NELSON PHYSICS UNITS 3 & 4

FOR THE AUSTRALIAN CURRICULUM

NEIL CHAMPION
ROBERT FARR
KATE WILSON

Key

atomic number → 26
Symbol of element: Fe
name of element → iron
standard atomic weight → 55.85

Symbol of element:
- gas at room temperature
- liquid at room temperature
- solid at room temperature
- synthetic (does not occur naturally)

- s block
- p block
- d block transition metals
- d block lanthanoids and actinoids

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 H hydrogen [1.007, 1.009]																	2 He helium 4.003
3 Li lithium [6.938, 6.997]	4 Be beryllium 9.012											5 B boron [10.80, 10.83]	6 C carbon [12.00, 12.02]	7 N nitrogen [14.00, 14.01]	8 O oxygen [15.99, 16.00]	9 F fluorine 19.00	10 Ne neon 20.18
11 Na sodium 22.99	12 Mg magnesium [24.30, 24.31]											13 Al aluminium 26.98	14 Si silicon [28.08, 28.09]	15 P phosphorus 30.97	16 S sulfur [32.05, 32.08]	17 Cl chlorine [35.44, 35.46]	18 Ar argon 39.95
19 K potassium 39.10	20 Ca calcium 40.08	21 Sc scandium 44.96	22 Ti titanium 47.87	23 V vanadium 50.94	24 Cr chromium 52.00	25 Mn manganese 54.94	26 Fe iron 55.85	27 Co cobalt 58.93	28 Ni nickel 58.69	29 Cu copper 63.55	30 Zn zinc 65.38(2)	31 Ga gallium 69.72	32 Ge germanium 72.63	33 As arsenic 74.92	34 Se selenium 78.96(3)	35 Br bromine [79.90, 79.91]	36 Kr krypton 83.80
37 Rb rubidium 85.47	38 Sr strontium 87.62	39 Y yttrium 88.91	40 Zr zirconium 91.22	41 Nb niobium 92.91	42 Mo molybdenum 95.96(2)	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 lanthanoids	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 [204.3, 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		

57 La lanthanum 138.9	58 Ce cerium 140.1	59 Pr praseodymium 140.9	60 Nd neodymium 144.2	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
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

This is a legal agreement between the you, (the "Customer") and Cengage Learning Australia Pty Limited (ABN 14 058 280 149) (the "Licensor") which provides the terms and conditions of this non-exclusive licence and the limited warranty for the Product. Use of the Product indicates an acknowledgement that the Customer has read and agreed to be bound by the terms and conditions of this Agreement. If you do not agree to these terms and conditions, return the Product to the place of purchase within 15 days of the date of purchase (with proof of purchase) for a full refund

1. **Licence Grant**
   You do not receive title to the Product. Copyright in the Product (which includes all images, photographs, video, animations, audio, music and text incorporated in the Product, including all of the accompanying printed material) is owned by the Licensor and/or its suppliers and is protected by Australian copyright laws. The Licensor grants you a non-exclusive licence to use the Product subject to the restrictions and terms set out in this Agreement.

2. **A Licence allows you to:**
   Use the Product on your computer. The Customer represents that they shall in no way place the Product in the public domain or in any way compromise our copyright in the Material. You agree to take reasonable steps to protect our copyright.

3. **You may not:**
   Alter, modify, translate, reverse engineer, decompile, or adapt the software or create derivative works based on the Product.

   Make further copies by any means technological, electronic, digital whatsoever without the written permission of the Licensor.

   Rent or transfer all or any part of your rights under this Agreement. Remove or alter any copyright or other proprietary notice or label attached to the software.

4. **Termination**
   Any failure to comply with the terms and conditions of this agreement will result in the automatic termination of this licence. Upon termination of this licence for any reason, the Customer must destroy or return to the Licensor all copies of the software and accompanying documentation.

5. **Warranties**
   To the extent permitted by law, the Licensor's liability for any breach of the warranty or any term implied by law into this licence is limited to the lowest cost of replacing the goods, acquiring equivalent goods or having the goods repaired.

This is a legal agreement between the you, (the "Customer") and Cengage Learning Australia Pty Limited (ABN 14 058 280 149) (the "Licensor") which provides the terms and conditions of this non-exclusive licence and the limited warranty for the Product. Use of the Product indicates an acknowledgement that the Customer has read and agreed to be bound by the terms and conditions of this Agreement. If you do not agree to these terms and conditions, return the Product to the place of purchase within 15 days of the date of purchase (with proof of purchase) for a full refund

1. **Licence Grant**
   You do not receive title to the Product. Copyright in the Product (which includes all images, photographs, video, animations, audio, music and text incorporated in the Product, including all of the accompanying printed material) is owned by the Licensor and/or its suppliers and is protected by Australian copyright laws. The Licensor grants you a non-exclusive licence to use the Product subject to the restrictions and terms set out in this Agreement.

2. **A Licence allows you to:**
   Use the Product on your computer. The Customer represents that they shall in no way place the Product in the public domain or in any way compromise our copyright in the Material. You agree to take reasonable steps to protect our copyright.

3. **You may not:**
   Alter, modify, translate, reverse engineer, decompile, or adapt the software or create derivative works based on the Product.

   Make further copies by any means technological, electronic, digital whatsoever without the written permission of the Licensor.

   Rent or transfer all or any part of your rights under this Agreement. Remove or alter any copyright or other proprietary notice or label attached to the software.

4. **Termination**
   Any failure to comply with the terms and conditions of this agreement will result in the automatic termination of this licence. Upon termination of this licence for any reason, the Customer must destroy or return to the Licensor all copies of the software and accompanying documentation.

5. **Warranties**
   To the extent permitted by law, the Licensor's liability for any breach of the warranty or any term implied by law into this licence is limited to the lowest cost of replacing the goods, acquiring equivalent goods or having the goods repaired.

NELSON PHYSICS UNITS 3 & 4

FOR THE AUSTRALIAN CURRICULUM

NEIL CHAMPION
ROBERT FARR
KATE WILSON

NELSON CENGAGE Learning

Australia • Brazil • Japan • Korea • Mexico • Singapore • Spain • United Kingdom • United States

NELSON
CENGAGE Learning®

Nelson Physics Units 3 & 4 for the Australian Curriculum
1st Edition
Neil Champion
Robert Farr
Kate Wilson

Publishing editor: Eleanor Gregory
Senior editor: Kelly Robinson
Editor: Marta Veroni
Proofreader: Sarah Endacott
Indexer: Rod Andrew
Cover design: Claire Atteia
Text design: Claire Atteia
Art direction: Luana Keays
Cover image: Getty Images/Tetra Images
Permissions researcher: Helen Mammides
Production controller: Julie McArthur
Typeset by: MPS Limited

Any URLs contained in this publication were checked for currency during the production process. Note, however, that the publisher cannot vouch for the ongoing currency of URLs.

For product information and technology assistance,
in Australia call **1300 790 853**;
in New Zealand call **0800 449 725**

For permission to use material from this text or product, please email
aust.permissions@cengage.com

National Library of Australia Cataloguing-in-Publication Data
Champion, Neil (Neil Douglas), author.
Nelson physics units 3 & 4 for the Australian curriculum /
Neil Champion, Robert Farr, Kate Wilson.

9780170242110 (paperback)
Includes index.
For secondary school age.

Physics--Textbooks.
Physics--Study and teaching (Secondary)
Physics--Problems, exercises, etc.

530

Cengage Learning Australia
Level 7, 80 Dorcas Street
South Melbourne, Victoria Australia 3205

Cengage Learning New Zealand
Unit 4B Rosedale Office Park
331 Rosedale Road, Albany, North Shore 0632, NZ

For learning solutions, visit **cengage.com.au**

Printed in China by China Translation & Printing Services.
1 2 3 4 5 6 7 19 18 17 16 15

CONTENTS

PREFACE

Nelson Physics Units 3 & 4 for the Australian Curriculum has been written to meet the requirements of the ACARA Australian Senior Secondary Curriculum – Physics. The text has been written to enable students to meet the A level Achievement Standard. It also allows all students to maximise their learning and results.

Physics deals with the wonderfully interesting and sometimes strange universe. Physicists investigate space and time (and space-time), from the incredibly small to the incredibly large, from nuclear atoms to the origin of the universe. They look at important, challenging and fun puzzles and try to work out solutions.

Physicists deal with the physical world where energy is transferred and transformed, where things move, where electricity and magnetism affect each other, where light and matter interact. As a result, physics has been responsible for about 95 per cent of the world's wealth – electricity supply and distribution, heating and cooling systems, computers, diagnostic and therapeutic health machines, telecommunications, safe road transport.

But physicists are not just concerned with observing the universe. They explain these observations, using models, laws and theories. Models are central to physics. Physicists use models to describe, explain, relate and predict phenomena. Models can be expressed in a range of ways – via words, images, mathematics (numerical, algebraic, geometric, graphical), or physical constructions. Models help physicists to frame physical laws and theories, and these laws and theories are also models of the world. Models are not static. As scientific understanding of concepts or physical data or phenomena evolves, so too do the models scientists use to describe, explain, relate and predict these. Thus, the text emphasises both the observations and quantitative data upon which physicists develop the models they use to explain the data. Central to this is the rigorous use of mathematical representations as a key element of physics explanations.

Nelson Physics Units 3 & 4 for the Australian Curriculum is written by academic and classroom teaching experts. They were chosen for their comprehensive knowledge of the physics discipline and best teaching practice in physics education at secondary and tertiary levels. They have written the text to make it accessible, readable and appealing to students. They have included numerous, current contexts to ensure students gain a wide perspective on the breadth and depth of physics. This contextual, mathematically rigorous and methodological approach is designed to ensure students can reach the highest possible standard. The intention has been to ensure all students achieve the level of depth and interest necessary to pursue tertiary studies in physics, engineering, technology and other scientific courses.

Each chapter follows a consistent pattern. Learning outcomes from the Science Understanding strand appear on the opening page. Learning outcomes from the Science as a Human Endeavour and Science Inquiry Skills strands are mapped on pages XII–XIII. The text is then broken into manageable sections under headings and subheadings. Relevant diagrams support the text. New terms are bolded and defined in a glossary at the end of the chapter and book. Important concepts are summarised in boxes to assist students to take notes.

