero [1]

ii Zero [1]

iii $8.0 \times 10^{-13}\ \text{J}$ [1]

c $W = \frac{1}{4\pi\varepsilon_0} \times \frac{Q_1 Q_2}{r}$ leading to

$r = \frac{9.0 \times 10^9 \times 79 \times 1.6 \times 10^{-19} \times 3.2 \times 10^{-19}}{8.0 \times 10^{-13}}$ [1]

$r = 4.5 \times 10^{-14}\ \text{m}$ [1]

d Radius of nucleus is $4.5 \times 10^{-14}\ \text{m}$ [1]

This is the maximum radius of the nucleus; the inverse square law is still valid. [1]

8 a Potential at a point is the work done in bringing unit charge [1]

from infinity to the point. [1]

b i $V = \frac{-1}{4\pi\varepsilon_0} \times \frac{Q}{r}$

$= \frac{-9.0 \times 10^9 \times 1.6 \times 10^{-19}}{1.04 \times 10^{-10}}$ [1]

$= -13.8\ \text{V}$ [1]

ii $E = qV = 1.6 \times 10^{-19} \times 13.8$ [1]
$= 2.2 \times 10^{-19}$ J [1]

c Use of gradient [1]
leading to $E = 1.4 \times 10^{11}$ V m^{-1} [1]

9 a $E = \frac{1}{4\pi\varepsilon_0} \times \frac{Q}{r^2}$

$= \frac{9 \times 10^9 \times 6.0 \times 10^{-8}}{(0.8 \times 10^{-2})^2}$ [1]

$= 8.4 \times 10^6$ V m^{-1} [1]

b i $F = \frac{1}{4\pi\varepsilon_0} \times \frac{Q_1 Q_2}{r^2}$

$= \frac{9 \times 10^9 \times 6.0 \times 10^{-8} \times (-4.5 \times 10^{-8})}{(5.0 \times 10^{-2})^2}$ [1]

$= 9.72 \times 10^{-3}$ N [1]

ii Reading = 0.0482 – 0.0097 = 0.0384 N [1]

c $W = \frac{1}{4\pi\varepsilon_0} \times \frac{Q_1 Q_2}{(1/r_2 - 1/r_1)}$

$= \frac{9 \times 10^9 \times 6.0 \times 10^{-8} \times (-4.5 \times 10^{-8})}{(1/0.035 - 1/0.050)}$ [1]

$= -2.1 \times 10^{-4}$ J [2]

(Lose one mark if the minus sign is missing.)

Answers to EOC questions

Chapter 24

1 $Q = CV = 470 \times 10^{-6} \times 9.0 = 4.2 \times 10^{-3}$ C [1]

2 $V = \frac{Q}{C} = \frac{0.033}{2200 \times 10^{-6}} = 15$ V [1]

3 $C = \frac{Q}{V} = \frac{2.0}{5000} = 4.0 \times 10^{-4} = 400$ μF [1]

4 $W = \frac{1}{2}CV^2 = 0.5 \times 470 \times 10^{-6} \times 12^2 = 0.034$ J [1]

5 $W = \frac{1}{2}QV = 0.5 \times 1.5 \times 10^{-3} \times 50 = 0.0375$ J [1]

6 **a** $W = \frac{1}{2}CV^2 = 0.5 \times 5000 \times 10^{-6} \times 24^2 = 1.44$ J [1]

b When charge is halved, p.d. is halved [1]
so energy stored = $0.5 \times 5000 \times 10^{-6} \times 12^2$
= 0.36 J [1]
Energy dissipated in lamp = 1.44 J – 0.36 J
= 1.08 J [1]

7 **a** $W = \frac{1}{2}CV^2 = 0.5 \times 4700 \times 10^{-6} \times 12^2 = 0.34$ J [1]

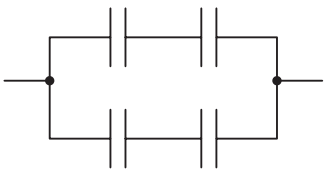
b Charge stored = CV
$= 4700 \times 10^{-6} \times 12 = 0.056$ C [1]

c Average current = $\frac{Q}{t} = \frac{0.056}{2.5} = 0.023$ A [1]

d average p.d. = 6.0 V [1]
$R = \frac{V}{I} = \frac{6.0}{0.023} = 260$ W [1]

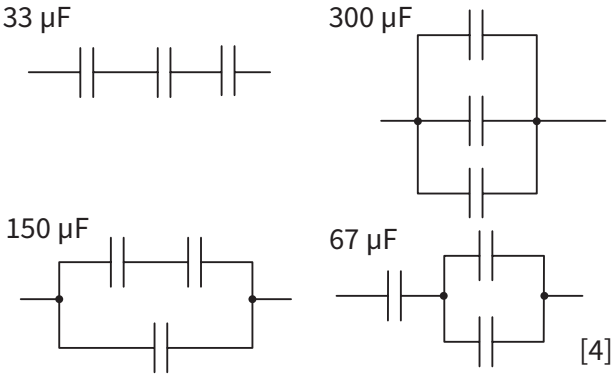
e Current is dependent on p.d., which decreases at a non-uniform rate. [1]

8



[1]

9



[4]

10 **a** $\frac{1}{\text{total capacitance}} = \frac{1}{100} + \frac{1}{200} + \frac{1}{600} = \frac{10}{600}$ [1]

so combined capacitance = $\frac{600}{10} = 60$ μF [1]

Charge stored = $1.5 \times 60 \times 10^{-6} = 90$ μC [1]
$Q_1 = +90$ μC, $Q_2 = -90$ μC, $Q_3 = +90$ μC,
$Q_4 = -90$ μC, $Q_5 = +90$ μC, $Q_6 = -90$ μC [2]

b Using $V = \frac{Q}{C}$ [1]

p.d. = 0.90 V across the 100 μF capacitor,
0.45 V across the 200 μF capacitor, [1]
0.15 V across the 600 μF capacitor [1]

11 **a** Time delay, anti-surge, anti-spark, etc. [1]

b **i** $R = \frac{V}{I} = \frac{9.0}{15 \times 10^{-3}}$ [1]

$R = 600\ \Omega$ [1]

ii p.d. decreases across capacitor [1]
as charge flows off, so less p.d. to drive
current through the resistor [1]

iii Evidence of using the area [1]
45 ± 5 mC [1]

iv $C = \frac{Q}{V} = \frac{45 \times 10^{-3}}{9.0}$ [1]
$= 5.0 \times 10^{-3}$ μF [1]

12 **a** **i** $V = \frac{1}{4\pi\varepsilon_0} \times \frac{Q}{r}$ leading to $Q = V \times r \times 4\pi\varepsilon_0$ [1]

$Q = 1.2 \times 10^{-6}$ C [1]

ii $C = \frac{Q}{V} = \frac{1.2 \times 10^{6}}{54\,000}$ [1]

$= 2.2 \times 10^{-11}$ F [1]

b $E = \frac{1}{2}QV$ **or** $\frac{1}{2}CV^2$

$= \frac{1}{2} \times 1.2 \times 10^{-6} \times 54\,000$ (**or** equivalent) [1]

initial energy = 0.0324 J [1]

p.d. after discharge = 13.5 kV [1]

therefore energy remaining = $\frac{1}{2}CV^2$
$= \frac{1}{2} \times 2.2 \times 10^{-11} \times 13\,500^2 = 0.0020$ J [1]

energy released = 0.0324 – 0.0020 ≈ 0.03 J [1]

c Not sufficient p.d. between sphere and plate [1]

to continue ionising the air. [1]

13 **a** $V = \frac{1}{4\pi\varepsilon_0} \times \frac{Q}{r}$ so $Q = 4\pi\varepsilon_0 \times V \times r$ [1]

$C = \frac{Q}{V} = \frac{4\pi\varepsilon_0 \times V \times r}{V} = 4\pi\varepsilon_0 \times r$ [1]

b **i** $V \frac{1}{4\pi\varepsilon_0} \times \frac{Q}{r} = \frac{9 \times 10^{9} \times 5.0 \times 10^{-8}}{0.1}$ [1]

$= 4500$ V [1]

ii $W = \frac{1}{2}QV$ (**or** otherwise)

$= \frac{1}{2} \times 5.0 \times 10^{-8} \times 4500$ [1]

$= 1.1 \times 10^{-4}$ J [1]

c i New charge on each sphere $= 2.5 \times 10^{-8}$ C [1]

new potential = 2250 V [1]

new energy (on each) = ½ × $2.5 \times 10^{-8} \times 250$

$= 2.8 \times 10^{-5}$ J [1]

ii Energy is lost to heating the spheres, as charge moves from one to the other. [1]

14 a Capacitance is the charge stored [1]

per unit potential difference across the capacitor plates. [1]

b i $Q = CV = 67 \times 10^{-6} \times 12$ [1]

$Q = 804 \approx 800\ \mu C$ [1]

ii $I = \frac{Q}{t} = \frac{804 \times 10^{-6}}{1/50}$ [1]

$= 0.04$ A [1]

iii $P = VI$ where V is average p.d. [1]

$= \frac{1}{2} \times 12 \times 0.04$ [1]

$P = 0.24$ W [1]

c Capacitance is halved [1]

Charge stored is halved [1]

Current is halved but (average) V is unchanged [1]

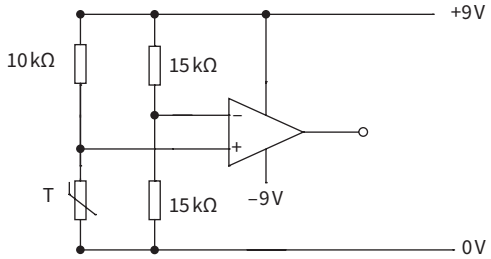
Power is halved [1]

(Maximum [2] if qualitative, i.e. capacitance reduced, etc.)

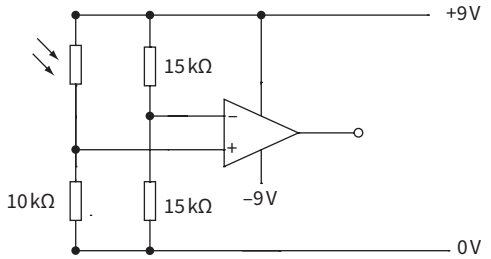
Answers to EOC questions

Chapter 25

1 Exchange T and the 10 kΩ resistor. [1]



2



[3]

3 **a** The process in which a part of the output of a device is passed back to the input. [1]

b Unless an op-amp is saturated, the p.d. between the two input terminals of an op-amp is almost zero. Since one of the terminals, usually the (+) terminal, is connected to 0 V (earth), the other terminal (–) is almost at earth potential. [2]

4 **a** Gain $= \frac{-R_f}{R_{in}}$ [1]

$= \frac{-200}{10} = -20$ [1]

b Input voltage $= \frac{\text{output voltage}}{\text{gain}}$ [1]

$= \frac{-8.0}{20} = -0.40\,\text{V}$ [1]

c Maximum input voltage for the op-amp to still have the gain calculated in **a** is equal to

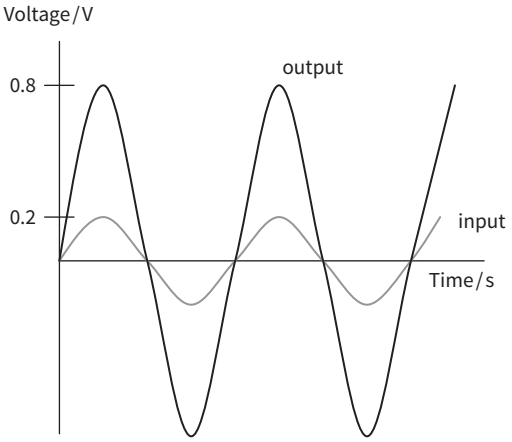
$\frac{V_s}{20} = \frac{12}{20} = 0.60\,\text{V}$ [1]

5 **a** The feedback resistor R_f is half the resistance of the input resistor R_{in}. This then means the output voltage is half of the input voltage and of opposite sign. [2]

b The gain of a non-inverting amplifier $= 1 + \frac{R_1}{R_2}$. The smallest value of $\frac{R_1}{R_2}$ is zero, so the smallest value of the gain is 1. [1]

When R_1 is zero, all of the output voltage is fed back to the negative input (–). As long as the op-amp is not saturated, the (+) and (–) terminals are at the same potential and the gain is 1.0. If R_1 is any bigger, a smaller fraction of the output voltage is fed back. Since this smaller fraction is the input voltage, then the output voltage must be larger in value than the input. [1]

6



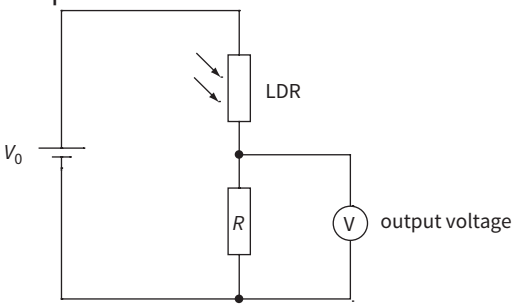
[3]

7 Non-inverting amplifier [1]

of gain $\frac{(9+1)}{1} = 10$ [2]