Worked examples, written to connect important ideas and solution strategies, are included throughout the text. Solutions are written in full, including algebraic transformations, substitution of values with units, and a proposed marking scheme. In order to consolidate learning, students are challenged to try similar questions on their own.

Question sets appear at the end of logical sections. There is a comprehensive set of Review questions at the end of each chapter. All question sets have been graded from lower to higher order thinking skills: Remembering, Understanding, Applying, Analysing and Reflecting. Numerical answers appear at the end of the book. Complete worked answers appear on the NelsonNet teacher website.

Experiments and Investigations demonstrate the high level of importance the authors attach to understanding-by-doing physics. These activities introduce, reinforce and enable students to practise Science Inquiry skills, especially experimental design, data collection, analysis and conclusions. The Scientific Investigations chapter consolidates important investigative concepts and values. It enables students to learn and reflect on their experience as puzzle-solvers and investigators. It is an invaluable tool for students undertaking the Extended experimental investigation (EEI).

Système Internationale (SI) units and conventions, including accuracy, precision, uncertainty and error are introduced in the Measurement chapter. This invaluable tool supports student learning through chapter questions, experiments and investigations as well as their EEI.

Case studies elucidate Examples in context, part of the Science as a Human Endeavour strand. Scientific literacy activities ensure students develop the key Science Inquiry Skills capability of comprehending and evaluating scientific claims and synthesising a response.

Nelson Physics Units 3 & 4 for the Australian Curriculum provides students with a comprehensive study of modern physics that will fully prepare them for exams and any future studies in the area.

Neil Champion

Series editor

AUTHOR AND REVIEWER TEAMS

Authors

Neil Champion

Neil Champion BSc(Hons) DipEd MScEd was directly involved in writing the Australian Senior Physics Curriculum. He is an experienced secondary physics teacher at Buckley Park College in Victoria and has held numerous school leadership positions in State and Independent schools, including Principal. Neil has taught university level Physics and Physics Teaching Method for pre-service teachers. As VCAA Science Manager, he was responsible for the development and implementation of the worlds first senior secondary curriculum in photonics and synchrotron physics. For over 20 years, Neil has published school science texts for VCE Physics, IB Middle Years Program and 7–10 Australian Science Curriculum.

Rob Farr

Rob Farr has been teaching Physics in NSW schools for 30 years, 18 of those as Head of Science. He is currently teaching at Brigidine College St Ives in Sydney. He has extensive experience working with the Board of Studies in various roles reviewing syllabuses and HSC examinations. He has had 25 years of marking HSC examinations. Rob co-authored the successful *Physics in Focus* texts and he contributed to the *Biology in Focus* series, along with the *iScience for NSW* texts being used for the NSW syllabus based on the ACARA Australian Curriculum. Rob's ability to bring his enthusiasm and passion for Physics into his writing is evident in his work, making Physics accessible to students around the country.

Dr Kate Wilson

Dr Kate Wilson is a senior lecturer in the School of Engineering and Information Technology at UNSW Canberra (ADFA). She has a PhD in physics and a Graduate Diploma in Secondary Teaching (Science). Kate is a past director of the Australian Science Olympiads Physics Program. She has been a member of the Physics Education Research Group at the University of Sydney and held an Innovative Teaching and Educational Technology Fellowship at UNSW. She was first year physics coordinator at ANU for several years and has extensive experience in teaching first year university physics. Kate has done research into student learning in physics, particularly in laboratories and tutorials. She has published numerous research papers in the physics and education literature, as well as an award-winning first year university physics textbook and the Workshop Tutorial instructor resources. She has also designed and delivered several workshops and associated resources to support the NSW HSC Physics curriculum.

Reviewers

The publisher and authors would like to thank the reviewers for their valuable assistance in developing the chapter manuscripts.

- Lachlan Champion – Brighton Secondary College, Melbourne, Vic
- Charles Coleiro – Ave Maria College, Melbourne, Vic
- Dr Allan Ernest – Charles Sturt University, Wagga Wagga, NSW
- Dr Stephen Hughes – Queensland University of Technology, Brisbane, Qld
- Dr Geoffrey Pang – Central Queensland University, Rockhampton, Qld
- Dr Geoff Swan – Edith Cowan University, Perth, WA
- Associate Professor Marjan Zadnik – Curtin University, Perth, WA
- Dr Stephen Zander – Christ Church Grammar School, Perth, WA

Special thanks to Anna Wilson and Elisabetta Barberio for their advice on the Standard Model chapter.

Special thanks also to Wayne Hutchison and Bashir Ahmad at the University of New South Wales, Canberra, for their kind assistance with some photographs.

USING *NELSON PHYSICS*

Nelson Physics Units 3 & 4 for the Australian Curriculum has been purposely crafted to enable students to achieve maximum understanding and success in this subject. Each page has been carefully considered to provide students with all the information that they need without appearing cluttered or overwhelming. Students will find it easy to navigate through each chapter and see connections between chapters. Practical work has been integrated within the text so the students can see the interconnectedness between the theoretical and practical aspects of physics.

Each chapter begins with a **Chapter opener**. This presents the content descriptions from the Science Understanding strand of the senior Physics Australian Curriculum that will be incorporated into the chapter.

The text has been authored and reviewed by experienced Physics educators, academics and researchers to enable students to achieve the maximum level of achievement of which they are capable. A number of devices have been utilised to improve literacy and understanding. One of these is the use of shorter sentences and paragraphs. This is coupled with clear and concise explanations and real-world examples. New terms are bolded as they are introduced and appear in an end-of-chapter as well as an end-of-book glossary.

Important concept

Throughout the text, important ideas, formulas and laws are summarised in the **Important concept** box.

WORKED EXAMPLE

Mathematical representations and relationships are presented in context. Step-by-step instructions on how to perform mathematical calculations are shown in the **Worked examples**. The logic behind each step is explained and approximate marks allocated so that students can see that they need to show their full working out. Students are then able to practise these steps by attempting the related problems presented at the end of the worked example.

Physics is a practical subject and students need to be given the opportunity to explore and discover through practical activities. These are presented in three different types of boxes throughout the text.

ACTIVITY

The **Activities** provide the opportunity for short hands-on tasks to clarify or reinforce a concept. The activities can be performed either individually or in groups.

EXPERIMENT

The **Experiments** introduce and reinforce the Science Inquiry Skills strand of the Australian Curriculum. Experiments contain guided instruction on the materials, procedure, collection and analysis of results and discussion.

INVESTIGATION

The **Investigations** allow students to practice Science Inquiry Skills. They provide students with the opportunity to design and carry out their own scientific investigation either individually or in a group. Students are prompted to consider ideas for improvement and further investigation to illustrate that science is an ongoing and improving process. Further information on how to conduct a scientific investigation can be found in the Scientific Investigations chapter on page 327.

Risk assessment

The **Risk assessment** table occurs within the experiment and investigation boxes. The table highlights the risks to the students and provides suggestions on how to minimise these risks. Teachers are able supplement this table by adding any further risks specific to their school situation.

Scientific literacy

Many so called scientific claims are used to promote an issue, product or idea. It is important that students are able to understand and analyse the information presented to them so they can make well-informed choices. The **Scientific literacy** boxes present scientific texts or media articles that enable students to use evidence to evaluate the claims and conclusions presented. This allows them to use reasoning and knowledge beyond the information presented to construct a valid scientific argument.

Case study

Case studies provide students with the opportunity to see how science is applied using an up-to-date and real-world example in context.

Margin note

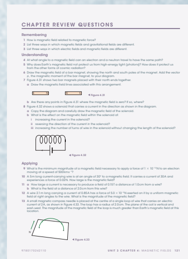
Full understanding of a concept is often constructed from many pieces of information. Due to the sequential nature of a book, this information cannot always be presented together as it is best placed in other chapters. Links between concepts that occur on other pages and chapters are indicated using **Margin notes**.

QUESTION SET

Review of student understanding is attained through the **Question sets** throughout each chapter. Questions are ordered from lower to higher order thinking skills. The addition of reflection questions gives students the opportunity to reflect upon not only what they are learning but why they are learning it, and how they are learning it.

The end of chapter review provides:

- a **Summary** of the important concepts presented within the chapter. This will be a valuable tool when students are revising for tests and exams
- a **Glossary** of all the new terms introduced within the chapter
- **Chapter review questions** that review understanding of concepts from the chapter. Questions are ordered from lower to higher order thinking skills and also include reflection questions.



Where answers to questions are numerical, they are provided in the back of the book.

NelsonNet

NelsonNet is your protected portal to the premium digital resources for Nelson textbooks located at www.nelsonnet.com.au. Once your registration is complete you will have access to an exciting and stimulating digital suite of resources for each chapter that are designed to enhance and reinforce learning.

Each chapter will also be supplemented with the following digital resources:

- **Prior learning activity sheet** to revise content from Year 11 that is a prerequisite to understanding
- **Activity sheets**, including theory and practical exercises
- **Revision sheets** to complete at home to revise class work
- A **review quiz** containing 20 auto-correcting multiple-choice questions to review understanding
- **Videos** to provide extra information or real-world examples. Pages that have videos associated with them are indicated with a blue icon in the footer
- **Links** to websites that contain extra information. These are hotspotted within the ebook and they can also be accessed at http://pac3and4.nelsonnet.com.au.