8 **a** Power supply, resistor and LDR in series [1]

Output marked across the series resistor [1]



b As light intensity increases, the resistance of the LDR decreases. [1]

The p.d. across the series resistor R is given by: output voltage $= \frac{V_0 \times R}{(R + \text{resistance of LDR})}$

As the resistance of the LDR decreases, the output voltage increases. [1]

c The whole circuit is the sensor (the LDR is the sensing device) and the voltmeter is the output device. [1]

The processor is the potential divider (two resistors and voltage supply). [1]

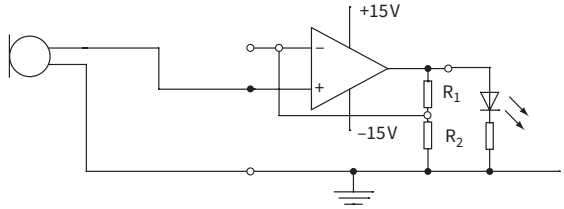
Cambridge International AS and A Level Physics © Cambridge University Press 2014

9 a The wire becomes longer (and thinner). [1]
Resistance $R = \frac{\rho L}{A}$ where ρ is the resistivity, L is the length and A is the cross-sectional area. [1]
If all quantities apart from L are constant, then, if the length increases to $L + \Delta L$,
$R + \Delta R = \frac{\rho(L + \Delta L)}{A}$
Subtracting gives
$\Delta R = \frac{\rho \Delta L}{A}$ so ΔR is proportional to ΔL [1]

b Voltmeter reading $= V_0 \times \frac{R_1}{(R_1 + R_2)} = 5 \times \frac{150}{450}$ [1]
$= 1.67 \approx 1.7\,\text{V}$ [1]

c The voltmeter reading increases. [1]
The resistance of the strain gauge is a larger fraction of the total resistance, and thus there is a larger fraction of the supply voltage across the strain gauge. [1]

10 a i Microphone connected to (+) input and 0 V. [1]
Resistors connected to op-amp making a non-inverting amplifier. [1]



ii LED connected to output and zero volt (earth) line [1]
Series resistor included [1]

b *Any two from:* less distortion, increased bandwidth, more stable gain, reduced gain (from very large value to a sensible value) [2]

c i Maximum gain $= \frac{15}{0.020}$ [1]
$= 750$ [1]

ii $1 + \frac{R_1}{R_2} = 750$ [1]
Any value where $R_1 = 749 R_2$
(e.g. $R_2 = 1\,\text{k}\Omega$ and $R_2 = 749\,\text{k}\Omega$)
with $300\,\text{W} \leq R_1 + R_2 < 10\,\text{MW}$ [1]

d As the potential at the (+) terminal of the op-amp changes / rises, the output voltage changes / increases.
Some (1/750th) of the increased output voltage is fed back to the (–) terminal of the op-amp. [1]
Output rises until the potential at the (+) and (–) terminals of the op-amp are (very nearly) equal. [1]

11 a i +9 V [1]

ii –9 V [1]

b $V_A = \frac{9 \times 300}{500}$ [1]
$= 5.4\,\text{V}$ [1]

c i Potential at A is greater than at B **or** potential at B is 3.4 V so V_{out} is +9 V. [1]

ii Potential at B is greater than at A *and* V_{out} is –9 V **or** motor runs in reverse. [1]

d The difference in level of light will fluctuate. Sometimes P will have more light than Q and, since the op-amp has such a high gain, its output will switch from +9 V to –9 V as the relative amount of light on P and Q fluctuates. [1]

12 a (Unless the op-amp is saturated) the p.d. between the two input terminals of an op-amp is almost nothing. [1]
Since the (+) terminal of the op-amp is connected to 0 V (earth), the other terminal (–) is almost at earth potential and called a virtual earth. [1]
The input resistance (or impedance) of the op-amp is very large **or** negligible current enters the input terminals of the op-amp. [1]
current in R_1 = – current in R_2 [1]

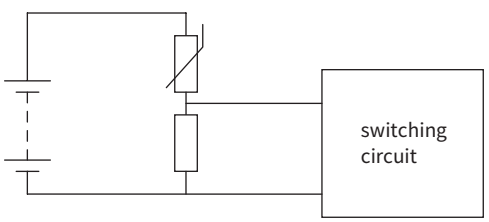
$$\frac{V_{out}}{R_1} = -\frac{V_{in}}{R_2} \text{ so}$$

gain $= \frac{V_{out}}{V_{in}} = -\frac{R_1}{R_2}$ [1]

b i Gain $= -\frac{20}{4} = -5$ [1]
$V_{out} = -5 \times 0.5 = -2.5\,\text{V}$ [1]

ii Total resistance of two parallel resistors $= 2\,\text{k}\Omega$ [1]
Gain becomes $-\frac{20}{2} = -10$ so $V_{out} = -5.0\,\text{V}$ [1]

13 a



[2]

b Increase the resistance in the potential divider circuit. [1]

14 a Light-dependent resistor (LDR). [1]
When the light level increases the resistance of the LDR decreases [1]
The current through the variable resistor increases, so the potential difference across it increases and the potential difference across the LDR falls [1]
or When the light level decreases the resistance of the LDR increases [1]
The current through the variable resistor decreases, so the potential difference across it decreases and the potential difference across the LDR rises. [1]

Answers to EOC questions

Chapter 26

1 a (The force F is given by $F = BIl\sin\theta$.)
The force is a maximum the angle θ between the wire and the magnetic field is 90° (i.e. when $\sin\theta = 1$). [1]

b The force is zero when the angle θ between the wire and the magnetic field is 0°. (The wire is parallel to the magnetic field.) [1]

2 a $F = BIl \propto I$ (force ∝ current) [1]
Hence, the force increases by a factor of 3.0 to a value of 1.41×10^{-2} N. [1]

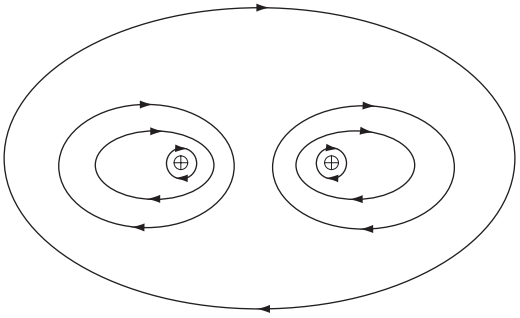
b $F = BIl \propto \Delta B$
(force ∝ change in magnetic flux density) [1]
Hence, the force is halved to a value of 2.35×10^{-3} N. [1]

c $F = BIl \propto \Delta l$
(force ∝ change in length of wire in the field) [1]
Hence, the force is reduced to 40% of its initial value to 1.88×10^{-3} N. [1]

3 a $F = BIl\sin\theta$ [1]
$B = \frac{F}{Il\sin\theta} = \frac{3.8 \times 10^{-3}}{1.2 \times 0.03 \times \sin 50°}$ [1]
$B = 0.138\,\text{T} \approx 0.14\,\text{T}$ [1]

b The direction is given by Fleming's left-hand rule. The wire experiences a force into the plane of the paper. [1]

4 a



Clockwise magnetic field lines around and close to each strip. [1]
Elliptical lines further away from the strips, eventually becoming elliptical around both strips (even farther away, the shape becomes circular, not shown in the diagram). [1]

b The force on strip A is towards strip B and the force on strip B is towards strip A, i.e. the strips attract each other. [1]
This is because strip A, on its own, produces a magnetic field vertically down the paper at strip B by the right-hand rule. The left-hand rule can then be applied to strip B, which has a current into the plane of the paper and a field down the paper, and so the force is to the left, towards strip A. [1]

5 a The current is from Y to X. This is because Q shows that the magnetic field above the wire, produced by the current, is from west to east. The right-hand rule then shows the current is upwards. [1]

b P points towards the north-west. [1]

c Q then points towards the north-west. [1]

6 a $F = BIl$ [1]
$= 4.5 \times 10^{-3} \times 2.5 \times 0.07$ [1]
$= 7.88 \times 10^{-4} \approx 7.9 \times 10^{-4}$ N [1]

b The magnetic field is parallel to the current (or the wire). [1]

c From Fleming's left-hand rule, PQ experiences a force out of the plane of the paper and RS experiences a force into the plane of the paper. [1]
Hence, the frame starts to rotate in a clockwise direction (when viewed from the end PS). [1]

d Torque $= F \times d = 7.88 \times 10^{-4} \times 0.04$ [1]
$= 3.15 \times 10^{-5}$ N m
≈ 3.1 **or** 3.2×10^{-5} N m [1]

7 a Left to right [1]

b Force = weight of paper tape
$F = mg = 60 \times 10^{-3} \times 10^{-3} \times 9.81$ [1]
$= 5.89 \times 10^{-4}$ N $\approx 5.9 \times 10^{-4}$ N [1]

c $B = \frac{F}{Il}$ [1]
$= \frac{5.89 \times 10^{-4}}{8.5 \times 0.052}$ [1]
$= 1.33 \times 10^{-3}$ T ≈ 1.3 mT [1]

d Moves up and down slowly [1]

8 a Diagram showing wire and magnetic field and a method of measuring the force, e.g. diagrams similar to either Figure 26.14 or 26.15 in the coursebook [1]
Measure I, the current, and F, the force. [1]
Method of measuring the force, e.g. mass of paper tape (in kg) × 9.81 **or** difference in readings on top-pan balance (in kg) × 9.81 [1]
Measure length of wire at right angles to magnetic field. [1]
$B = \frac{F}{Il}$ [1]

b i Using the left-hand rule, the field is horizontally towards the north, the force is upwards and thus the current is from west to east. [1]

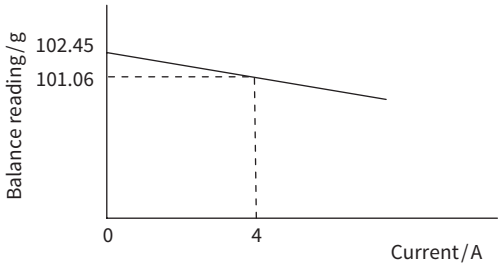
ii $I = \frac{F}{Bl}$ [1]
$= \frac{0.02}{1.6 \times 10^{-5} \times 3.0}$ [1]
$= 417 \approx 420\,\text{A}$ [1]

9 a The wire is carrying a current in a magnetic field and experiences a force. [1]
There is an upwards force on the top-pan balance, reducing the reading. [1]

b By Newton's third law, the force is downwards on the wire to produce an upwards force on the top-pan balance. [1]
By the left-hand rule, the current in the wire is from left to right. [1]

c Force on wire = $(102.45 - 101.06) \times 10^{-3} \times 9.81$
$= 0.0136\,\text{N}$ [1]
$B = \frac{F}{Il} = \frac{0.0136}{4.0 \times 5.0 \times 10^{-2}} = 0.068\,\text{T}$ [1]

d Axes labelled and balance reading when current is zero marked [1]
Balance reading decreases linearly with current [1]



10 a Magnetic flux density is the force per unit length on a wire carrying a current of 1 A [1] when placed at right angles to the magnetic field. [1]
Both electric field strength and magnetic field strength are defined in terms of the force on an object. For the electric field strength, the object is a charge of one coulomb (1 C); for the magnetic field strength, the object is a wire of length 1 m carrying a current of 1 A. [1]

b i The two wires attract one another **or** there is a force upwards on the wire on the bench. [1]
The wire above the bench produces a horizontal magnetic field on the wire lying on the bench. [1]
This magnetic field interacts with the current in the wire lying on the bench to produce a force upwards. [1]
The currents flow in the same direction causing the wires to attract each other. [1]

ii $B = \frac{2.0 \times 10^{-7} \times 4.0}{0.03} = 2.67 \times 10^{-5}\,\text{T}$ [1]
$F = BIl = 2.67 \times 10^{-5} \times 4.0 \times 1$ [1]
$= 1.07 \times 10^{-4} \approx 1.1 \times 10^{-4}\,\text{N}$ [1]

Answers to EOC questions

Chapter 27

1 The force is at right angles to the velocity v [1] and does no work on the electron, since work = force × distance moved in the direction of the force. [1]
As no work is done on the electron, the kinetic energy of the electron is constant, as is its speed. [1]

2 $\frac{1}{2}mv^2 = eV = e \times 1600$ [1]
$r = \frac{mv}{Be}$ [1]
$\frac{e}{m} = \frac{2V}{B^2r^2} = \frac{3200}{0.017^2 \times (8 \times 10^{-3})^2}$ [1]
$= 1.73 \times 10^{11} \approx 1.7 \times 10^{11}\,\text{C kg}^{-1}$ [1]

3 a Ratio $\frac{\text{mass of α-particle}}{\text{mass of β-particle}}$ [1]
$= \frac{4 \times 1.67 \times 10^{-27}}{9.11 \times 10^{-31}} = 7333 \approx 7300$ [1]

b Ratio $\frac{\text{charge of α-particle}}{\text{charge of β-particle}}$ [1]
$= \frac{3.2 \times 10^{-19}}{1.6 \times 10^{-19}} = 2.0$ [1]

c Ratio $\frac{\text{force on α-particle}}{\text{force on β-particle}}$ [1]
= ratio of charges = 2.0 [1]

d $\frac{r_a}{r_b} = \frac{m_a Q_b}{m_b Q_a} = \frac{7333}{2} = 3666 \approx 3700$ [1]

4 The force is always at right angles to the direction of the magnetic field and is in the same direction (or the reverse) in an electric field. [1]
The force is proportional to the velocity of the electron in a magnetic field but does not depend on the velocity of the electron in an electric field. [1]

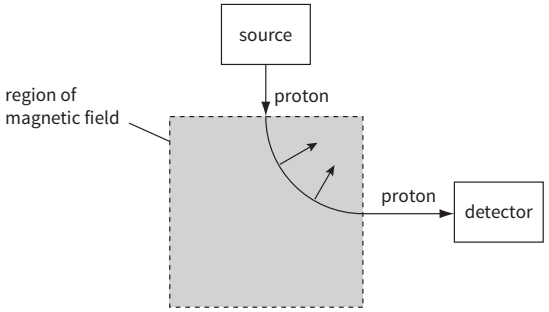
5 a i The electron is charged. **or** A moving charge is the equivalent of a current. [1]
The magnetic field of the moving charge interacts with the uniform magnetic field. [1]

ii The electron ionises atoms as it moves and loses (kinetic) energy. [1]
The speed of the electron decreases and the radius decreases as $r = \frac{mv}{B}$. [1]

b i $F = Bev = 0.25 \times 1.6 \times 10^{-19} \times 1.0 \times 10^7$ [1]
$= 4.0 \times 10^{-13}\,\text{N}$ [1]

ii $r = \frac{mv^2}{F} = \frac{9.11 \times 10^{-31} \times (1.0 \times 10^7)^2}{4.0 \times 10^{-13}}$ [1]
$= 2.278 \times 10^{-4} \approx 2.3 \times 10^{-4}\,\text{m}$ [1]

6 a i Quarter of a circle linking exit and entry points [1]
Force at right angles to the circle marked at two points [1]
Both forces towards the centre of the circle (top right-hand corner of square) [1]



ii Into the plane of the paper. [1]

b i $F = BQv = 0.25 \times 1.6 \times 10^{-19} \times 4.0 \times 10^6$
$= 1.6 \times 10^{-13}\,\text{N}$ [1]

ii $r = \frac{mv^2}{F} = \frac{1.67 \times 10^{-27} \times (4.0 \times 10^6)^2}{1.6 \times 10^{-13}}$ [1]
$= 0.167 \approx 0.17\,\text{m}$ [1]

c The direction of the magnetic field is reversed. [1]
The magnetic field is reduced in strength. [1]

7 a i Diagram showing – on right side and + on left side [1]
Force on moving charges in magnetic field [1]
Left-hand rule explained [1]

ii Force due to electric field or charges on sides cancels magnetic force. [1]

b Larger number density [1]
More electrons have enough energy to go from the valence band to the conduction band [1]
Smaller drift velocity [1]
Same charge flow per second but larger number of charge carriers [1]
Smaller Hall voltage [1]

Smaller force on each charge carrier due to lower drift velocity needs a smaller electric field to cancel force due to magnetic field [1]

8 a i The electron is charged and is in an electric field. [1]
The electric field from the plates is downwards and the electron is negative, so there is a constant force upwards. [1]

ii A (small) spot is formed on the screen. [1]
Higher-speed electrons are deflected less and the spot would become larger. [1]

b $\frac{1}{2}mv^2 = e \times V_{ac}$ [1]

c i Into the plane of the paper / page [1]

ii $Bev = eE$ [1]

$v = \frac{E}{B}$ [1]

iii $\frac{1}{2}mv^2 = e \times V_{ac}$

so $v = \sqrt{(2 \times 1000 \times 1.76 \times 10^{11})}$

$= 1.88 \times 10^7 \approx 1.9 \times 10^7 \text{ m s}^{-1}$ [1]

$E = \frac{V}{d} = \frac{500}{5.0 \times 10^{-2}} = 1.0 \times 10^4 \text{ V m}^{-1}$

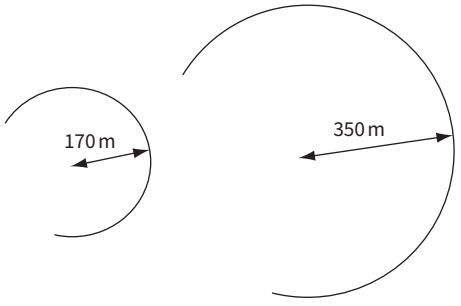
$B = \frac{E}{v} = \frac{1.0 \times 10^4}{1.88 \times 10^7} = 5.3 \times 10^{-4} \text{ T}$ [1]

9 a $\frac{mv^2}{r} = BQv$ [1]

$r = \frac{mv}{BQ} = \frac{1.67 \times 10^{-27} \times 1.0 \times 10^6}{6.0 \times 10^{-5} \times 1.6 \times 10^{-19}}$ [1]

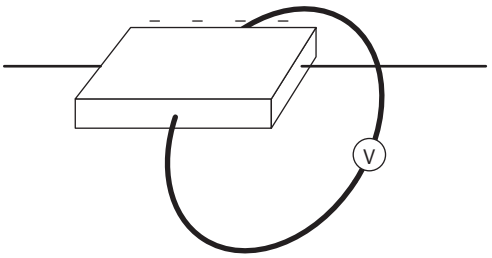
$r = 174 \approx 170 \text{ m}$ [1]

b Two paths shown, both circles or parts of circles of different radius [1]
Radius of helium nucleus twice as large as radius of proton (350 m) [1]



radius of path of proton radius of path of helium nucleus

10 a



i The opposite face to the one marked X. [1]

ii Between X and the opposite face [1]

b Electrons move until the magnetic force (Bev) is equal to the electric force (eV_H/d). [1]

$Bev = \frac{eV_H}{d}$ so $V_H = Bvd$ [1]

$I = nAve$ or $v = \frac{I}{nAe}$ [1]

$V_H = Bvd = \frac{BdI}{nAe} = \frac{BI}{nte}$ (working shown) [1]

c i $v = \frac{I}{nAe}$

$= \frac{0.04}{8.5 \times 10^{28} \times 3.0 \times 10^{-5} \times 9 \times 10^{-3} \times 1.6 \times 10^{-19}}$ [1]

$= 1.1 \times 10^{-5} \text{ m s}^{-1}$ [1]

ii $V_H = \frac{BI}{nte} = \frac{0.60 \times 0.04}{8.5 \times 10^{28} \times 3.0 \times 10^{-5} \times 1.6 \times 10^{-19}}$ [1]

$= 5.9 \times 10^{-8} \text{ V}$ [1]

d i Electrons move faster [1]
Larger magnetic force so larger voltage (or electric field) is needed to repel them or obtain equilibrium. [1]

ii Hall voltage V_H or value calculated is very small with metal [1]
Semiconductor has smaller value of n so larger Hall voltage V_H [1]

e Electrons are forced in opposite direction when strip is rotated through 180°. [1]
either V_H is maximum when plane and B-field are normal to each other.
or V_H is zero / minimum when plane and B-field are parallel to each other.
or Voltmeter reading or V_H depends on sine of angle between plane and B-field. [1]

Answers to EOC questions

Chapter 28

1 The most obvious demonstration is to show that the secondary coil is made of insulated wire, so no current can flow from the core to the secondary coil. (Alternatively, if you arrange for a small gap in the core, perhaps a piece of paper, then there is still an induced e.m.f. even though paper is an insulator. The e.m.f. will be reduced because the amount of flux in the core is reduced if there is not a complete circuit of iron.) [1]
An electrical current is induced because there is a change in the magnetic flux linking the secondary coil. This changing flux is caused by the changing current in the primary coil. [1]

2 **a** $\Phi = BA$ [1]
$= 20 \times 10^{-3} \times (5.0 \times 10^{-2})^2 = 5.0 \times 10^{-5}$ Wb [1]

b e.m.f. $E = \frac{\Delta(N\Phi)}{\Delta t}$ [1]
$= \frac{100 \times 5.0 \times 10^{-5}}{0.1}$ [1]
$= 5.0 \times 10^{-2}$ V [1]

3 $= \frac{\Delta(N\Phi)}{\Delta t} = Blv$ [1]
$= 5.0 \times 10^{-5} \times 40 \times 300$ [1]
$= 0.60$ V [1]

4 An induced current flowing in a mass of metal, such as the core of a transformer. [1]
Eddy currents are used in electromagnetic braking, for example in trains, where energy for the eddy currents comes from the kinetic energy of the train and the eddy currents themselves provide a braking force. [1]
Eddy currents are a disadvantage in the cores of transformers, where energy is wasted in raising the internal energy of the core. [1]

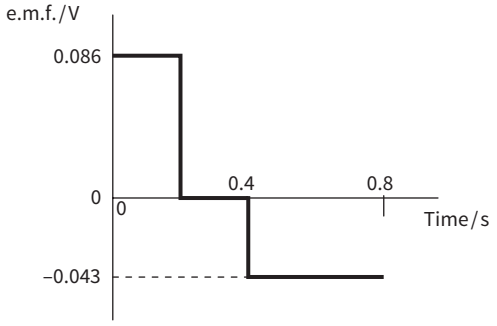
5 When there is no flux linkage, the flux is changing at the greatest rate and so the induced e.m.f. is a maximum. [2]
When the flux linkage is a maximum, it is, instantaneously, not changing and thus there is no induced e.m.f. [2]

6 **a** A magnetic flux linkage of 1 Wb exists if a coil of 1 turn and area 1 m^2 [1]
has a field of 1 T passing at right angles through it. [1]
(Other areas and numbers of turns are possible.)

b **i** $\frac{\Delta\Phi}{\Delta t} = \frac{0.6 \times 1.2 \times 10^{-4}}{0.2}$ [1]
3.6×10^{-4} Wb s^{-1} [1]

ii $E = \frac{\Delta(N\Phi)}{\Delta t} = 240 \times 3.6 \times 10^{-4}$ [1]
$E = 0.0864 \approx 8.6 \times 10^{-2}$ V [1]

iii E.m.f. correct and constant between 0 and 0.2 s and zero between 0.2 and 0.4 s [1]
Negative e.m.f. of half the value in **ii** between 0.4 and 0.8 s [1]



7 **a** **i** Time taken = $\frac{\text{distance}}{\text{speed}} = \frac{0.02}{0.05} = 4.0 \times 10^{-2}$ s [1]

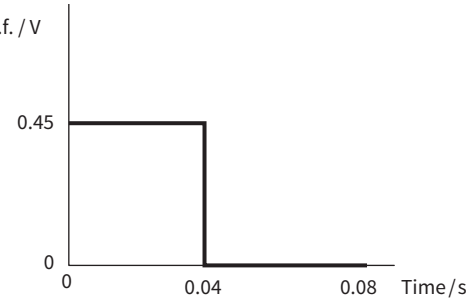
ii Flux linkage = NBA
$= 150 \times 0.30 \times (0.02 \times 0.02)$ [1]
flux linkage = 1.8×10^{-2} Wb [1]

b The rate of change of magnetic flux is constant. [1]

c Initial flux linkage = 0 and final flux linkage = 1.8×10^{-2} Wb [1]
Magnitude of induced e.m.f. = rate of change of magnetic flux linkage [1]
$E = \frac{1.8 \times 10^{-2} - 0}{4.0 \times 10^{-2}}$ [1]
$E = 0.45$ V (magnitude only) [1]

d When the coil is completely within the field, the induced e.m.f. is zero. [1]
The reason for this is that there is no change in the magnetic flux linkage. [1]

e



Correct axes and labels with e.m.f. constant between 0 and 0.04 s [1]
Zero e.m.f. between 0.04 and 0.08 s [1]

8 a The e.m.f. induced is proportional to the rate of change [1]
of magnetic flux linkage. [1]

b i $\Phi = BA = 50 \times 10^{-3} \times \pi \times (0.1)^2$ [1]
$\Phi = 1.57 \times 10^{-3} \approx 1.6 \times 10^{-3}$ Wb [1]

ii Change in flux linkage = $\Delta(N\Phi)$
$= 600 \times 1.57 \times 10^{-3} - 0$ [1]
$\Delta(N\Phi) = 0.942 \approx 0.94$ Wb (magnitude only) [1]

iii e.m.f. $E = \frac{\Delta(N\Phi)}{\Delta t} = \frac{0.942}{0.12}$ [1]
$E = 7.85 \approx 7.9$ V [1]

9 a Lenz's law states that the direction of induced e.m.f. / current opposes the change that is producing it. [1]
Plunge magnet into coil / magnet on spring oscillates in and out of coil [1]
either Ends of coil connected together [1]
experience force resisting movement of magnet / magnet oscillates slower when coil ends are connected together and induced current flows
or Use an ammeter / galvanometer in series with coil. [1]
Prove that, if north pole of magnet enters coil, then that end of coil is north [1]

b i An induced e.m.f. in the aluminium disc causes an induced current to flow. [1]
either Current in the disc causes heating of the disc / increase in internal energy of disc. [1]
By conservation of energy, the kinetic energy of the car is reduced. [1]
or Currents in the disc are flowing in the magnetic field created by the electromagnets. [1]
These currents experience a force backwards, opposing the rotation of the disc. [1]
(maximum of [3])

ii There is a greater rate of change of flux (in any area on the disc). [1]
A larger e.m.f. is induced by Faraday's law, thus larger induced currents flow. [1]