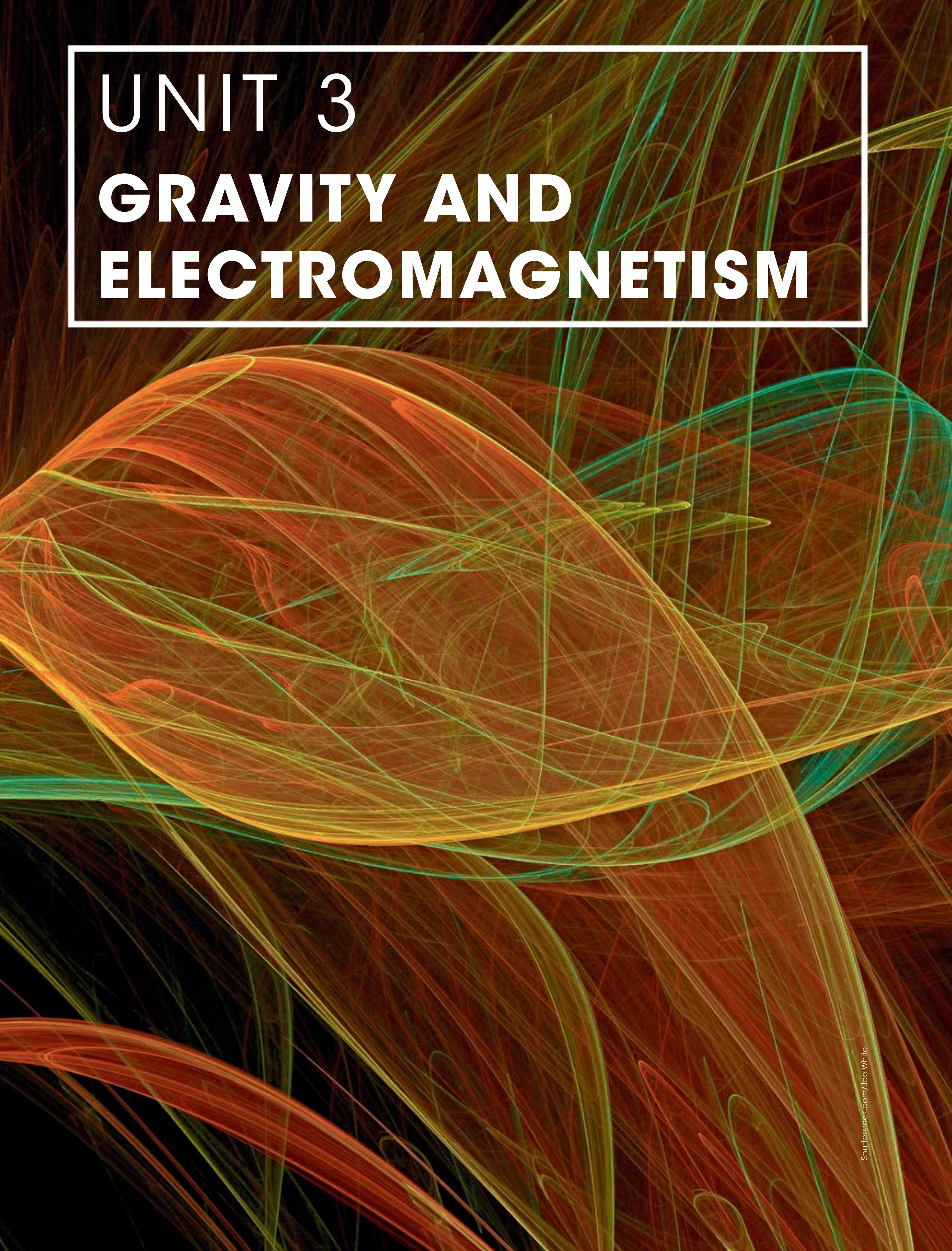
Please note that complimentary access to NelsonNet and the NelsonNetBook is only available to teachers who use the accompanying student textbook as a core educational resource in their classroom. Contact your sales representative for information about access codes and conditions.

CURRICULUM GRID

		Unit 3					Unit 4						
		Chapter											
		1	2	3	4	5	6	7	8	9	10	11	12
Science Inquiry Skills	Identify, research and construct questions for investigation; propose hypotheses; and predict possible outcomes (ACSPH078 AND ACSPH114)					✓							✓
	Design investigations, including the procedure to be followed, the materials required, and the type and amount of primary and/or secondary data to be collected; conduct risk assessments; and consider research ethics (ACSPH079 AND ACSPH115)					✓							✓
	Conduct investigations, including • the manipulation of force measurers and electromagnetic devices (ACSPH080) • use of simulations and manipulation of spectral devices (ACSPH116) safely, competently and methodically for the collection of valid and reliable data				✓	✓	✓	✓	✓	✓			✓
	Represent data in meaningful and useful ways, including using appropriate SI units, symbols and significant figures; organise and analyse data to identify trends, patterns and relationships; identify sources of uncertainty and techniques to minimise these uncertainties; utilise uncertainty and percentage uncertainty to determine • the uncertainty in the result of calculations (ACSPH081) • the cumulative uncertainty resulting from calculations (ACSPH117) and evaluate the impact of measurement uncertainty on experimental results; and select, synthesise and use evidence to make and justify conclusions				✓	✓		✓	✓			✓	✓
	Interpret a range of scientific and media texts, and evaluate processes, claims and conclusions by considering • the accuracy and precision of available evidence (ACSPH082) • the quality of available evidence (ACSPH118) and use reasoning to construct scientific arguments	✓	✓		✓	✓	✓						✓

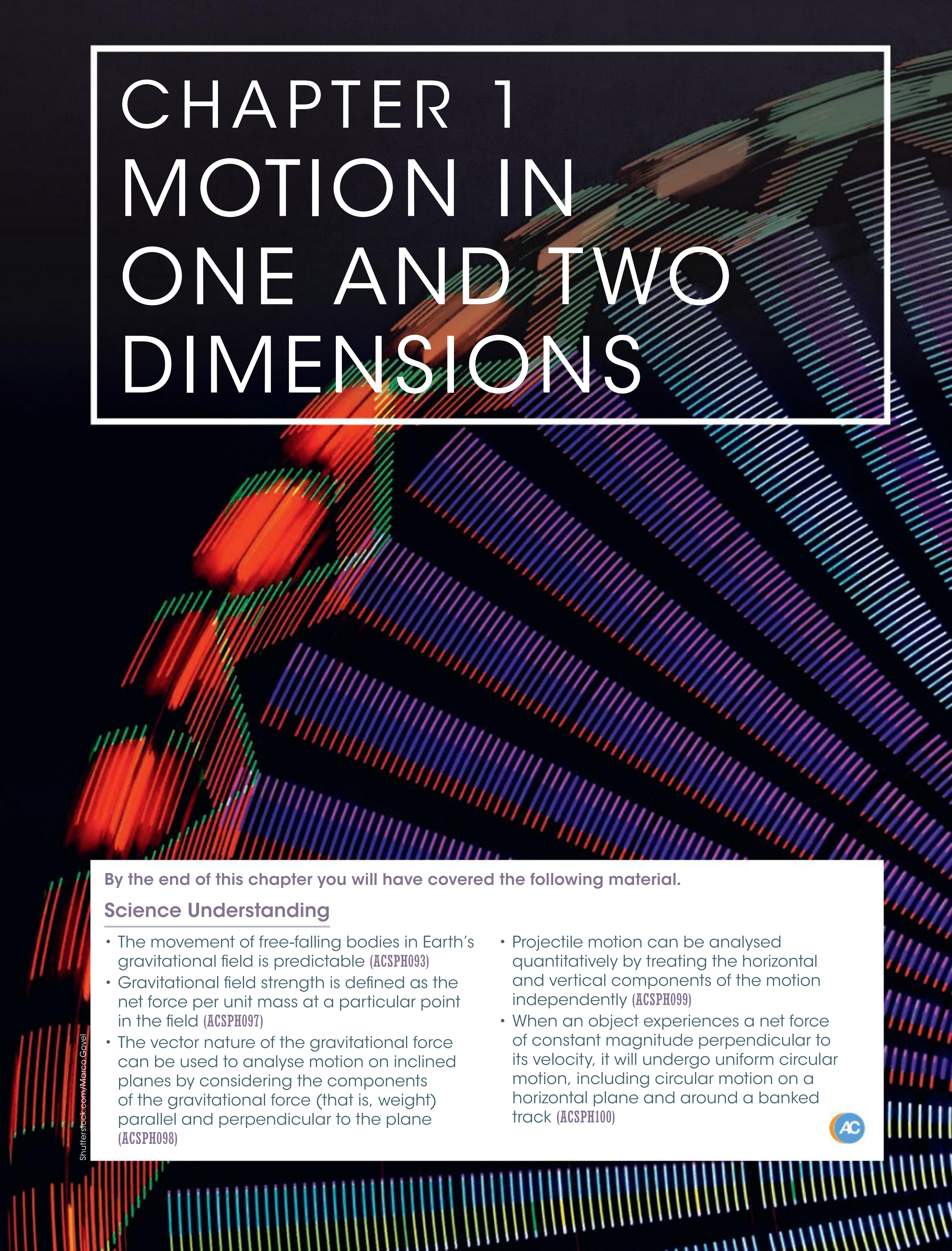
		Unit 3						Unit 4					
		Chapter											
		1	2	3	4	5	6	7	8	9	10	11	12
	Select, construct and use appropriate representations, including text and graphic representations of empirical and theoretical relationships												
	• vector diagrams, free body/force diagrams, field diagrams and circuit diagrams (ACSPH083)	✓		✓	✓	✓							
	• simulations, simple reaction diagrams and atomic energy level diagrams (ACSPH119)						✓	✓	✓	✓	✓		✓
	to communicate conceptual understanding, solve problems and make predictions												
	Select, use and interpret appropriate mathematical representations, including linear and non-linear graphs and algebraic relationships representing physical systems, to solve problems and make predictions (ACSPH084 AND ACSPH120)	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓
	Communicate to specific audiences and for specific purposes using appropriate language, nomenclature, genres and modes, including scientific reports (ACSPH085 AND ACSPH121)			✓		✓	✓		✓		✓	✓	✓
Science as a Human Endeavour	ICT and other technologies have dramatically increased the size, accuracy and geographic and temporal scope of datasets with which scientists work (ACSPH086 AND ACSPH122)	✓	✓			✓		✓			✓		
	Models and theories are contested and refined or replaced when new evidence challenges them, or when a new model or theory has greater explanatory power (ACSPH087 AND ACSPH123)		✓	✓	✓	✓	✓	✓	✓		✓		
	The acceptance of science understanding can be influenced by the social, economic and cultural context in which it is considered (ACSPH088 AND ACSPH124)		✓		✓			✓					
	People can use scientific knowledge to inform the monitoring, assessment and evaluation of risk (ACSPH089 AND ACSPH125)				✓			✓					
	Science can be limited in its ability to provide definitive answers to public debate; there may be insufficient reliable data available, or interpretation of the data may be open to question (ACSPH090 AND ACSPH126)					✓		✓	✓		✓		
	International collaboration is often required when investing in large-scale science projects or addressing issues for the Asia-Pacific region (ACSPH091 AND ACSPH127)				✓	✓		✓			✓		
	Scientific knowledge can be used to develop and evaluate projected economic, social and environmental impacts and to design action for sustainability (ACSPH092 AND ACSPH128)				✓			✓					

UNIT 3
GRAVITY AND ELECTROMAGNETISM



Shutterstock.com/Joe White

CHAPTER 1
MOTION IN ONE AND TWO DIMENSIONS



Shutterstock.com/Marco Govel

By the end of this chapter you will have covered the following material.

Science Understanding

- The movement of free-falling bodies in Earth's gravitational field is predictable (ACSPH093)
- Gravitational field strength is defined as the net force per unit mass at a particular point in the field (ACSPH097)
- The vector nature of the gravitational force can be used to analyse motion on inclined planes by considering the components of the gravitational force (that is, weight) parallel and perpendicular to the plane (ACSPH098)
- Projectile motion can be analysed quantitatively by treating the horizontal and vertical components of the motion independently (ACSPH099)
- When an object experiences a net force of constant magnitude perpendicular to its velocity, it will undergo uniform circular motion, including circular motion on a horizontal plane and around a banked track (ACSPH100)

See Chapter 7 in Nelson Physics Units 1 & 2 for the Australian Curriculum.

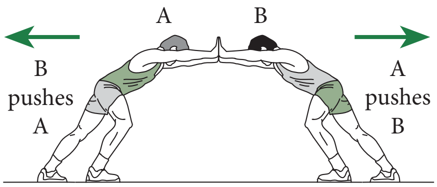


Figure 1.1 ▲ Person A contacts person B, so A acts on B: *F*(by A on B). Notice that at the same time B contacts A, so B acts on A: *F*(by B on A)

Introduction

A force is applied by one object on another (Figure 1.1).

If, in an interaction involving two objects A and B, object A is the external agent that acts on the receiver, B, we denote this by:

$$\vec{F}(\text{by A on B})$$

Simultaneously, B acts on A, so that, from this viewpoint, B is the agent and A the receiver:

$$\vec{F}(\text{by B on A})$$

This basic idea, that forces act on objects, leads to Newton's three laws:

Newton 1: If the **vector** sum of all forces acting on an object is zero, then its velocity remains constant.

Newton 2: The acceleration, $\vec{a}$, of an object is dependent on the vector sum of all forces acting on it and its mass, m:

$$\vec{a} = \frac{\Sigma F(\text{on object})}{m}$$

Newton 3: $\vec{F}(\text{by A on B})$ and $\vec{F}(\text{by A on B})$:

- are equal in magnitude
- are opposite in direction
- have the same fundamental nature

AND

- each force acts on a different object.

In *Nelson Physics Units 1 & 2 for the Australian Curriculum* we distinguished between contact and non-contact forces. In this chapter we now abandon the idea of contact force. Instead, we develop the more powerful concept of field and show how it is used to explain interactions between objects. Fields mediate the forces: object A affects object B by way of A's field stretching out beyond A, and B acts reciprocally on A via its own field.

Consequently, all forces may be considered to be **action-at-a-distance** effects. The gravitational field of a mass affects all other masses. Similarly, we shall see in chapters 3 and 4 that electrostatic and magnetic forces act at a distance from charges and magnets respectively.

Figure 1.2 shows a person standing on the surface of Earth.

The electrostatic force applied by the surface, $\vec{F}(\text{by S on P})$, acts upwards while the gravitational force on the person, $\vec{F}(\text{by } M_E \text{ on } m_P)$, acts downwards. The respective Newton 3 pairs for these forces are also shown. Only the forces acting on the person affect the person's motion.

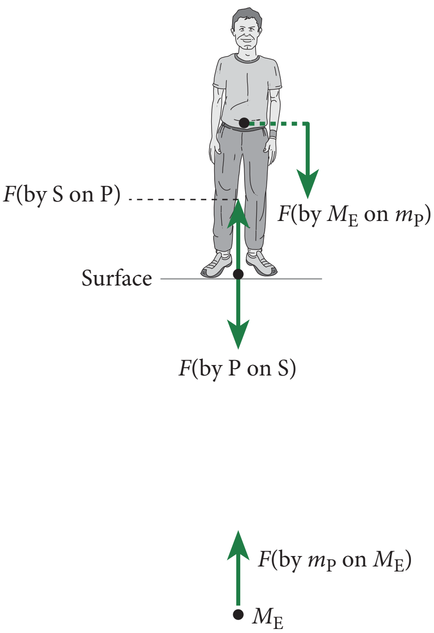


Figure 1.2 ▲ The feet and Earth's surface are in contact and act on each other, but the masses of Earth and the person act on each other from a distance

Movement of free-falling bodies in Earth's gravitational field

At any point near Earth's surface, an object experiences the effect of the mass of Earth. Newton was the first to realise that this same effect was the force that held the Moon in orbit around Earth and the planets in their orbits about the Sun.

NEWTON'S CONSIDERATION

Newton's consideration of how Earth's gravitational field emanates through space led him to the invention of integration.

Gravitational field near Earth's surface

Due to Earth's nearly-spherical shape, an object anywhere near Earth's surface is about the same distance from the centre of mass of Earth. This point can be taken as the point in space from where Earth's gravitational field emanates throughout the universe. In theory, this field will apply a force to every object in the universe. However, the great distances within our own solar system, galaxy and group of galaxies mean that, in reality, Earth's gravitational field really only has an influence on objects within a few hundred million kilometres.

Weight force near Earth's surface

The gravitational field strength around any mass is determined by the distance from the centre of mass and the mass itself. The force applied to any mass within the gravitational field is, in turn, determined by the strength of the gravitational field. Near Earth's surface, a 1 kg mass has a force of 9.8 N applied to it by Earth's gravitational field. The gravitational field strength is:

$$g = \frac{F(\text{by mass of Earth on mass } m)}{m} \text{ N kg}^{-1}$$

According to Newton's third law, Earth's mass and a 1 kg mass will apply the same magnitude of gravitational force on each other. These forces will be in opposite directions, along a line joining their centres of mass. The forces will apply to different masses: the 1 kg mass acts on Earth's mass, and Earth's mass acts on the 1 kg mass.

F(by mass of Earth on 1 kg mass) and F(by 1 kg mass on mass of Earth):
- are equal in magnitude.
- are opposite in direction.
- act on different things.
- cannot be added to make a net force.

If the 1 kg mass is released, the mass will fall to Earth and Earth will fall towards the 1 kg mass. Earth will, in fact, fall an immeasurably small distance, while the 1 kg mass will fall through the distance between it and the ground below.

Weight, w, is a force, the force applied by the mass of Earth on a nearby mass:

$$w = F(\text{by mass of Earth on nearby mass})$$

For any mass m, the magnitude of the force applied to it by Earth's gravitational field, the weight force w, is given by:

$$w = mg$$

where g is the gravitational field strength. The value of g near Earth's surface is 9.8 N kg^{-1}.

Gravitational acceleration

Newton's second law, $\vec{a} = \frac{\vec{F}_{net}}{m}$, can be applied to the gravitational force applied to any mass m near Earth's surface. Using $\vec{F}_{net} = w$, we get:

$$\vec{a} = \frac{\vec{F}_{net}}{m}$$

$$= \frac{w}{m}$$

so: $w = m\vec{a}$

we get: $\vec{a} = \vec{g}$

Therefore, the acceleration of a mass free to move near Earth's surface is 9.8 m s^{-2}.

As g is the acceleration due to gravity, the direction of the acceleration must be towards the centre of mass of Earth as well. This also tells us that the gravitational field, g, actually *is* the gravitational acceleration.

You may have already guessed this from the units. 1 N kg^{-1} is equal to 1 m s^{-2}, which is the unit for acceleration. (Recall that 1 N = 1 kg m s^{-2}, therefore 1 N kg^{-1} = (1 kg m s^{-2})(1 kg^{-1}) = 1 m s^{-2}.)

Beware of assuming that if two quantities have the same units they are the same thing. Although this is often the case, it is not always so, as we shall see when we look at torque. Torque can be written in the same units as work (Nm) but it is not the same thing.

This gives us a very simple way of measuring the gravitational field at any point in space. We simply have to ensure no forces other than gravity are acting on a test mass, and then observe its acceleration.

The gravitational field at any point is equal to the acceleration of a mass due to the gravitational force at that point, $g = \frac{F_{gravitational}}{m}$.

Close to Earth's surface:

$$g_{\text{close to Earth}} = 9.8\,\text{N}\,\text{kg}^{-1} = 9.8\,\text{m}\,\text{s}^{-2}$$

EXPERIMENT 1.1

EARTH'S GRAVITATIONAL FIELD STRENGTH

The period, T, of a pendulum is dependent on two variables: the length of the pendulum, ℓ, and the strength of the gravitational field, g, in which it is swinging. To a good approximation, for small displacements. $T = 2\pi\sqrt{\frac{\ell}{g}}$

Aim

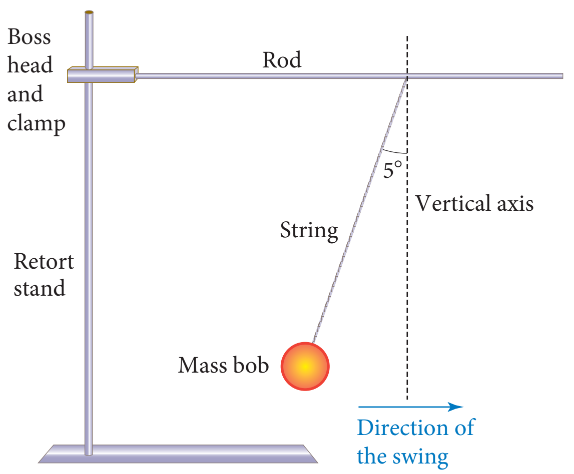
To measure the strength of Earth's gravitational field, g, near the surface using a pendulum

Materials

- retort stand
- boss head and clamp
- string, mass bob
- stopwatch or timing device
- ruler
- data logging apparatus optional

Procedure

1 Set up the apparatus as shown in Figure 1.3.



◀ **Figure 1.3**
Experimental set-up

2 As accurately as possible, measure the effective length of the pendulum from the top of the string to the centre of mass of the bob. This length should be about 1 m.