10 a i A change in magnetic flux causes an induced e.m.f. [1]
either The spokes are cutting magnetic lines of flux.
or The circuit (containing a spoke and the connections) is sweeping out an area of magnetic flux. [1]

ii Increase the strength of the magnetic field. This causes a greater magnetic flux linking the circuit and thus a greater rate of change of the magnetic flux. [1]
Rotate the coil faster. Each change in magnetic flux occurs in a smaller time, and thus a greater rate of change of flux occurs. [1]

b i Area per second = $\pi R^2 f = \pi \times (0.15)^2 \times 5$ [1]
Area per second = $0.353 \approx 0.35\ m^2\ s^{-1}$ [1]

ii $E = \frac{\Delta(N\Phi)}{\Delta t} = \frac{\Delta(BA)}{\Delta t} = 0.353 \times 5 \times 10^{-3}$ [1]
$E = 1.77 \times 10^{-3} \approx 1.8 \times 10^{-3}$ V [1]

Answers to EOC questions

Chapter 29

1 **a** $V = V_0 \sin \omega t$, where $\omega = 2\pi f$ [1]

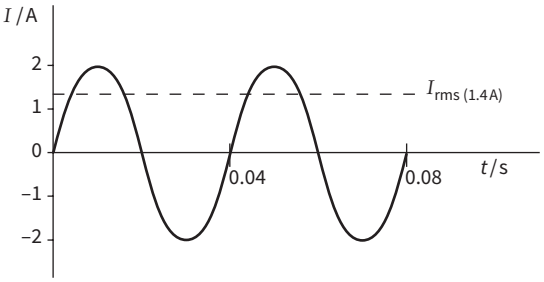
b $I = I_0 \sin \omega t$ [1]

c $P = I_0 V_0 \sin^2 \omega t = (I_0)^2 R \sin^2 \omega t = \frac{(V_0)^2 \sin^2 \omega t}{R}$ [1]

2 **a** 2 A [1]

b $2\pi f = 50\pi$ [1]
so $f = 25$ Hz [1]

c



[2]

d $I_{rms} = \frac{I_0}{\sqrt{2}} = 1.41 \approx 1.4$ A [1]

e 0.005 and 0.015 s in the first cycle, and 0.045 and 0.055 s in the second cycle [2]
I_{rms} marked on graph [1]

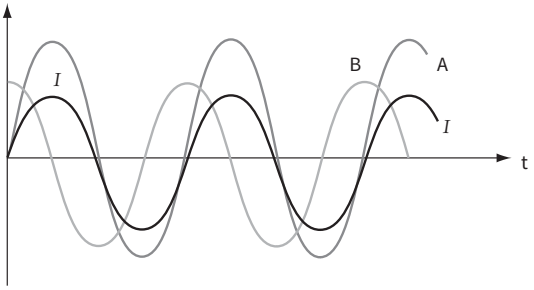
3 **a** $N = \frac{1200 \times 6.0}{240} = 30$ [1]

b **i** $P = \frac{(V_{rms})^2}{R} = \frac{6.0^2}{6} = 6.0$ W [1]

ii Assuming the transformer is 100% efficient, for primary coil $I_{rms} V_{rms} = 6.0$ W [1]
$I_{rms} = \frac{6.0}{240} = 0.025$ A [1]
$I_0 = I_{rms} \times \sqrt{2} = 0.025\sqrt{2} = 0.0353 \approx 0.035$ A [1]

4 **a** **i** Magnetic flux is in phase with the current. [1]

ii The induced e.m.f. and current are out of phase by 90°. [1]



b The two graphs differ in phase by 90° since the induced e.m.f. is the rate of change of magnetic flux linkage. [1]

5 **a** Period = 8 × 5 = 40 ms = 0.040 s [1]
Frequency = $\frac{1}{T} = 25$ Hz [1]

b $V_0 = 3 \times 0.5 = 1.5$ V [1]
$V_{rms} = \frac{1.5}{\sqrt{2}} = 1.06 \approx 1.1$ V [1]

c $I_{rms} = \frac{V_{rms}}{R} = \frac{1.06}{200} = 5.3 \times 10^{-3}$ A [1]

d $<P> = I_{rms} V_{rms} = 1.06 \times 5.3 \times 10^{-3}$ [1]
$= 5.6 \times 10^{-3}$ W [1]

6 **a** $I_0 = \sqrt{2} \times I_{rms}$ [1]

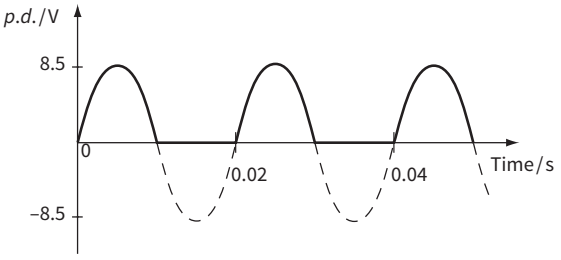
b **i** The current using the a.c. supply is not steady and is mostly below 2.0 A. [1]
The heating effect is greater using the d.c. supply. [1]

ii $\left\langle \frac{P_{dc}}{P_{ac}} \right\rangle = \frac{(I_{dc})^2}{(I_{rms})^2}$ [1]

$\left\langle \frac{P_{dc}}{P_{ac}} \right\rangle = 2$ [1]

c Transformers only work with alternating current. The use of transformers allows the mains voltage to be altered without much power loss to suit various devices that do not need such large voltages. [1]

7 **a** Correct dotted line showing a.c. [1]
Peak voltage $V_0 = \sqrt{2} \times 6.0 = 8.49 \approx 8.5$ V marked as peak on graph. [1]
Time for one cycle $\frac{1}{50} = 0.02$ s marked on graph. [1]
Correct half-wave rectified shape shown. [1]



b **i** $\frac{3}{4} \times 0.02 = 0.015$ s [1]

ii $I_{mean} = \frac{V_{mean}}{R} = \frac{5.7}{400}$ [1]
$= 0.014\,25 \approx 1.4 \times 10^{-2}$ A [1]

iii $Q = I_{mean} t = 0.014\,25 \times 0.015$ [1]
$= 2.14 \times 10^{-4} \approx 2.1 \times 10^{-4}$ C [1]

iv p.d. across capacitor falls from 8.49 to 4.24 V
$C = \frac{Q}{V_{final} - V_{initial}} = \frac{2.14 \times 10^{-4}}{8.49 - 4.24}$ [1]
$= 5.04 \times 10^{-5} \approx 5.0 \times 10^{-4}$ F [1]

Cambridge International AS and A Level Physics © Cambridge University Press 2014

c Diode has very large / infinite resistance [1]
when the p.d. across the diode is negative (i.e. it is reverse biased) or when top contact of supply is at a smaller potential than top plate of the capacitor. [1]

8 **a** Full-wave rectification occurs. [1]
Four diodes are used. [1]
Circuit showing correct bridge rectifier with four diodes, input and output (Figure 29.15 in the coursebook). [1]
Circuit with all diodes connected in correct direction. [1]
Correct explanation describing the two diodes that conduct when the supply voltage has positive potential at top and negative potential at bottom. [1]
Correct explanation describing the other two diodes that conduct when the supply voltage has negative potential at top and positive potential at bottom. [1]

b Capacitor charges up to 1.0 V [1]
When the p.d. across capacitor exceeds the supply p.d., the diode stops conducting. [1]
Capacitor provides current in resistor and discharges (exponentially). [1]

c **i** $t = 0.013$ to 0.015 s **or** 0.002 s [1]

ii $V_{mean} = 0.90$ to 0.92 V [1]

iii $\langle P \rangle = \frac{(V_{mean})^2}{R} = \frac{0.91^2}{1000}$ [1]

$= 8.28 \times 10^{-4} \approx 8.3 \times 10^{-4}$ W [1]

9 **a** Peak value of mains supply $= 230\sqrt{2} = 325$ V

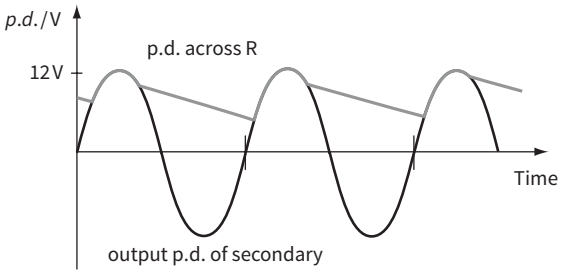
or r.m.s value of output $= \frac{12}{\sqrt{2}} = 8.49$ V [1]

$N_s = N_p \times \frac{V_s}{V_p} = 2000 \times \frac{12}{325}$ [1]

$N_s = 73.8 \approx 74$ [1]

b **i** Sine wave with correct peak value 12 V, at least 1.5 waves drawn [1]

ii Correct charging portion of p.d. across R [1]
Correct discharging portion of p.d. across R, approximately exponential [1]



c Smoothing capacitor now has twice the capacitance. [1]
Capacitor holds more charge and can supply the current / charge with less p.d. drop. [1]
Less 'ripple', i.e. the p.d. falls less during the discharge from the capacitor. [1]

10 **a** **i** The steady direct voltage that delivers the same energy / power [1]
to a resistance / resistive load as the alternating voltage. [1]

ii $V_0 = \sqrt{2} \times V_{rms} = \sqrt{2} \times 400\,000$ [1]

$= 566\,000 \approx 5.7 \times 10^5$ V [1]
or the p.d. between two conductors is twice this value

b **i** $I_{rms} = \frac{P}{V_{rms}} = \frac{500 \times 10^6}{400 \times 10^3}$ [1]

$= 1250 \approx 1.2 \times 10^3$ or 1.3×10^3 A [1]

ii $P = (I_{rms})^2 R = 1250^2 \times 4.0$ [1]

$P = 6.25 \times 10^6 \approx 6.2 \times 10^6$ **or** 6.3×10^6 W [1]

iii When the voltage is stepped up, the current is reduced (for the same input power). [1]
Smaller currents produced less energy / power loss in the line's resistance. [1]

c A changing current in the primary coil causes a changing magnetic flux in the iron core. [1]
This changing magnetic flux passes to the secondary coil where there are fewer turns and a smaller e.m.f. is induced. [1]

Answers to EOC questions

Chapter 30

1 $E = hf$ [1]
$= 6.63 \times 10^{-34} \times 4.0 \times 10^{18} = 2.7 \times 10^{-15}$ J [1]

2 For shortest wavelength:
$E = \frac{hc}{\lambda}$ [1]
$= \frac{6.63 \times 10^{-34} \times 3 \times 10^{8}}{0.005} = 4 \times 10^{-23}$ J [1]
so range is from 4×10^{-23} J to 4×10^{-25} J [1]

3 **a** $E = 1.02 \times 10^{-5} \times 1.60 \times 10^{-19} = 1.63 \times 10^{-24}$ J [1]

b $f = \frac{E}{h} = \frac{1.63 \times 10^{-24}}{6.63 \times 10^{-34}} = 2.46 \times 10^{9}$ Hz [1]

c $\lambda = \frac{c}{f} = \frac{3.0 \times 10^{8}}{2.46 \times 10^{9}} = 0.12$ m [1]

4 **a** $E = 5 \times 10^{6} \times 1.6 \times 10^{-19} = 8.0 \times 10^{-13}$ J [1]

b **i** 10 000 eV [1]

ii $E = 10\,000 \times 1.6 \times 10^{-19} = 1.6 \times 10^{-15}$ J [1]

c $E = \frac{6 \times 10^{-21}}{1.6 \times 10^{-19}} = 3.8 \times 10^{-2}$ eV [1]

5 **a** Charge = $2e$ [1]
so energy in eV = $2 \times 7500 = 15\,000$ eV [1]

b $E = 15\,000 \times 1.6 \times 10^{-19} = 2.4 \times 10^{-15}$ J [1]

c $v^2 = \frac{2E}{m} = \frac{4.8 \times 10^{-16}}{6.8 \times 10^{-27}} = 7.06 \times 10^{10}$ m^2 s^{-2} [1]

$v = 8.4 \times 10^{5}$ m s^{-1} [1]

6 **a** photon energy in eV = $\frac{2.5 \times 10^{-18}}{1.6 \times 10^{-19}} = 15.625$ eV [1]
Maximum energy
= photon energy – work function [1]
= 15.625 – 4.3 = 11.3 eV [1]

b Converting this to joules:
maximum energy
$= 11.3 \times 1.6 \times 10^{-19} = 1.8 \times 10^{-18}$ J [1]

7 Minimum frequency = $\frac{\text{work function}}{h}$ [1]
$= \frac{4.9 \times 1.6 \times 10^{-19}}{6.63 \times 10^{-34}} = 1.2 \times 10^{15}$ Hz [1]