3 Pull the mass bob back until it makes an angle of about 5° with the vertical.

4 If it is available, use data logging apparatus to record the period of oscillation of the pendulum. Otherwise, record the time taken for 10 complete oscillations of the pendulum.

5 Change the length of the string to about 0.8 m and repeat steps 3 and 4.

6 Repeat again with string lengths of about 0.6 m and 0.4 m.

7 Calculate the period of the pendulum and the value of g and record it in the following data table.

▶

Results

Record the data in a table similar to the one shown below.

Length of pendulum (ℓ) (m)	Time for 10 oscillations (s)	Period of pendulum (T) (s)

Analysis of results

Two methods can be used to calculate the value of g.

Method 1: Using equations to model the pendulum

Use: $T = 2\pi\sqrt{\frac{\ell}{g}} \Rightarrow T^2 = (2\pi)^2 \frac{\ell}{g}$

$$g = \frac{4\pi^2 \ell}{T^2}$$

Substitute each period and length into the equation. Calculate a value for g each time and then take an average of the four values found.

Method 2: Using graphical representations to model the pendulum

Plot the relationship between ℓ and T^2. Plot ℓ on the x axis, as it is the independent variable. T^2, the dependent variable, is plotted on the y axis. The data points should be plotted and a line of best fit drawn.

Using: $T = 2\pi\sqrt{\frac{\ell}{g}}$

$$T^2 = (2\pi)^2 \frac{\ell}{g}$$

$$T^2 = \left(\frac{4\pi^2}{g}\right)\ell$$

This indicates that the value of the gradient of the graph of T^2 vs ℓ is $\frac{4\pi^2}{g}$.

Using any two points on the line of best fit, find the gradient of the graph and hence calculate the value of g using this method.

Discussion

1 What are the sources of uncertainty in this experiment?
2 Suggest ways in which these uncertainties could be minimised.
3 Explain why the time for 10 oscillations was measured and then divided by 10 to find the period, T.
4 If the length of the pendulum was always being over-estimated, how would this effect the value of g obtained for each analysis method?
5 Give your best estimate of g, including the uncertainty.

Taking it further

Investigate how accelerometers are used in cars and how they work.

Vertical motion

An object that is free to fall near Earth's surface will experience a gravitational force in a vertically downwards direction, regardless of its direction of motion. The gravitational force results in the object accelerating vertically downwards with a value of acceleration equal to g.

As the value of g is 9.8 m s^{-2} downwards, we get:

$$a = \frac{v_y - u_y}{t} \text{ (the definition of acceleration)} \quad (1)$$

$$v_y = u_y + at$$

$$\text{so: } v_y = u_y + gt$$

Notice that written in this way, if upwards is taken as being the positive direction, then change in speed is negative, and the value of g must be written as −9.8 m s^{-2} (downwards direction).

The velocity versus time graph for an object with constant acceleration is shown in Figure 1.4. The area under the v–t graph is a triangle plus rectangle in area, and is equal to the value of the distance interval, s.

Use the rectangle plus the triangle to find the area:

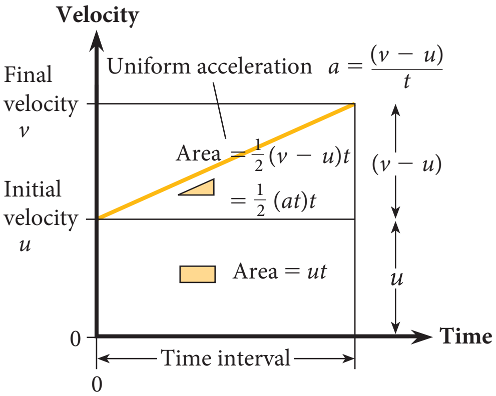


Figure 1.4 ▲ Motion of a point mass along a straight line

$$s = ut + \frac{1}{2}(v - u)t$$

$$s = ut + \frac{1}{2}(at)t$$

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

Considering the vertical motion of a free-falling object for which $a = g$, we get:

$$s = u_y t + \frac{1}{2}gt^2 \quad (2)$$

Combining equations (1) and (2) we can derive a third useful equation.

From $s = ut + \frac{1}{2}at^2$ and $v = u + at$ (hence, $t = \frac{v - u}{a}$), we can show that:

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

Again, if we consider the vertical acceleration only, with $a = g$, we have:

$$v_y^2 = u_y^2 + 2gy \quad (3)$$

The symbol y replaces s, the displacement interval, as only the vertical **component** of the object's motion is being considered here.

Equations (2) and (3) are sometimes written:

$$s = \frac{1}{2}gt^2 + u_y t$$

$$v_y^2 = 2gy + u_y^2$$

WORKED EXAMPLE 1.1

A rock is dropped out of a hot air balloon that is hovering stationary 200 m above the ground.

a With what speed does the rock hit the ground?

b How long does the rock take to fall?

Answers

a Identify known and required variables:

u	a	y	v
0	−9.8 m s^{-2}	−200 m	?

Select the equation and solve:

$v_y^2 = u_y^2 + 2gy$

$v_y^2 = 0^2 + 2 \times -9.8\,\text{m}\,\text{s}^{-2} \times -200\,\text{m}$

$v_y^2 = 3920$

$v_y = 63$ m s^{-1} vertically down

Logic

Select the appropriate equation. 1 mark

Substitute known values and solve the equation. 1 mark

Use correct units and number of significant figures. 1 mark

b Identify known and required variables:

u	a	s	t
0	-9.8 m s^{-2}	-200 m	?

Select the equation and solve: — Select the appropriate equation. 1 mark

$$s = u_y t + \frac{1}{2}gt^2$$

$$s = \frac{1}{2}gt^2$$

Rearrange for t and substitute known values. 1 mark

$$t = \sqrt{\frac{2s}{g}}$$

$$t = \sqrt{\frac{2 \times -200\text{m}}{-9.8\text{ms}^{-2}}}$$

$t = 6.4$ s

Use correct units and number of significant figures. 1 mark

Try these yourself

A ball is thrown vertically upwards at 20 m s^{-1}.

a What is the maximum height reached by the ball? (Hint: $v = 0$ at the maximum height.) (3 marks)

b How long will the ball take to fall back to its original position? (Hint: The motion of the ball is symmetrical about the maximum height reached.) (3 marks)

In the example above, the upwards direction was chosen as being positive. Quantities with a downwards direction, such as acceleration and displacement, are therefore assigned negative values. It would be possible to assign the reverse – assigning negative values to quantities with an upwards direction. If this was done the answer would have the same physical meaning and value.

QUESTION SET 1.1

Remembering

1 What are the units of g?

2 Can you be weightless when you are in a spacecraft orbiting Earth? Use a sketch to aid your answer.

Understanding

3 Identify the purpose of using the notation v_y rather than simply v for vertical motion.

4 Explain why $a_y = g$ for free-falling objects near Earth's surface.

5 When finding the maximum height reached by a tennis ball hit vertically upwards, the value of v is assigned to be zero. Why is this done?

Applying

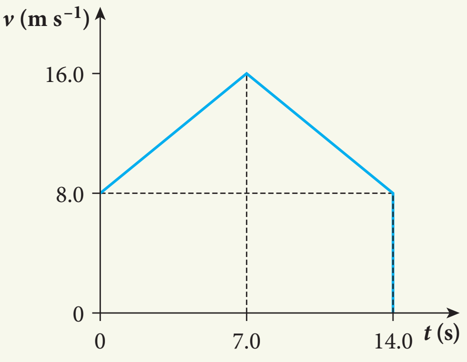
6 A rock is dropped from a very high cliff. How long does it take until the rock is moving at 60 m s^{-1}?

7 A ball is thrown from a window with an initial downwards speed of 4.4 m s^{-1}. It hits the ground after 1.4 s. How high is the window?

8 With what minimum speed must a student throw a pencil case vertically upwards so that it reaches a window 5.5 m high?

Analysing

9 The shape of the v versus t graph for an object is shown in Figure 1.5. Find the displacement interval, s, for the object over the time shown.



▲Figure 1.5

Reflecting

10 Given that the acceleration of objects near Earth's surface is directed vertically downwards, explain why a bullet shot horizontally from a height *h* hits the ground at the same time as one dropped from the same height.

The vector nature of gravitational force

The force exerted on an object within a gravitational field is applied in a direction towards the centre of mass of the object. Gravitational fields exist even for the smallest of masses. Therefore, we can say that a gravitational force is acting between you and the person nearest you. However, because this force is so small, it goes unnoticed. The very large mass of Earth results in a gravitational force on any object near it that cannot be ignored. The position of the centre of mass of Earth means that the gravitational force is exerted vertically downwards. Only tiny variations in the direction of this force are caused by local influences such as mountains or dense bodies of rock beneath the surface.

Vector addition of forces

Newton's second law, $\vec{a} = \frac{\vec{F}_{net}}{m}$, allows for the fact that any number of forces may be acting on a mass m at any time. The symbol $\vec{F}_{net}$ signifies the resultant force (or **net force**), or the sum of all the forces acting.

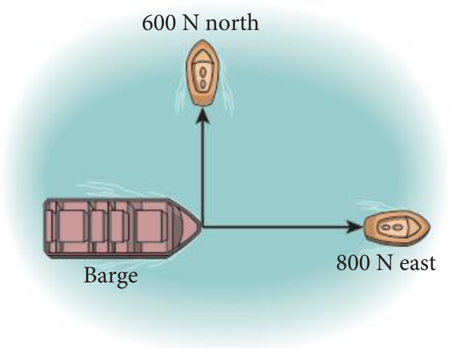
$$\sum \vec{F}(\text{by other objects on A}) = \vec{F}_{\text{net on A}}$$

To find the sum of the forces acting on an object, the magnitudes, or sizes, of the forces cannot simply be added. The vector nature of force means that the directions of the individual forces must be taken into account.

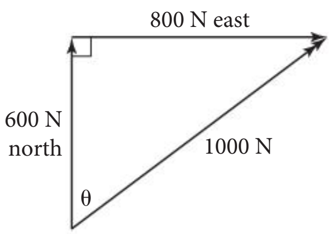
When adding two force vectors acting on an object at the same time, the resultant force may be found either geometrically or by using a scale diagram.