8 **a** Energy required = 54.4 eV [1]
$= 54.4 \times 1.6 \times 10^{-19} = 8.7 \times 10^{-18}$ J [1]

b Energy change = 13.6 – 6.1 = 7.5 eV [1]
$f = \frac{E}{h} = \frac{7.5 \times 1.6 \times 10^{-19}}{6.63 \times 10^{-34}} = 1.8 \times 10^{15}$ Hz [1]
This lies in the ultraviolet region. [1]

c The drop in energy from $n = 2$ to $n = 1$ is much more than that from $n = 3$ to $n = 2$, [1]
so the frequency of the light emitted is much higher. [1]

9 **a** $E = \frac{hc}{\lambda} = \frac{6.63 \times 10^{-34} \times 3 \times 10^{8}}{590 \times 10^{-9}}$ [1]
$E = 3.4 \times 10^{-19}$ J [1]

b **i** Electrons have zero potential energy at infinity, and less than this near the nucleus. [1]

ii Electron is excited from the -5.8×10^{-19} J level to the -2.4×10^{-19} J level, [1]
absorbing all the energy of the incident photon. [1]

iii Incident light is directional; light is re-emitted in all directions. [1]

10 **a** Electron is totally removed from the nucleus. [1]
(accept atom is ionised)

b **i** The potential energy of the electron is less in level 2 than in level 3, [1]
so energy is given out (as a photon). [1]

ii Change in energy = –2.4 – (–5.4) = 3.0 eV **or** 4.8×10^{-19} J [1]
$E = \frac{hc}{\lambda}$ leads to $\lambda = \frac{hc}{E} = \frac{6.63 \times 10^{-34} \times 3 \times 10^{8}}{4.8 \times 10^{-19}}$ [1]
$= 4.1 \times 10^{-7}$ m [1]

iii If E is inversely proportional to n^2, then En^2 is a constant.
$n = 1$: $En^2 = -21.9 \times 1 = -21.9$
$n = 2$: $En^2 = -5.4 \times 4 = -21.6$
$n = 3$: $En^2 = -2.4 \times 9 = -21.6$ [1]
All products are approximately the same. [1]

Alternative route is to compare ratios of $E_1 : E_n$ and $\frac{1}{(n_1)^2} : \frac{1}{(n_n)^2}$ [1]
with linking comment. [1]

11 **a** **i** Electromagnetic radiation displays properties associated both with particles [1]
and with waves. [1]

ii Radiation below certain frequency will not produce photoelectrons / maximum energy of photoelectrons increases linearly with frequency. [1]

Energy of a packet (particle) depends on frequency (wave). [1]

b Energy of the photon is less than the work function, [1]
which is the minimum energy required to remove an electron from the metal surface. [1]

c Read off frequency = 5.6×10^{14} Hz [1]
$E = hf = 6.63 \times 10^{-34} \times 5.6 \times 10^{14}$ [1]
$= 3.7 \times 10^{-19}$ J [1]

d Gradient, e.g.
$\text{gradient} = \frac{(1.82 - 0) \times 1.6 \times 10^{-19}}{4.4 \times 10^{14}}$ [1]
$h = 6.62 \times 10^{-34} \approx 6.6 \times 10^{-34}$ J s [1]

12 **a** Electrons show wave–particle duality. [1]
The de Broglie wavelength is the wavelength of an electron (of a definite energy). [1]

b **i** 5.0 keV **or** 8.0×10^{-19} J [1]

ii $E = \frac{1}{2}mv^2 \rightarrow Em = \frac{1}{2}m^2v^2 \rightarrow mv = p = \sqrt{2Em}$ [1]
$p = \sqrt{2 \times 8.0 \times 10^{-19} \times 9.1 \times 10^{-31}}$ [1]
$= 3.8 \times 10^{-23}$ N s [1]

iii $\lambda = \frac{h}{p} = \frac{6.63 \times 10^{-34}}{3.8 \times 10^{-23}}$ [1]
$\lambda = 1.7 \times 10^{-11}$ m [1]

c Neutrons have a greater mass than electrons, so greater momentum [1]
for same energy, [1]
so shorter wavelength [1]

13 **a** Valence band electrons are not free to move / are bound to atoms [1]
conduction band above valence band [1]
valence band full and conduction band empty or only a few electrons. [1]

b $E = \frac{hc}{\lambda} = \frac{6.63 \times 10^{-34} \times 3 \times 10^{8}}{450 \times 10^{-9}} = 4.42 \times 10^{-19}$ [1]
$= \frac{4.42 \times 10^{-19}}{1.60 \times 10^{-19}} = 2.76$ eV [1]

c **i** Photons have enough energy [1]
to promote electrons from the valence band to the conduction band. [1]

ii Greater intensity of light causes more photons (per second) to hit the LDR. [1]
More electrons are promoted, the current is larger and resistance is smaller. [1]

d $\lambda = \frac{hc}{E} = \frac{6.63 \times 10^{-34} \times 3 \times 10^{8}}{2.4 \times 1.6 \times 10^{-19}}$ [1]
$= 5.18 \times 10^{-7}$ m **or** 520 nm [1]

14 **a** Metal: *any 3 from:*
Conduction and valence bands overlap. [1]
There are always electrons in conduction band. [1]
Number of electrons in conduction band does not change with temperature. [1]
Higher temperature increases lattice / atomic vibration, slowing movement of electrons / Reducing drift velocity. [1]

b Semiconductor: *any 3 from:*
At low temperatures, conduction band has few electrons. [1]
Band gap in energy exists between valence and conduction bands. [1]
Electrons need energy to go from valence band to conduction band. [1]
Higher temperature gives energy to electrons to cross band gap. [1]

Answers to EOC questions

Chapter 31

1 a Using $E = mc^2$ [1]
energy released $= 2 \times 1.67 \times 10^{-27} \times (3.0 \times 10^8)^2$ [1]
$= 3.0 \times 10^{-10}$ J [1]

b 1 mole contains N_A particles [1]
so energy released $= 3.0 \times 10^{-10} \times N_A$ [1]
$= 3.0 \times 10^{-10} \times 6.023 \times 10^{23} = 1.8 \times 10^{14}$ J [1]

2 Using $E = mc^2$, $m = \frac{E}{c^2}$ [1]
$= \frac{1.0}{(3.0 \times 10^8)^2} = 1.1 \times 10^{-17}$ kg [1]

3 Using $E = mc^2$, energy released per second
$= 70 \times 10^{-6} \times (3.0 \times 10^8)^2$ [1]
$= 6.3 \times 10^9$ J [1]
But energy per second = power,
so power = 6.3 GW [1]

4 Mass deficit = 221.970 – 217.963 – 4.002
= 0.005 u [1]
$= 0.005 \times 1.660 \times 10^{-27}$ kg
$= 8.30 \times 10^{-30}$ kg [1]
energy released $= 8.30 \times 10^{-30} \times (3 \times 10^8)^2$ [1]
$= 7.47 \times 10^{-13}$ J [1]
kinetic energy of the α-particle [1]
electromagnetic radiation (the γ-ray) [1]

5 a Mass defect in u
$= 6 \times (1.007\,276 + 1.008\,665 + 0.000\,548) - 12.000$
= 0.098 934 u [1]
Mass defect in kg $= 0.098\,934 \times 1.660 \times 10^{-27}$
$= 1.64 \times 10^{-28}$ kg [1]

b Binding energy = mass defect × c^2 [1]
$= 1.64 \times 10^{-29} \times (3.0 \times 10^8)^2 = 1.48 \times 10^{-11}$ J [1]

c number of nucleons = 12 [1]
so BE per nucleon $= \frac{1.48 \times 10^{-1}}{12} = 1.23 \times 10^{-12}$ J [1]

6 a Decrease in mass =
3.015 500 + 2.013 553 – 4.001 506 – 1.007 276 [1]
= 0.020 271 u [1]
$= 0.020\,271 \times 1.660 \times 10^{-27} = 3.365 \times 10^{-29}$ kg [1]

b Energy released = mc^2 [1]
$= 3.365 \times 10^{-29} \times (3.0 \times 10^8)^2 = 3.028 \times 10^{-12}$ J [1]

c energy released per mole
= energy per atom × N_A [1]
$= 3.028 \times 10^{-12} \times 6.023 \times 10^{23} = 1.823 \times 10^{12}$ J [1]

7 a 1 mole contains N_A atoms. [1]
Using $A = \lambda N$,
$\lambda = \frac{A}{N}$ [1]
$= \frac{8.02 \times 10^{21}}{6.023 \times 10^{23}} = 1.33 \times 10^{-2}\ s^{-1}$ [1]

b $\lambda t_{1/2} = 0.693$, so
$t_{1/2} = \frac{0.693}{\lambda}$ [1]
$= \frac{0.693}{1.33 \times 10^{-2}} = 52.0$ s [1]

8 a From the graph, $2 \times t_{1/2} = 28$ s [1]
so $t_{1/2} = \frac{28}{2} = 14$ s [1]
(**or** use the point (14, 80))

b $\lambda t_{1/2} = 0.693$, so
$\lambda = \frac{0.693}{t_{1/2}}$ [1]
$= \frac{0.693}{14} = 4.95 \times 10^{-2}\ s^{-1}$ [1]

9 a $t_{1/2} = \frac{0.693}{\lambda} = \frac{0.693}{4.9 \times 10^9}$ [1]
$= 1.4 \times 10^{-10}\ y^{-1}$ [1]

b Using $\ln\left(\frac{N}{N_0}\right) = -\lambda t$, [1]
$\ln 0.992 = -1.4 \times 10^{-10} t$ [1]
$t = 5.7 \times 10^7$ y [1]

10 a i Graph drawn using these figures, single smooth line, points plotted as crosses, suggested scales:
activity (*y*-axis)
50 Bq per cm
time (*x*-axis) 2 minutes per cm [1]

ii There is a random element of radioactive decay, which becomes more apparent at lower levels of activity. [1]

b *from the graph*, $t_{1/2}$ ~ 3.8 minutes [1]

c All count rates would be greater [1]
but the time for the rate to halve would remain the same. [1]

11 No change in count rate [1]
because temperature does not affect rate of radioactive decay. [1]

12 a 92 protons, 143 neutrons [1]

b $V = \frac{4}{3}\pi r^3 = \frac{4\pi(1.41\times10^{-15})^3\times235}{3} = 2.76\times10^{-42}\,\text{m}^3$ [1]

$r = \frac{3.89\times10^{-25}}{2.76\times10^{-42}}\,\text{m}^3$ [1]

$= 1.41\times10^{17} \approx 1.4\times10^{17}\,\text{kg m}^{-3}$ [1]

c When the nucleons are combined in the nucleus they have less energy (= binding energy) than when separated. [1]
Less energy means less mass (**or** energy has mass) [1]

d Sum the total mass of the separate protons and neutrons. [1]
Subtract from that the mass of the uranium nuclide. [1]
Apply $\Delta E = \Delta mc^2$ [1]
ΔE is the binding energy. [1]

13 **a** Nuclear fusion is the joining together of two (or more) light nuclei to form a heavier nucleus. [1]
The repulsive electrostatic forces between nuclei must be overcome. [1]
High temperatures mean particles moving very fast / high energy. [1]

b **i** Electrostatic forces larger as the charge on each nucleus is twice that on hydrogen (H). [1]

ii $\Delta m = 12 - (3\times4.001\,506) = 0.004\,518\,\text{u}$ [1]
$\Delta m = 0.004\,518\times1.660\times10^{-27}$
$= 7.500\times10^{-27}\,\text{kg}$ [1]
$\Delta E = 7.500\times10^{-27}\times(3.0\times10^8)^2$
$\Delta E = 6.75\times10^{-13}\,\text{J}$ [1]

14 **a** Alpha-particles have a very low penetration and those from outside the body are stopped by the layer of dead skin cells. [1]
Dust can be inhaled, bringing the α-particles inside the body, where they are very dangerous. [1]

b $N = \frac{2.4}{218\times6.02\times10^{23}}$ particles $(= 6.63\times10^{21})$ [1]

$\lambda = \ln\frac{2}{183} = 3.78\times10^{-3}\,\text{s}^{-1}$ [1]
$A = \lambda N = 3.78\times10^{-3}\times6.63\times10^{21}$ [1]
$A = 2.50\times10^{19}\,\text{Bq}$ [1]

c $\lambda t = \ln\left(\frac{A_0}{A}\right) = \ln\left(\frac{2.50\times10^{19}}{10}\right)$ $(= \ln(2.50\times10^{18}))$ [1]

$t = \ln\left(\frac{2.50\times10^{18}}{3.78\times10^{-3}}\right)$ $(= 11\,200\,\text{s})$ [1]
$= 3$ hours [1]

15 **a** Kinetic energy [1]
of the fission fragments [1]

b **i** Mass defect
$= 3.90\times10^{-25} - (1.44\times10^{-25} + 2.42\times10^{-25} + 1.67\times10^{-27}\times2)$
$\approx 0.01\times10^{-25}\,\text{kg}$ [1]
(accept a mass defect of $6.6\times10^{-28}\,\text{kg}$)
Energy released $= 0.01\times10^{-25}\times(3\times10^8)^2$
$= 9\times10^{-11}\,\text{J}$ [1]

ii Fissions per second $= \frac{200\times10^6}{9\times10^{-11}} = 2.2\times10^{18}$ [1]

iii Fissions per year $= 2.2\times10^{18}\times3.15\times10^7$
$= 6.93\times10^{25}$ [1]
moles per year $= \frac{6.93\times10^{25}}{6.02\times10^{23}} = 115$ [1]
mass per year $= 115\times235$
$= 2.7\times10^4\,\text{g} = 27\,\text{kg}$ [1]
(accept a final answer of 41 kg if a mass defect of $6.6\times10^{-28}\,\text{kg}$ is used)

16 **a** Emission of α-particles, β-particles and γ-particles [1]
from unstable nuclei [1]

b It is impossible to predict when a particular nucleus will decay. [1]
The rate of decay of many nuclei varies above and below an average value. [1]

c The decay of a nucleus is not affected by other nuclei around it. [1]
The decay of a nucleus is not affected by external conditions such as temperature. [1]

Answers to EOC questions

Chapter 32

1 **a** Radiation that can cause ionisation of the air (or of any other material it passes through). [1]
It is likely to have sufficient energy to cause damage to DNA and hence cell mutation. [1]
b X-ray shadow imaging, CAT scan [1]

2 Minimum wavelength $= \frac{hc}{E}$ [1]
$= \frac{6.63 \times 10^{-34} \times 3 \times 10^{8}}{20\,000 \times 1.6 \times 10^{-19}} = 6.2 \times 10^{-11}$ m [1]

3 There is a large difference between the acoustic impedance of air and skin.
Consequently, a very large percentage of the ultrasound is reflected. [1]
The gel is used to match the impedances. [1]

4 $\frac{I_r}{I_0} = \frac{(Z_2 - Z_1)^2}{(Z_2 + Z_1)^2}$ [1]
$= \frac{(1.78 - 1.63)^2}{(1.78 + 1.63)^2}$ [1]
$= 0.0019 = 0.19\%$ [1]

5 Wavelength $= \frac{\text{speed}}{\text{frequency}} = \frac{5200}{800\,000} = 0.0065$ m [1]
Optimum thickness is half this, i.e.
3.25×10^{-3} m $\approx 3.3 \times 10^{-3}$ m [1]

6 The angular frequency (or velocity) of the precession of a proton's magnetic axis around the direction of the applied field. [1]

7 Large (relatively) exposure of patient to the ionising radiation with consequent risks [1]
Expensive, because sophisticated equipment is required [1]

8 **a** Diagram showing basic tube with cathode, anode (target) [1]
Cathode, anode, vacuum labelled / referred to elsewhere [1]
High potential (50 kV or more) between anode and cathode [1]
Reference to material for anode / reference to cooling of anode [1]
Reference to acceleration of electrons across the tube [1]
b Electric potential energy
→ kinetic energy of electrons [1]
→ internal energy on collision with anode [1]
plus X-radiation [1]

9 **a** **i** Peaks are formed by excitation of an inner electron in an atom to a higher level. [1]
Energy is given out when it drops back to ground state. [1]
ii Band is produced by many interactions, each absorbing different energies. [1] [1]
b **i** Low-energy X-rays are absorbed by flesh so do not contribute to the overall picture. [1]
Filtering these rays reduces the patient's overall exposure to radiation. [1]
ii Aluminium absorbs these low-energy X-rays. [1]
c Use of 115 keV [1]
$E = hf$ leading to
$f = \frac{115 \times 10^{3} \times 1.6 \times 10^{-19}}{6.63 \times 10^{-34}}$ [1]
$= 2.78 \times 10^{18} \approx 2.8 \times 10^{18}$ Hz [1]

10 **a** *Any 5 points for [5] marks:*
The patient lies in a vertical ring of X-ray detectors and the X-ray tube rotates around the ring. [1]
The body part is split up into tiny cubes called voxels. [1]
A voxel is a unit of graphic information that defines a point in three-dimensional space. [1]
Slices through the body are viewed. [1]
The computer combines slices to build up a three-dimensional image of the patient. [1]
Images of 'slices' through the patient are shown on the computer screen. [1]
A traditional X-ray picture is formed by absorption. [1]
Denser materials absorb more than less dense materials, producing a shadow image. [1]
b CAT gives a better contrast of soft tissues. [1]
MRI gives better image of dense tissue such as bone. [1]

11 **a** **i** $I = I_0 e^{-\mu x}$ [1]
$= 4.0 \times 10^5 \times \exp(-250 \times 5 \times 10^{-2})$ [1]
$= 1.5\,W$ [1]

ii The long-wavelength / low-energy X-rays are absorbed, [1]
which do not contribute to the overall picture / are absorbed by soft tissue, [1]
reducing the overall exposure to radiation. [1]

b *Any 5 points for [5] marks:*
X-rays are scattered as they pass through the body. [1]
This scattering reduces the sharpness of the image. [1]
An anti-scatter grid consists of plates that are opaque to X-rays, [1]
which alternate with transparent material. [1]
The grid is placed between the patient and the screen. [1]
Grid absorbs scattered X-rays. [1]

12 **a** Acoustic impedance of a material
= density × speed of sound in the material [1]
When ultrasound goes from one material to another, some reflects. [1]
The proportion of the signal reflected is determined by the acoustic impedance of the two materials. [1]

b $Z = rc = 1.04 \times 10^3 \times 1.58 \times 10^3\,J$ [1]
$Z = 1.64 \times 10^6\,m^2\,s^{-1}$ [1]

c **i** Some energy is reflected at the first surface, so incident energy is lower at the second surface. [1]

ii Time between pulses $= 3.5 \times 10 \times 10^{-6}\,s$ [1]
Distance of travel of pulse $= vt$
$= 1.58 \times 10^3 \times 3.5 \times 10 \times 10^{-6}\,s$ [1]
Diameter of head $= \frac{1}{2} \times$ distance
$= 2.76 \times 10^{-2} \approx 28\,mm$ [1]

13 **a** The ratio of the Larmor frequency to the applied magnetic field [1]

b **i** $\omega_0 = 2\pi \times 48 \times 10^6 = 3.01 \times 10^8\,rad\,s^{-1}$ [1]
$\gamma = \frac{\omega_0}{B_0} = \frac{3.01 \times 10^8}{1.8}$ [1]
$= 1.67 \times 10^8\,rad\,s^{-1}\,T^{-1}$ [1]

ii The frequency of the RF signal is equal to the natural frequency of the proton precession, [1]
so the proton absorbs the energy. [1]

iii Protons fall back to a lower energy state, [1]
emitting (RF) radiation as they do so. [1]
The different relaxation times (**or** words to that effect) from different tissues enable a picture to be built up. [1]

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Answers to EOC questions

Chapter P2

1 **a** Dependent variable is frequency; independent variable is volume [1]

b Temperature (or air pressure) [1]

c Loudspeaker connected to signal generator **or** tuning forks of different frequencies [1]

d Connect to a microphone and oscilloscope. Measure T, the time for one wave on x-axis and $f = \frac{1}{T}$ [1]

e Fill with water and tip water into measuring cylinder [1]

f Add water [1]

g Not too loud **or** use ear plugs [1]

2 **a** Dependent variable is terminal velocity; independent variable is bubble volume **or** radius [1]

b Temperature / mass of air in bubble / atmospheric pressure [1]

c Place marks on side of glass tube in which bubble rises and time bubble between the marks [1]

d Squirt air from a small glass tube under water and change the radius of the tube [1]

3 **a** Dependent variable is count rate; independent variable is distance [1]

b Amount of radioactive material / atmospheric pressure / long half-life [1]

c Wrap source in a few millimetres of aluminium and subtract background count [2]

d Diagram of source and GM tube, with GM tube connected to counter [1]
Record change in reading on counter and divide count by time [1]

e Record for as long as possible (but much less than half-life) [1]

f The source is quite large and it is difficult to know which part to measure to / difficult to put a ruler over the distance [1]
Put a fiducial mark on the bench on the outside of the source to mark its middle [1]

g Use lead absorbers / stand far away / stay close for a short time [1]

4 **a** Dependent variable is volume / radius / diameter; independent variable is atmospheric pressure [1]

b Temperature / type of balloon [1]

c Bell jar **or** container to hold balloon [1]
Means to change pressure – e.g. vacuum pump [1]

d Turn on vacuum pump [1]
Pressure gauge shown or manometer (pressure may be measured in cm of liquid) [1]

e Wear goggles / stand behind safety screen (as bell jar may implode) [1]

5 **a** 6.0 ± 0.7 [1]

b 1.5 ± 0.2 [1]

c 9.0 ± 1.2 **or** 9 ± 1 [1]

d 1.0 ± 0.3 [1]

e 5.0 ± 1.6 **or** 5 ± 2 [1]

f 1.73 ± 0.06 [1]

6 Best-fit line – an even balance of points about the line along its whole length [1]
Worst-fit line – line with most different gradient to best-fit line but that still passes through as many error bars as possible [1]

In questions **7**–**11** only brief descriptions are given.

7 **a** dependent variable R [1]
independent variable d [1]
constant – intensity of bulb [1]

b circuit diagram [1]
current and voltage measured [1]
distance d between source and LDR changed [1]
d measured and shown [1]
method of measuring d [1]

c $R = \frac{V}{I}$ [1]
graph of $\ln R$ against $\ln d$ should be a straight line (not through origin) [1]

d avoid touching hot lamp / do not stare at bright light [1]

e *any [4] points from*
calculation of currents and choice of sensible meter ranges [1]
lamp is a line filament [1]
detail of how to measure distance to filament inside bulb [1]
repeat measurement and average [1]
trial experiment to find sensible distances [1]
avoidance of outside light / reflections [1]

8 a dependent variable T [1]
independent variable l [1]
constant – mass [1]

b displace mass downwards and release [1]
method of measuring l using rule [1]
method to measure to centre of mass [1]
time 10 oscillations and divide by 10 [1]
use stopwatch [1]

c graph of ln T against ln l [1]
gradient of graph n [1]

d avoid masses falling on foot / mass flying off / avoid breaking ruler with large amplitudes [1]

e *any [4] points from*
use small amplitudes [1]
trial experiment to find sensible value of mass [1]
use of electronic timer [1]
fiducial aid [1]
measure time from centre of oscillation [1]

9 a dependent variable B [1]
independent variable I [1]
constant – coil turns and area [1]

b diagram showing coil and Hall probe with voltmeter [1]
coil connected to d.c. power supply [1]
ammeter to measure I [1]
probe at right angles to direction of magnetic field [1]
method to locate centre of coil [1]

c graph of B against I **or** graph of ln B against ln I [1]
straight line through origin **or** graph of ln B against ln I has slope 1 [1]

d large current causes heating, so switch off when not in use [1]

e *any [4] points from:*
method to create a large magnetic field [1]
reasoned method to keep probe in same orientation (e.g. set square, fix to rule) [1]
calibrate Hall probe [1]
repeat experiment with probe reversed [1]
avoid external magnetic fields [1]

10 a dependent variable E [1]
independent variable v [1]
constant – current in coil [1]

b labelled diagram with magnet falling vertically through coil [1]
voltmeter or c.r.o. connected to coil [1]
method to change speed of magnet [1]
measurements to find v, distance or time measurements [1]
determining v, e.g. $v = \sqrt{2gh}$ or $v = \frac{2h}{t}$ [1]

c plot graph of E against v **or** graph of ln E against ln v [1]
valid if straight line through origin **or** gradient of log graph = +1 [1]

d keep away from falling magnet / use sand tray to catch magnet [1]

e *any [4] points from:*
detailed measurement of maximum E, use of datalogger / storage oscilloscope [1]
use video including playback frame by frame [1]
use short magnet or thin coil so v is constant [1]
cardboard tube for magnet to fall down [1]
how to support coil or tube vertically [1]
repeat for each v and average [1]

11 a dependent variable V_S [1]
independent variable f [1]
constant – number of turns / input current / voltage to coil [1]

b labelled diagram with a.c. generator [1]
voltmeter connected to output [1]
c.r.o. or frequency meter connected [1]
measurement of time for one or more wave on c.r.o. screen used to find f [1]
method of measuring V_S; method of changing frequency [1]

c plot graph of V_S against f **or** graph of ln V_S against ln f [1]

valid if straight line through origin **or** gradient of log graph = +1 [1]

d heating of coil, e.g. switch off when not in use to avoid overheating coil / do not touch [1]

e *any [4] points from:*
choice of number of turns to give reasonable output [1]
$f = \frac{1}{\text{period}}$ [1]
V_S found using *y*-gain [1]
changing r.m.s. to peak from voltmeter [1]
check that input voltage or current is constant [1]
repeat and average [1]

12 **a** gradient = $\frac{4\pi^2}{g}$ [1]

b **i**

T / s	T^2 / s^2
1.11	1.23 ± 0.02
1.28	1.64 ± 0.03
1.42	2.02 ± 0.03
1.59	2.53 ± 0.03
1.69	2.86 ± 0.04
1.81	3.28 ± 0.04

[2]

ii all points plotted correctly [1]
all error bars correct [1]

iii line of best fit [1]
worst acceptable straight line [1]

iv 4.1 s^2 m^{-1} uncertainty between ±0.1 and ±0.3 [2]

v 9.6 m s^{-2} uncertainty between ±0.2 and ±0.6 [2]

vi using $g = 9.6 \pm 0.2$ m s^{-2} gives $t = 19.2 \pm 0.2$ s [2]

13 **a** gradient *n*; *y*-intercept lg *k* [2]

b

lg(T / K)	lg(R / W)
2.436	2.740 ± 0.008
2.452	2.681 ± 0.009
2.467	2.625 ± 0.010
2.481	2.568 ± 0.012
2.496	2.519 ± 0.014

[2]

c **i** all points plotted correctly [1]
all error bars correct [1]

ii line of best fit [1]
worst acceptable straight line [1]

iii –3.7 ± 0.2 [2]

iv Use of a point on the graph and $y = mx + c$ to find *c* about 12 [1]
Use of point on worst line graph to find value of *c* about 12 ± 1 [1]

v $n = -3.7 \pm 0.2$ [1]
$k = 10^{\text{intercept}}$ [1]
Use of worst value of intercept to give worst value of *k* [1]

Recommended resources

Chapter 1

Further reading

Lang, H. (2008) Head First Physics. O'Reilly Media.
Contains explanations, examples and further reading about kinematics and displacement–time and velocity–time graphs in particular.