WORKED EXAMPLE 1.2

A barge has two tug boats pulling it with ropes attached in the directions shown in Figure 1.6. What is the vector sum of these two applied forces? (3 marks)



◀ **Figure 1.6** Using the tip-to-tail method, the two force vectors form two sides of a right triangle.

Answer	Logic	
 ◀ Figure 1.7 A right triangle is formed by the two forces and the resultant of these two forces.	Construct a correct vector diagram.	1 mark
The resultant, or net force, on the barge can be found by using Pythagoras' theorem: $F_{net} = \sqrt{600^2\,N + 800^2\,N} = 1000\,N$	Find the magnitude of the force.	1 mark
Force is a vector, so its direction is required: $\theta = \tan^{-1}\frac{800\,N}{600\,N} = 53°$ $F_{net} = 1000$ N, N53°E An alternative method can be used using a scale diagram. The scale chosen may, in this case, be 1 cm = 100 N.	Find the resultant angle.	1 mark

Try this yourself

Two ropes attached to a tree stump apply forces of 900 N to the north and 1600 N to the east. Find the resultant force on the stump. Use geometric methods first by constructing a scale diagram and then check using trigonometry. (5 marks)

Components of forces

When a resultant force is applied in one direction but motion can only occur in a different direction, only a part of the applied force will be effective. The object will only accelerate in the direction in which the motion is allowed.

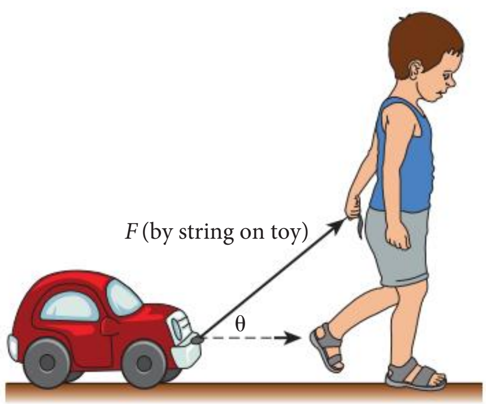
Figure 1.8 shows a child pulling a toy car on a string. The direction of the force applied to the toy by the string is at an angle θ to the direction of the motion of the toy. We shall assume that the wheels are fixed and the toy is sliding across the floor. **Friction** is negligible.

Figure 1.9 shows how a right triangle can be used to find the **component** of the applied force vector that is in the direction of motion of the toy car.

The component of this force in the direction of motion, which in this case is horizontal, is found by:

$$\cos\theta = \frac{F_H}{F(\text{string on toy})}$$

$$F_H = F(\text{string on toy})\cos\theta$$



▲ Figure 1.8 θ is the angle between the direction of the force applied to the toy car and the car's direction of motion

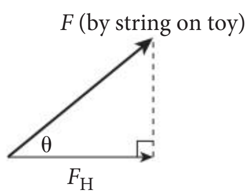


Figure 1.9 ▶ The component of the applied force F(string on toy) by the string is F_H

WORKED EXAMPLE 1.3

In the example shown in Figure 1.8, the force applied by the string on the toy car is 15.0 N. The angle made by the string with the horizontal direction of motion is 25°. What is the component of the force by the string in the direction of motion? (1 mark)

Answer	Logic	
$F_H = F(\text{string on toy})\cos\theta$ $= 15\cos 25°$ $= 13.6$ N	Select the correct equation and substitute.	1 mark

Try this yourself

A toy truck is being pulled by its string at 15° above the horizontal. The force applied by the string on the truck is 17.5 N. What force is acting in the direction of the motion of the toy truck? (1 mark)

Motion down an inclined plane

The normal force and friction were described in Nelson Physics Units 1 & 2 for the Australian Curriculum. Remember that the normal force and the friction force are the perpendicular and parallel components of the force applied by one surface on another surface.

Any object free to move down an inclined plane does so because of the component of the gravitational force acting in the direction of motion. For our purposes, we will model the motion of such objects as if they are free to slide down the inclined surface with no **friction** forces acting.

Of course, in reality, balls actually roll and boxes slide. The rolling action of a ball means that the ball possesses rotational kinetic energy as well as translational kinetic energy. Hence the gravitational potential energy is transformed into two different types of kinetic energy. This means that, if you fail to account for the rotational kinetic energy, a ball rolling down an inclined plane will have a speed at the bottom that is less than that predicted by energy conservation. The effect of friction on the rotation of wheels is beyond the scope of this book, but you should be aware of this effect in practical situations. For our purposes, objects will usually be analysed as sliding without rolling.

A sliding box will have a friction force applied to it in a direction opposite to its direction of motion. Again, this will mean that the speed of the box down the inclined place will be less than if the friction force did not act.

Figure 1.10 shows the two forces acting on a car on a frictionless inclined plane. The force arrows on the car are shown starting at the centre of mass of the car. We are modelling the car's mass as being a point mass for our purposes.

The weight of the car, which is the gravitational force by the mass of the Earth on the car, acts directly downwards. There is also a **contact force** acting on the car due to the road surface. To analyse the forces, we decompose the forces into components parallel and perpendicular to the surface. We choose parallel and perpendicular to the surface rather than vertical and horizontal because we know from experience that if there is any acceleration it will be along the direction of the slope.

The weight can be broken into two components, one parallel to the surface, w(parallel), and one perpendicular to the surface, w(perpendicular):

$$w(\text{parallel}) = w\sin\theta = mg\sin\theta$$
$$w(\text{perpendicular}) = w\cos\theta = mg\cos\theta$$

The contact force has two components: a component perpendicular to the surface and a component parallel to the surface.

The component of the force perpendicular to the surface, which is applied by the road surface on the car, is called the **normal force**, N. The friction force, F(friction), is the component of the force parallel to the surface, which is applied by the road surface on the car (Figure 1.10).

Perpendicular to the surface

No acceleration occurs perpendicular to the surface:

$$\Rightarrow F(\text{net})_{\perp} = 0$$
$$\Rightarrow N - mg\cos\theta = 0$$
$$\Rightarrow N = mg\cos\theta$$

The normal force, N, balances the perpendicular component of the weight.

Parallel to the slope

For a sloping surface, we expect the acceleration of the car to be affected by the friction on the object by the road and the angle of the slope:

$$\Rightarrow F(\text{net})_{\parallel} = ma$$
$$\Rightarrow mg\sin\theta - F(\text{friction}) = ma$$

If the hand brake is on so that the wheels cannot rotate, the car will exhibit one of two possible fates. If it is stationary, it will remain at rest, or if it is travelling at constant velocity, it will continue at constant velocity (Newton's first law). In these cases, $F(\text{net}) = 0$ (Newton's second law).

$$\Rightarrow F(\text{net})_{\perp} = 0$$
$$\Rightarrow mg\sin\theta - F(\text{friction}) = 0$$
$$\Rightarrow F(\text{friction}) = ma - mg\sin\theta$$

If a car rolls down a hill, it is not because of an absence of friction between the tyres and the road. Friction is required to make the wheels turn. On a completely frictionless slope a car will slide without the wheels turning.

Alternatively, if the wheels are prevented from rotating, the car may accelerate down the slope. But it will slide – the wheels will not rotate.

In general, when the wheels rotate and the car slides, motion is a bit more complicated to analyse. This is because the friction force on the wheels acts at the surface but the axis of rotation, the axle, is at another place; the friction force causes rotational motion.

If we make the approximation that the surface is frictionless, the net force acting is directed down the slope and is equal to the component of the gravitational force parallel to the slope:

$$F(\text{net}) = mg\sin\theta \qquad \textit{for a frictionless slope}$$

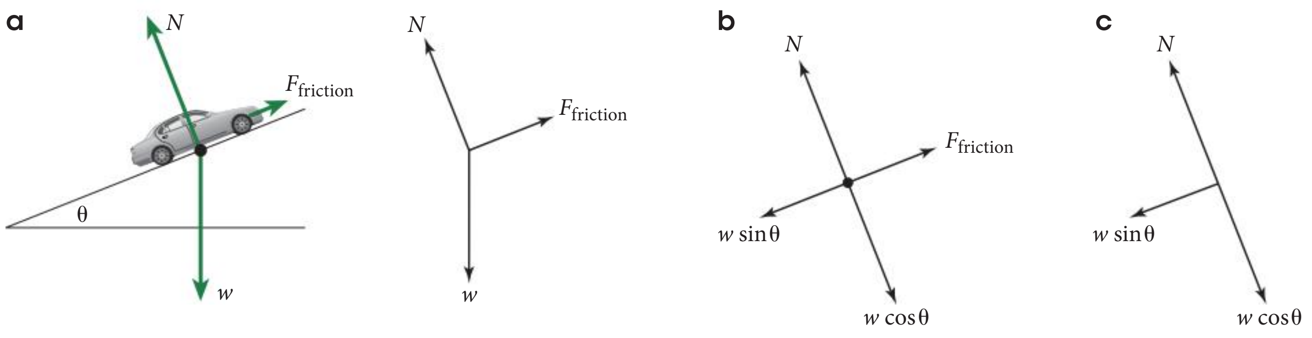


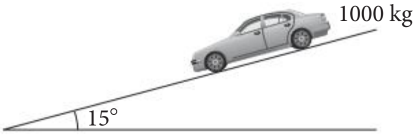
Figure 1.10
a) Forces acting on a car on a slope. b) The forces acting shown as components parallel and perpendicular to the slope. Note that $N = w\cos\theta$. c) If the slope is frictionless, then the net force $F(\text{net}) = w\sin\theta$ down the slope.

We often make the approximation that there is no friction force acting on one surface by another surface with which it is in contact. We know from experience that this is usually not realistic. However, it does give us a simple model that we can use to give us an indication of how an object will behave, even if it doesn't exactly match the real situation.

Every model that we use is an approximation. Sometimes it is reasonable to ignore friction, but sometimes it is not. You need to think about the situation you are modelling, and decide which approximations are reasonable to make and which ones are not.

WORKED EXAMPLE 1.4

A 1000 kg car is on a frictionless inclined plane that is at an angle of 15° to the horizontal, as shown in Figure 1.11.