Kirkpatrick, L.D. and Francis, G.E. (2006) Physics: A World View, 6th edn. Cengage Learning.
Provides many different examples of vector addition as well as other topics you will encounter.

Online and interactive resources

www.physicsclassroom.com/class/vectors/Lesson-1/Relative-Velocity-and-Riverboat-Problems
Examples showing how to calculate relative velocity in a variety of problem situations.

www.launc.tased.edu.au/online/sciences/PhysSci/done/kinetics/grap_eqn/Grmotion.htm
Interactive resources for examining distance–time graphs and velocity–time graphs.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 2

Further reading

Collier, P. (2013) *A Most Incomprehensible Thing*. O'Reilly Media.
Explains equations of motion for constant acceleration in detail and also has some general reading around physics linking into relativity.

Lambourne, R. and Durrant, A. (2000) *Describing Motion: The Physical World*. Taylor & Francis.
Deals with many aspects of motion in physics. There is a detailed analysis of projectile motion. The book also demonstrates some useful mathematical tools.

Johnson, K., Hewett, S., Holt, S. and Miller, J. (2000) *Advanced Physics for You*. Nelson Thornes.
Aimed at A level students, this book explains equations of motion, velocity–time graphs and projectile motion using examples.

Online and interactive resources

phet.colorado.edu/sims/projectile-motion/projectile-motion_en.html
Projectile motion simulation that allows you to investigate varying angle and initial velocity to see how the motion of the projectile varies.

planetcalc.com/981/
This has a tool for entering three of the five parameters used in the equations of motion (displacement s, initial velocity u, final velocity v, acceleration a and time t) and it will calculate the others. You can use this to check your calculations using the equations of motion.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 3

Further reading

Zimba, J. (2009) *Force and Motion – An Illustrated Guide to Newton's Laws*. Johns Hopkins University Press.
An illustrated explanation of Newton's laws. The book looks in detail at concepts that students can find difficult, and irons out misconceptions and misinterpretations of the laws.

Jefferson, B. and Beadsworth, T. (2000) *Introducing Mechanics*. Oxford University Press.
Thorough explanation of Newton's laws of motion with a mathematical focus. Ideal for physics students who are also studying further maths to deepen their understanding of mechanics and Newton's laws.

Online and interactive resources

www.sciencechannel.com/games-and-interactives/newtons-laws-of-motion-interactive.htm
Interactive introduction to Newton's laws of motion with a quiz to test understanding.

teachertech.rice.edu/Participants/louviere/Newton/law1.html
Online lesson with simple animations showing each of Newton's laws of motion.

Video

www.space.com/23274-astronaut-uses-single-hair-to-show-newton-s-laws-video.html
An astronaut uses a single hair to demonstrate Newton's laws of motion.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 4

Further reading

Kirkpatrick, L.D. and Francis, G.E. (2006) *Physics: A World View*, 6th edn. Cengage Learning.
Detailed explanations of resolving vectors, calculating moments and the torque of a couple.

Johnson, K., Hewett, S., Holt, S. and Miller, J. (2000) *Advanced Physics for You*. Nelson Thornes.
Explains resolving forces, calculating moments and torques. States the conditions for a body to be in equilibrium.

Bird, J. and Ross, C. (2012) *Mechanical Engineering Principles*, 2nd edn. Routledge.
Extends the topic beyond A level; provides further reading and an in-depth study of the topic through examples.

Online and interactive resources

www.animatedscience.co.uk/as-chapter-07-forces-in-equilibrium
Summary of the principle of moments and notes on how to apply it to questions about balancing objects.

www.walter-fendt.de/ph14e/forceresol.htm
Animation showing how forces can be resolved into vectors.

Video

www.youtube.com/watch?v=cDly6AtOfMM&feature=youtu.be
Simple introduction to the idea of the principle of moments.

Recommended resources

Chapter 5

Further reading

Pople, S. (2001) *Advanced Physics Through Diagrams*. Oxford University Press.
An explanation of conservation of energy and kinetic and potential energy using diagrams.

Johnson, K., Hewett, S., Holt, S. and Miller, J. (2000) *Advanced Physics for You*. Nelson Thornes.
Explanation and further reading about the topics in this chapter, including defining the different types of energy and power.

Online and interactive resources

phet.colorado.edu/en/simulation/energy-skate-park
Applies the conservation of energy to a simulation of a skater in a skateboard park. The skateboard park can be on Earth, on different planets or in space.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 6

Further reading

Lambourne, D. and Durrant, A. (2000) *Describing Motion: The Physical World*. Taylor & Francis.
Explains conservation of momentum in one and two dimensions.

Johnson, K., Hewett, S., Holt, S. and Miller, J. (2000) *Advanced Physics for You*. Nelson Thornes.
Defines momentum and provides explanations and examples of the conservation of momentum in one and two dimensions.

Online and interactive resources

zebu.uoregon.edu/nsf/mo.html
Interactive resource on using and applying the conservation of linear momentum.

phet.colorado.edu/en/simulation/collision-lab
Interactive investigation of conservation of momentum in one and two dimensions.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 7

Further reading

Ohring, M. (1995) *Engineering Materials Science*. Academic Press.
Analysis of materials in terms of force constant and the Young modulus. Takes the reader beyond A level, providing further reading and a chance to delve deeper into the subject.

Carter, C. (2001) *Facts and Practice for A-level Physics*. Oxford University Press.
Notes on all the key topics from this chapter in concise form. This book contains many practice questions, which will reinforce key ideas.

Online and interactive resources

www.4physics.com/phy_demo/HookesLaw/HookesLaw.html
Virtual interactive Hooke's law experiment.

www.walter-fendt.de/ph14e/hydrostpr.htm
Interactive investigation of pressure in liquids, leading to the equation $p = \rho gh$.

www.matter.org.uk/schools/content/youngmodulus/experiment_3.html
Virtual experiment to find the Young modulus of a metal.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 8

Further reading

Kirkpatrick, L.D. and Francis, G.E. (2006) *Physics: A World View*, 6th edn. Cengage Learning.
Examines the uniform electric field between two parallel plates and explains the force on a charge.

Carter, C. (2001) *Facts and Practice for A-level Physics*. Oxford University Press.
Notes on all the key topics from this chapter in concise form. This book also contains many practice questions, which will reinforce key ideas.

Lang, H. (2008) *Head First Physics*. O'Reilly Media.
Contains explanations, examples and further reading about electric fields, forces on charges and the strength of electric fields.

Online and interactive resources

www.falstad.com/emstatic/
Choose the option 'charged planes' to investigate the electric field between two charged plates. Shows field lines and equipotential lines.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 9

Further reading

Bird, J. (2001) *Electrical Circuit Theory and Technology*. Routledge.
Further reading for students who want to find out more about electric circuits. Material from the chapter is covered, as well as more advanced circuit theory.

Carter, C. (2001) *Facts and Practice for A-level Physics*. Oxford University Press.
Notes on all the key topics from this chapter in concise form. This book also contains many practice questions, which will reinforce key ideas.

Online and interactive resources

phet.colorado.edu/en/simulation/circuit-construction-kit-dc
Interactive electric circuit. Use the circuit to check your understanding of current, voltage and resistance.

Video

www.youtube.com/watch?v=KgbqPKZU5IA
Explanation of drift velocity of electrons in circuits.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 10

Further reading

Bond, T. and Hughes, C. (2013) *GCE A Level Physics Oh-My-God Drill Questions and Solutions*. Yellowreef Ltd.
Questions with solutions on d.c. circuits, including Kirchhoff's laws, provide practice in solving d.c. circuit problems.

Carter, C. (2001) *Facts and Practice for A-level Physics*. Oxford University Press.
Notes on all the key topics from this chapter in concise form. This book also contains many practice questions, which will reinforce key ideas.

Online and interactive resources

www.flashscience.com/electricity/kirchhoff.htm
Kirchhoff's laws presented in a slightly different way to further your understanding.

labs.physics.dur.ac.uk/skills/skills/kirchhoff.php
Short notes on Kirchhoff's laws, with a link to animated examples of both laws in circuits.

Video

www.youtube.com/watch?v=Z2QDXjG2ynU
Walkthrough solution to a circuit problem using Kirchhoff's laws.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 11

Further reading

Johnson, K., Hewett, S., Holt, S. and Miller, J. (2000) *Advanced Physics for You*. Nelson Thornes.
Practice using the equations $V = IR$ and $R = \frac{\rho L}{A}$. Also provides information about thermistors and other semiconductors, including their I–V characteristics.

Carter, C. (2001) *Facts and Practice for A-level Physics*. Oxford University Press.
Practice using Ohm's law and resistivity equations.

Online and interactive resources

micro.magnet.fsu.edu/electromag/java/filamentresistance/
Animation explaining resistance in an ohmic conductor.

phet.colorado.edu/sims/ohms-law/ohms-law_en.html
Animation demonstrating Ohm's law.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 12

Further reading

Breithaupt, J. (2000) *New Understanding Physics for A Level*. Nelson Thornes.
Detailed explanation of e.m.f., internal resistance and potential dividers.

Carter, C. (2001) *Facts and Practice for A-level Physics*. Oxford University Press.
Provides practice using equations for e.m.f., internal resistance and potential dividers.

Online and interactive resources

people.clarkson.edu/~jsvoboda/eta/designLab/VoltageDividerDesign.html
Interactive resource which allows you to change the input resistors of a potential divider and predict the output voltage. The resource then displays the output voltage.

Video

www.youtube.com/watch?v=rcYqC2C9CeE
Clip explaining e.m.f. and internal resistance and how the behaviour of a circuit relates to the equation for e.m.f.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 13

Further reading

Ellse, M. and Honeywill, C. (2003) *Waves and Our Universe*. Nelson Thornes.
Explanation of transverse and longitudinal waves, relating them to 'real' waves.

Staelin, D.H. (1994) *Electromagnetic Waves*. Prentice Hall PTR.
Provides all the information needed at A level, but then takes the topic further for students who want to delve deeper.

Online and interactive resources

earthguide.ucsd.edu/eoc/special_topics/teach/sp_climate_change/p_emspectrum_interactive.html
An interactive electromagnetic spectrum that allows you to investigate energies and wavelengths of the different types of wave.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 14

Further reading

King, G.C. (2013) *Vibrations and Waves*. John Wiley & Sons.
Detailed explanations of the topics in this chapter including interference and diffraction.

Johnson, K., Hewett, S., Holt, S. and Miller, J. (2000) *Advanced Physics for You*. Nelson Thornes.
Detailed explanation of interference and diffraction and the superposition of waves.

Online and interactive resources

phet.colorado.edu/en/simulation/wave-interference
Simulation of waves passing through a single slit and a double slit, showing interference.

micro.magnet.fsu.edu/primer/java/diffraction/diffractionorders/index.html
Simulation showing the effect of wavelength (colour) on diffraction of light through a grating and linking to the diffraction grating equation in this chapter.

Recommended resources

Chapter 15

Further reading

Johnson, K., Hewett, S., Holt, S. and Miller, J. (2000) *Advanced Physics for You*. Nelson Thornes.
Explanations of the formation of stationary waves on strings.

Carter, C. (2001) *Facts and Practice for A-level Physics*. Oxford University Press.
Notes on all the key topics from this chapter in concise form. This book also contains many practice questions, which will reinforce key ideas.

Online and interactive resources

phet.colorado.edu/sims/wave-on-a-string/wave-on-a-string_en.html
Animation showing how standing waves can be formed on a string.

www.walter-fendt.de/ph14e/stlwaves.htm
Animation showing how standing waves are formed in air columns such as in wind instruments.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 16

Further reading

Johnson, K., Hewett, S., Holt, S. and Miller, J. (2000) *Advanced Physics for You*. Nelson Thornes.
Thoroughly covers the topic of radioactivity.

Kirkpatrick, L.D. and Francis, G.E. (2006) *Physics: A World View*, 6th edn. Cengage Learning.
Explains radioactivity in a different way to a standard textbook.