▲ Figure 1.11

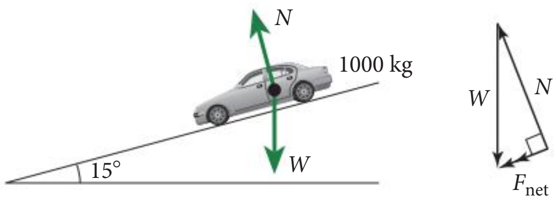
a Copy the diagram and draw the force vectors acting on the car and the force vector diagram to show F_{net}. (2 marks)

b Find F_{net}, the force acting on the car in the direction of motion. (2 marks)

c Hence find the car's acceleration. (2 marks)

Answers

a



▲ Figure 1.12

Logic

Correct force vectors shown on car. 1 mark

Correct vector diagram drawn. 1 mark

b $F_{net} = W \sin\theta$

$= mg \sin\theta$

$= 1000\,\text{kg} \times 9.8\,\text{N}\,\text{kg}^{-1} \times \sin 15°$

$= 2.5 \times 10^3$ N down the incline

Correct calculation. 1 mark

Direction is given. 1 mark

c $\vec{a} = \frac{\vec{F}_{net}}{m}$

$= \frac{2.5 \times 10^3\,\text{N}}{1000\,\text{kg}}$

$= 2.5\ \text{m s}^{-2}$ down the incline

Correct calculation. 1 mark

Direction is given. 1 mark

Try these yourself

An 800 kg car rests on a frictionless inclined roadway at 10° to the horizontal. (7 marks)

a Draw a diagram to show this and then draw the force vectors acting on the car and the force vector diagram to show F_{net}.

b Find F_{net}, the force acting on the car in the direction of motion.

c Hence find the car's acceleration, assuming it slides down the slope.

d What is the minimum force by the brakes on the car so that the car does not begin to roll down the hill?

EXPERIMENT 1.2

ACCELERATION DOWN AN INCLINED PLANE

Aim

To measure the acceleration of an object moving down an inclined plane

Materials

- steel ball
- ramp or desk placed on an incline
- ruler
- stopwatch or data logger and electronic timing gate

Procedure

1. Set up the ramp or desk at an incline and measure the length of the ramp, ℓ, and the height of the raised end, *h*, above the lower end.
2. Measure the length of a course for the ball to roll down the incline.
3. Use a stopwatch or electronic timing apparatus to record the time taken for the ball to roll down the course when it is released from rest.
4. Repeat step 3 at least twice more.

Results

Record your results in a table similar to the one below.

Length of course = ______m	Trial 1	Trial 2	Trial 3	Average
Time for ball to roll (s)				

Analysis of results

1. Find the average time interval taken for the ball to roll down the course. Include the uncertainty in your result.
2. Use the appropriate equation to find the speed of the ball at the bottom of the slope. Estimate the uncertainty in this value.
3. Use the appropriate equation to find the average acceleration of the ball down the slope. Estimate the uncertainty in this value.

Discussion

1. Calculate the expected value of the acceleration and the final speed of the ball, ignoring its rotational motion. To do this, take $a = g \sin\theta$ where θ is the angle between the horizontal and the slope.
2. Compare the expected speed and acceleration with the measured values. Do they agree? Don't forget that you need to consider the experimental uncertainties to say whether the values agree or not.
3. How can these differences (if any) be explained by rotational motion?
4. When we calculate a value of *g* from an experiment like this, we are modelling a rolling ball as an object sliding without friction. Comment on the appropriateness of using this model in this situation. Do you think there is a better model that could be used?

Taking it further

The surface must apply friction to the ball for it to roll, otherwise it would slide. However, we cannot measure the frictional force between the ball and the slope using this particular experiment. Design an experiment or investigation to measure the frictional forces between different surfaces using objects sliding, rather than rolling, down a slope.

QUESTION SET 1.2

Remembering

1 Which two methods can be used to find the resultant of two force vectors?

2 Why are the force vectors acting on a car shown originating from the centre of the car?

3 What is the direction of the net force by a roadway on a car, assuming no friction?

Understanding

4 If the component of the net force acting on an object down a frictionless incline is $F_{net} = W \sin\theta$, what would be the acceleration of the object when:

a $\theta = 0°$?

b $\theta = 90°$?

5 Would there be a limit to the number of force vectors acting on an object simultaneously for which a resultant F_{net} could be found?

Applying

6 a Using a scale drawing, find the resultant of forces A, B and C on an object.
A = 500 N north, B = 400 N west, C = 200 N south

b Verify your answer to part a using geometric methods.

7 Find the north and east components of a force vector of 650 N in a direction N38°E.

8 A 1200 kg car is parked on an incline of 8°. What friction force is required to hold the car in position when the handbrake is released?

Analysing

9 Explain the consequences of the total reaction force of a roadway on a car being at an angle to the surface on a horizontal road. When would this occur?

10 At the end of many steep declines there are emergency exit points for trucks that have sand as a base rather than solid roadway. Explain why these are necessary.

Projectile motion

A projectile is an object that moves relatively freely through space. Near Earth, projectiles are mainly affected by the gravitational field and friction effects. Projectiles can be launched vertically, horizontally or at any angle. Their initial speed, u, is determined by an initial force such as an explosion, a throw or a hit.

Real projectiles on Earth move quite differently from projectiles launched in a gravitational field where there is no atmosphere, such as on the Moon. Drag force is the frictional force applied by the air to a moving body. In general, drag acts in the opposite direction to that of the motion of the projectile relative to the air. The drag force depends on the air density, the speed of the projectile, its shape and surface material. In general, the smaller and more curved the surface, the less air drag there is. The faster its motion, the greater the drag. Details of the surface of an object, such as its texture, also affect the way air flows around it and, hence, the magnitude and direction of various forces applied to it. The dimples on a golf ball (Figure 1.13) actually cause the ball to fly further than one without dimples.

Shutterstock.com/cloki

Figure 1.13 ▲ The dimples on golf balls actually assist it to move through the atmosphere with less resistance.

In this chapter, we will ignore the effects of friction on projectiles and model their motion as if the only force applied to them after being launched is the gravitational force. This force is always directed vertically downwards.

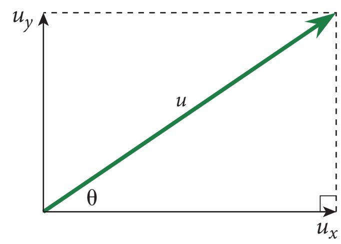
Horizontal component

The gravitational force acts vertically; it has no horizontal. Consequently, if we ignore friction, the projectile has zero acceleration in the horizontal direction, and we can say that its speed in the horizontal direction is constant. This is consistent with Newton's first law.

To find the magnitude of the initial horizontal component of velocity, u_x, a vector diagram is drawn as shown in Figure 1.14.

The horizontal component of the initial velocity is given by:

$$u_x = u \cos\theta$$



▲ **Figure 1.14** The horizontal and vertical components of the initial velocity, *u*

WORKED EXAMPLE 1.5

a A projectile is fired at an angle of 30° to the horizontal with a speed of 120 m s^{-1}. Find the horizontal component of the projectile's initial velocity. (1 mark)

b Some time later, the projectile hits the ground. With what horizontal speed does this occur? Explain your answer. (1 mark)

Answers	Logic	
a $u_x = u \cos\theta$ $= 120 \text{ m s}^{-1} \times \cos 30°$ $= 104 \text{ m s}^{-1}$	Select the correct equation and substitute.	1 mark
b $v_x = u_x = 104 \text{ m s}^{-1}$ There is no force applied to the projectile in the horizontal direction, so its horizontal speed remains constant throughout its flight.	Correct assumption made with reason.	1 mark

Try these yourself

Draw a sketch of the problem and then find the horizontal component of the initial velocity of projectiles fired: (4 marks)

a at an angle of 65° above the horizontal and a speed of 45 m s^{-1}.

b from a high cliff at an angle of 25° below the horizontal at 50 m s^{-1}.

Vertical component

When considering the vertical component of projectile motion, the action of the gravitational force must be taken into account. As has been seen previously, this force acts on any object near Earth's surface. The vertical force, F_v acting on a projectile is therefore its weight force, w, where $w = mg$. We take the value of g as 9.8 m s^{-2} downwards.

The vertical component of a projectile's motion can be treated independently of its horizontal motion. The two components can then be combined at a later stage to get a complete picture of the projectile's motion in two dimensions.

The initial vertical component of a projectile's speed, u_y can be found using Figure 1.14. It can be seen that:

$$u_y = u \sin\theta$$

The vertical component of a projectile's motion can be treated independently of its horizontal motion.

As the weight of the projectile, mg, acts all the time (assuming the projectile stays near to Earth's surface), the acceleration of the projectile is equal to $-g$, $-9.8\ \text{m s}^{-2}$, i.e. directed downwards.

The same equations used for vertical motion are used to analyse the projectile's vertical component of motion.

WORKED EXAMPLE 1.6

A projectile is fired at an angle of 45° above the horizontal with a speed of $86\ \text{m s}^{-1}$.

a Find the vertical component of the initial velocity of the projectile. (2 marks)

b Find the maximum height (vertical displacement) reached by the projectile. (4 marks)

c What is the time taken for the projectile to reach its maximum height? (2 marks)

Answers	Logic	
a $u_y = u \sin\theta$ $= 86\ \text{m s}^{-1} \sin 45°$ $= 61\ \text{m s}^{-1}$	Correct equation and substitution.	2 marks

b

u_y	a_y	v_y	y
$+61\ \text{m s}^{-1}$	$-9.8\ \text{m s}^{-2}$	0	?

Answers	Logic	
$v^2 = u_y^2 + 2gy$	Correct equation.	1 mark
$y = \frac{u_y^2}{2g}$	Rearranging for y, noting that $v = 0$.	1 mark
$= \frac{(61\ \text{m s}^{-1})^2}{2 \times 9.8\ \text{m s}^{-2}}$	Correct substitution.	1 mark
$= 190\ \text{m}$	Correct answer given.	1 mark

c

u_y	a_y	v_y	t
$+61\ \text{m s}^{-1}$	$-9.8\ \text{m s}^{-2}$	0	?