Online and interactive resources

hps.org/students/interactivesimulations.html
A wide range of animations on the theme of radioactivity. Includes simulations of alpha, beta and gamma decay and information about isotopes.

Video

www.youtube.com/watch?v=Z1IqewSDDcc
Very concise summary of particle physics covering all the points needed at A level.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 17

Further reading

Meyers, R. (2006) *The Basics of Physics*. Greenwood Publishing Group.
Very detailed analysis of circular motion. Further reading for those who want to delve deeper.

Lowe, T.L. and Rounce, J.F. (2002) *Calculations for A-Level Physics*. Nelson Thornes.
Practice using the many equations associated with this topic of physics.

Online and interactive resources

www.mhhe.com/physsci/physical/giambattista/circular/circular.html
Applet that shows the direction of centripetal force, acceleration and velocity of an object moving in a circle.

www.walter-fendt.de/ph14e/circmotion.htm
Animation showing the direction of velocity, force and acceleration for an object moving in a circle but also has calculations which can be confirmed by students.

Recommended resources

Chapter 18

Further reading

Johnson, K., Hewett, S., Holt, S. and Miller, J. (2000) *Advanced Physics for You*. Nelson Thornes.
Thorough explanation of gravitational fields and gravitational potential.

Pople, S. (2001) *Advanced Physics Through Diagrams.* Oxford University Press.
Uses diagrams, which lend themselves very well to explaining this topic.

Video

www.youtube.com/watch?v=Y50HeIUS4tk
Newton's law of gravitation explained with a simulation.

www.youtube.com/watch?v=zdQ54siEfvc
Gravitational fields revision clip.

Recommended resources

Chapter 19

Further reading

Young, C.Y. (2010) *Precalculus*. John Wiley & Sons.
Analysis of SHM and damping from a maths perspective for keen mathematicians.

Kirkpatrick, L.D. and Francis, G.E (2006) *Physics: A World View*, 6th edn. Cengage Learning.
SHM and damping explained thoroughly.

Online and interactive resources

phet.colorado.edu/sims/pendulum-lab/pendulum-lab_en.html
SHM covered using a pendulum as the oscillating object. Friction can be varied and the velocity and acceleration analysed as well as the displacement.

phet.colorado.edu/sims/mass-spring-lab/mass-spring-lab_en.html
Mass on a spring simulation. Vary the mass and the spring constant to investigate oscillations.

phet.colorado.edu/sims/resonance/resonance_en.html
Resonance and damping simulation with masses on springs.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 20

Further reading

Bissell, C.C. and Chapman, D.A. (1992) *Digital Signal Transmission*. Cambridge University Press.
Explains digital transmission from the original analogue waveform, through sampling, all the way to transmission.

Syngress (2001) *Designing a Wireless Network.* Syngress.
Covers this topic in detail. Applies the theory to setting up a real network. Considers the advantages and disadvantages of different methods of setting up networks.

Video

www.youtube.com/watch?v=W2-FP7twy8s
Introduces the basics of ADC and DAC using music sampling.

Recommended resources

Chapter 21

Further reading

Leach, J.A. (2007) *Engineering Thermodynamics*. Juta and Company Ltd.
University level text giving the definition of the first law of thermodynamics and examples of this law in practice.

Ellse, M. and Honeywill, C. (2004) *Electricity and Thermal Physics*. Nelson Thornes.
Notes and examples on the whole topic at A level standard.

Online and interactive resources

phet.colorado.edu/en/simulation/states-of-matter
Interactive applet on states of matter and kinetic theory.

eo.ucar.edu/skymath/tmp2.html
Web page summarising temperature and its measurement and linking temperature to kinetic theory.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 22

Further reading

Kirkpatrick, L.D. and Francis, G.E. (2006) *Physics: A World View*, 6th edn. Cengage Learning.
Provides explanations of ideal gas behaviour and practice using the ideal gas equations.

Carter, C. (2001) *Facts and Practice for A-level Physics*. Oxford University Press.
Brief notes on ideal gases and practice using the ideal gas equations. Also offers further explanations of the relationships between the quantities volume, pressure and temperature.

Online and interactive resources

phet.colorado.edu/en/simulation/gas-properties
Simulation for investigating the properties of ideal gases.

www.phy.davidson.edu/brownian.html
Brownian motion applet comparing Brown's observations and Einstein's explanation.

Recommended resources

Chapter 23

Further reading

Mansfield, M. and O'Sullivan, C. (2012) *Understanding Physics*, 2nd edn. John Wiley & Sons.
Concise yet thorough explanation of the key terms in this chapter and practice using the Coulomb's law and electric field equations.

Carter, C. (2001) *Facts and Practice for A-level Physics*. Oxford University Press.
Notes on all the key topics from this chapter in concise form. This book contains many practice questions, which will reinforce key ideas on the topic of electric fields.

Online and interactive resources

webphysics.davidson.edu/physlet_resources/bu_semester2/menu_semester2.html
Lots of mini-applets demonstrating the concepts in this chapter. A menu on the left allows you to select any that are of use.

phet.colorado.edu/en/simulation/charges-and-fields
Simulation showing the electric field between point charges and equipotential lines.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 24

Further reading

Breithaupt, J. (2000) *New Understanding Physics for A Level*. Nelson Thornes.
Explanation of capacitance and capacitors with the detail needed at this level.

Johnson, K., Hewett, S., Holt, S. and Miller, J. (2000) *Advanced Physics for You*. Nelson Thornes.
Explanation and further reading about the topics in this chapter, including explanations of how capacitors work, how charge is related to voltage and also the energy stored in a capacitor.

Online and interactive resources

phet.colorado.edu/en/simulation/capacitor-lab
Introduction to capacitors: how they work, what factors affect the amount of charge being stored, energy stored and direction of field.

Video

www.youtube.com/watch?v=CfloCOQ2_1U
Example calculation of a complex arrangement of series and parallel capacitors. Also includes revision of the capacitor equation, $Q = CV$.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 25

Further reading

Scherz, P. and Monk, S. (2013) *Practical Electronics for Inventors*, 3rd edn. Tab Electronics.
A book for students who want to know how to use and incorporate sensors into circuits in a variety of practical applications.

Mancini, R. (2012) *Op Amps for Everyone*. Newnes.
Further reading for students who want to find out more about op-amps and their uses in circuits. Contains A level theory and takes it much further.

Online and interactive resources

www.antonine-education.co.uk/Pages/ELectronics_1/Electronic_Components/Op-amp/intro_8.htm
Op-amps explained in more detail, and an example calculation using an op-amp as a voltage comparator.

www.electronics-micros.com/simulation/op-amp-comparator
Simple applet explaining how op-amps work.

Recommended resources

Chapter 26

Further reading

Muncaster, R. (1993) *A Level Physics.* Nelson Thornes.
Classic textbook explaining magnetic fields in detail.

Westbrook, C., Roth, C.K. and Talbot, J. (2011) *MRI in Practice*, 4th edn. Wiley-Blackwell.
This book is for professionals, including radiographers, radiologists and sales staff in the industry. Further reading for those who are interested in this area of physics.

Online and interactive resources

phet.colorado.edu/en/simulation/mri
This is a simulator in a java applet of an MRI scan. All the theory required at A level is included here.

Video

www.youtube.com/watch?v=HTTA30sEv6o
This clip explains the force on a current-carrying conductor in a magnetic field.

Recommended resources

Chapter 27

Further reading

Johnson, K., Hewett, S., Holt, S. and Miller, J. (2000) *Advanced Physics for You*. Nelson Thornes.
Classic explanations of how charged particles move in electric fields and in magnetic fields.

Carter, C. (2001) *Facts and Practice for A Level Physics*. Oxford University Press.
Covers all aspects of A level physics including a detailed analysis of the motion of charged particles in both electric and magnetic fields.

Online and interactive resources

www.magnet.fsu.edu/education/tutorials/java/halleffect
Interactive Java tutorial demonstrating the Hall effect.

Video

www.youtube.com/watch?v=XWkhUwX4D5s
This YouTube clip explains how to work out the force on a charge moving through a magnetic field.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 28

Further reading

Pople, S. (2001) *Advanced Physics Through Diagrams*. Oxford University Press.
Explains this topic with diagrams. Faraday's law and Lenz's law are both analysed in detail.

Bond, T. and Hughes, C. (2013) *GCE A-Level Physics Demanding Learn-By-Example*. Yellowreef.
Explanations of the topics in this chapter and example questions to test understanding.

Online and interactive resources

phet.colorado.edu/en/simulation/faraday
Move a magnet around near a coil to light a bulb. Learn more about Faraday's law with this applet.

phet.colorado.edu/en/simulation/generator
Investigate generators using this Java applet.

micro.magnet.fsu.edu/electromag/java/lenzlaw/
Applet demonstrating Lenz's law. Includes a written explanation and an animation showing how the direction of the induced current causes effects which oppose the change that is producing it.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 29

Online and interactive resources

www.britannica.com/EBchecked/topic/182915/electricity/71572/Alternating-current-circuits
Short but clear description of alternating circuits and the voltages and currents produced, as well as root-mean-square voltage and current.

www.school-for-champions.com/science/ac.htm#.UytsiNxcIds
Introduction to alternating current and the differences compared with direct current.

science.howstuffworks.com/environmental/energy/power.htm
Further explanation of how electrical power is distributed across a country; the focus is on the United States so references are to 120 V supply. Also explains three-phase alternating current, which goes into more detail than needed for A level but helps to explain practical aspects of power distribution.

Video

www.youtube.com/watch?v=2s6SvCJ-HOU
Further explanation of transformers and how they work.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 30

Further reading

There are many accessible books available that tackle the fundamentals of quantum physics in a non-mathematical way. Here is a short selection.

Hey, T. and Walters, P. (2003) *The New Quantum Universe*, 2nd edn. Cambridge University Press.
The first part deals with wave–particle duality and the evidence from experiments; the book goes on to explore de Broglie's matter waves and energy levels. Much of the book goes well beyond A level concepts, but is an informative and accessible primer for studies beyond A level.

Al-Khalili, J. (2012) *Quantum: A Guide for the Perplexed*. Phoenix.
A very wordy but friendly introduction to some of the key historical theories and experiments that led to the development of quantum theory. Particularly useful explanations of the double-slit experiment and Einstein's explanation of the photoelectric effect. Takes in several more advanced concepts.

Flynn, S. (2013) *The Science Magpie*. Icon Books.
An entertaining read comprising numerous short science stories; pages 142–145, 'The birth of the quantum', explains a little of the history behind Planck's equation.

Online and interactive resources

www.cyberphysics.co.uk/topics/atomic/Photoelectric%20effect/Photoelectric%20effect.htm
Detailed but accessible explanation of the photoelectric effect, with a number of videos included.

www.osram.co.uk/osram_uk/news-and-knowledge/led-home/professional-knowledge/led-basics/index.jsp
Information about LEDs, including how they work, the colours of light they produce, how they were developed and some of their uses.

Video

www.youtube.com/watch?v=TbFehcC4MHc
Clear explanation of the origins of line emission and absorption spectra. Uses as its starting point the concept of a black body emitting radiation, but quickly moves on to spectral lines observed in the light of the Sun, and spectra more generally.

www.nationalstemcentre.org.uk/elibrary/resource/2015/electron-diffraction-tube
Very clear description of the electron diffraction tube and a demonstration of it in use.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 31

Further reading

Muncaster, R. (1995) *Nuclear Physics and Fundamental Particles*. Nelson Thornes.
An older textbook but with detailed explanations and further exploration of nuclear physics.

Online and interactive resources

hps.org/students/interactivesimulations.html
Interactive simulations of alpha and beta decay, as well as illustrations of half-life.

www.miniphysics.com/2010/11/nuclear-physics.html
Comprehensive set of explanatory resources aimed at A level students, covering the main topics of nuclear physics including radioactive decay, half-life, binding energy and the relationship between mass and energy.

hyperphysics.phy-astr.gsu.edu/hbase/nucene/nucbin.html
Explanation and examples of binding energy calculations.

Video

www.youtube.com/watch?v=yTkojROg-t8
Explanation of binding energy and the stability curve, fission and fusion.

Cambridge International AS and A Level Physics © Cambridge University Press 2014

Recommended resources

Chapter 32

Further reading

Whitehouse, M. and Mould, R.F. (1995) *Beyond the Visible: One Hundred Years of X-rays*. Association for Science Education.
Explanations of the discovery of X-rays, their use in diagnostic imaging and other applications. Includes a number of activities for data analysis and comprehension.

Online and interactive resources

science.howstuffworks.com/cat-scan.htm
Simplified explanation of computerised axial tomography.

Video

www.nationalstemcentre.org.uk/elibrary/resource/7518/x-ray-imaging
Lecture explaining Roentgen's discovery of X-rays and the process of medical imaging using X-rays.

www.youtube.com/watch?v=AEel1PY_HoU
and
www.youtube.com/watch?v=FT2qygn4XjQ
Explanation of ultrasound scanning; the second video covers A- and B-scans.

www.youtube.com/watch?v=1CGzk-nV06g
Sixty-second explanation of magnetic resonance imaging.