Answers	Logic	
Use: $v_y = u_y + gt$	Correct equation and substitution.	1 mark
$t = \frac{u_y}{g}$	Rearranging for t, noting that $v = 0$.	1 mark
$t = \frac{61\ \text{ms}^{-1}}{9.8\ \text{ms}^{-2}}$	Correct substitution.	1 mark
$= 6.2\ \text{s}$	Correct answer given	1 mark

Try these yourself

1 A projectile is fired with an initial speed of $210\ \text{m s}^{-1}$ at an angle of 70° above the horizontal. (8 marks)

a Find the vertical component of the initial velocity of the projectile.

b Find the maximum height (vertical displacement) reached by the projectile.

c What is time taken for the projectile to reach its maximum height?

2 A golf ball is hit and given an initial speed of $50\ \text{m s}^{-1}$ at an angle of 19° above the horizontal. Find its height 2.4 s after being hit. (3 marks)

Combining horizontal and vertical components

Once launched, a projectile will move with a constant horizontal velocity while simultaneously moving vertically with an acceleration of 9.8 m s^{-2} downwards.

The two components of motion can be calculated independently of each other and then combined. The time of flight for the projectile depends only on the vertical speed, u_y, of the projectile. If launched from ground level and moving over level ground, the projectile's motion will end when it hits the ground, when $y = 0$. Finding the time of this event will allow the range of the projectile to be found.

As the projectile moves with a constant horizontal speed u_x, we have:

$$x = u_x t$$

WORKED EXAMPLE 1.7

How far will a projectile fly over level ground if it is launched from ground level with a speed of 35 m s^{-1} at an angle 55° above the horizontal? (7 marks)

Answer	Logic	
First, the time of flight must be found for the projectile. This can be done by analysing the vertical component of motion.	Find the initial vertical component of velocity.	1 mark

$$u_y = u \sin\theta$$
$$= 35 \text{ m s}^{-1} \sin 55°$$
$$= 28.7 \text{ m s}^{-1}$$

The time of flight can be found from the vertical component of motion only. Using the symmetrical nature of the motion, where $v_y = -u_y$.

u_y	a_y	v_y	t
+28.7 m s^{-1}	−9.8 m s^{-2}	−28.7 m s^{-1}	?

Use: $v_y = u_y + gt$ — Correct equation and substitution. 1 mark

$$t = \frac{v_y - u_y}{g}$$ — Rearrange for t. 1 mark

$$= \frac{-28.7 \text{ m s}^{-1} - 28.7 \text{ m s}^{-1}}{-9.8 \text{ m s}^{-2}}$$

$$t = 5.86 \text{ s}$$ — Correct substitution and answer. 1 mark

Projectile moves with constant horizontal component of speed: — Correct horizontal component used. 1 mark

$$u_x = u \cos\theta$$
$$= 35 \text{ m s}^{-1} \cos 55°$$ — Correct equation used. 1 mark
$$= 20.1 \text{ m s}^{-1}$$

The range of the projectile is:

$$x = u_x t$$
$$= 20.1 \text{ m s}^{-1} \times 5.86 \text{ s}$$
$$= 120 \text{ m}$$ — Correct answer calculated with correct significant figures and units. 1 mark

Try these yourself

1 What is the maximum height reached by a netball that is thrown at 6.0 m s^{-1} at 45° above the horizontal? (5 marks)

2 What is the range of a cricket ball that is struck at 37 m s^{-1} and an angle of 50°? (7 marks)

EXPERIMENT 1.3

ANALYSING THE MOTION OF A PROJECTILE

Aim

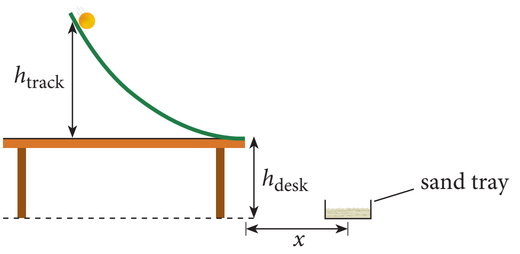
To find the launch velocity of a projectile by analysing its motion

Materials

- curved ramp such as a toy car track
- ball bearing or toy car
- ruler
- sand tray

Procedure

1 Set up the curved track on the edge of a desk as shown in Figure 1.15. Ensure the end of the track is horizontal.
2 Place the sand tray on the floor a small distance from the desk.
3 Measure the height h_{desk} of the end of the track above the sand tray.
4 Release the ball bearing or toy car from a height h_{track} above the end of the track.
5 Position the sand tray so that the car or ball lands in the sand. Measure the horizontal distance x from the impact point in the sand to the end of the track.
6 Repeat for several attempts and find the average distance x.



▲ **Figure 1.15** Experimental set-up

Results

1 h_{desk}, the height of the end of the track above the sand tray: __________ m
2 h_{track}, the height the ball or toy rolls through down the track: __________ m
3 Horizontal distance flown by ball or toy, x:

Attempt 1: __________ m

Attempt 2: __________ m

Attempt 3: __________ m

Average: __________ m

Analysis of results

1 Choose the appropriate equation to relate time of flight to the height of the desk, h_{desk}.
2 Use your average value of x to calculate the launch speed u_x.

Discussion

1 Using conservation of energy, where gained kinetic energy = lost gravitational potential energy, find the theoretical launch speed of the projectile. Note that we are ignoring the rotational motion of the ball.
2 Compare this value to the value of u_x found in the experiment. Do you think the model you have used, which ignores the rotational kinetic energy of the ball, is an appropriate one for this situation? Explain your answer.

Taking it further

Suggest ways in which this experiment could be improved.

QUESTION SET 1.3

Remembering

1 List the kinematic equations associated with projectile motion. Define all variables.

2 On a vector diagram show the horizontal and vertical components of the initial velocity of a projectile. Why do we assume that there is no horizontal component of acceleration for a projectile?

Understanding

3 Why is the time of flight for a projectile independent of the projectile's horizontal speed?

4 Sketch separate speed versus time graphs for the horizontal and vertical motions of a projectile that is launched directly upwards and lands below the launch position.

Applying

5 How long will it take a ball thrown at 12 m s^{-1} at an angle of 70° above the horizontal to reach a height of 4.0 m above its launch position?

6 A rock is thrown from a 55 m high cliff at an angle of 20° below the horizontal at an initial speed of 35 m s^{-1}. How long will it take the rock to land?

7 Find the range of a cannon ball that is launched with a velocity of 300 m s^{-1} at an angle of 35° above the horizontal and lands at the same height as the launch site.

Analysing

8 A projectile is launched at an angle of 60° above the horizontal and reaches a maximum height of 25 m. With what speed was it launched?

Reflecting

9 How does the trajectory of a real projectile differ from that of the model of projectile motion used in this chapter?

Circular motion

Newton's first law states that an object travelling in a straight line at a constant speed has no net force acting on it. What happens when the object travels at constant speed going around a corner? As the direction of motion is changing, clearly a net force must be acting on the object. This net force can be applied by such forces as the friction by the surface on the object, or a **tension** force exerted by an attached rope or string. It can also be exerted by the gravitational force by one mass on another mass.

If you are in a vehicle that is turning a sharp corner you will feel as if you are being thrown to the side. If you carefully examine your situation, you will realise that as the vehicle changes direction, you are trying to maintain your velocity in a straight line. The road surface pushes on the car through the tyres. In turn, the side of the vehicle and the seat push you towards the inside of the curve (see Figure 1.16). Your body is still trying to obey Newton's first law.

Shutterstock.com/supergenijalac

▲ **Figure 1.16**
A car on a horizontal road turns as a result of the inwards directed net force applied to the tyres by the road.

Uniform circular motion

An object going around a circle of radius r at a constant speed v takes a period of time T to complete one revolution around the circumference.

The period is the time taken to go around the circle. It has the unit of time, second (s). A related quantity is frequency f. This has the unit per second, s^{-1}, or hertz (Hz):

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

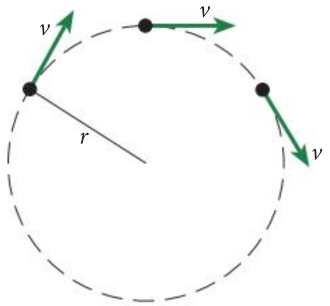


Figure 1.17 ▲ The velocity of the object at any point on the circle is the tangent to the circle at that point.

Newton is often credited with the invention of calculus, although in fact the mathematician Leibniz published the central ideas of calculus first. Newton claimed to have developed the same ideas much earlier, but without publishing them. This led to a long and bitter argument between the two.

The speed of the object is the distance it covers per unit time interval. In circular motion the object travels one circumference in one time period, so the speed is:

$$v = \frac{\text{circumference}}{\text{time}} = \frac{2\pi r}{T} = 2\pi r f$$

The object's direction is changing, so its velocity is changing. The velocity at a point on a circle is the speed in the direction along the tangent to the circle at that point (Figure 1.17).

The velocity is constantly changing, so the object is accelerating. Acceleration is the rate of change of velocity. In Figure 1.18(a), $\vec{v}_1$ and $\vec{v}_2$ are velocity vectors separated by a time interval Δt. The average acceleration, a_{ave}, over the time interval is:

$$\vec{a}_{ave} = \frac{\Delta \vec{v}}{\Delta t}$$

The instantaneous velocity can be found by making the time intervals smaller and smaller until you reach the limit:

$$\vec{a}_{inst} = \lim_{\Delta t \to 0} \frac{\Delta \vec{v}}{\Delta t}$$

In this case, the speed is constant, so the average acceleration is the same magnitude as the instantaneous acceleration, $\vec{a}_{inst}$, halfway through the time interval. Also $\vec{a}_{inst} \propto \Delta \vec{v}$, so $\vec{a}_{inst}$ points in the direction of $\Delta \vec{v}$, that is, towards the centre of the circle.

The concept of instantaneous acceleration uses a mathematical trick by taking the limit of Δt close to zero without actually being zero. If Δt was equal to zero, a division by Δt would not be possible. Here we are modelling by saying that the instant in time is so short it doesn't really matter, while at the same time saying it cannot be zero. This forms the basis for calculus and all other mathematical concepts based on taking the limit of a value.

In the same time interval Δt, the distance covered by the object is $v\Delta t$ and the change in velocity is Δv. The angle between the initial and final vectors is the same in each case.

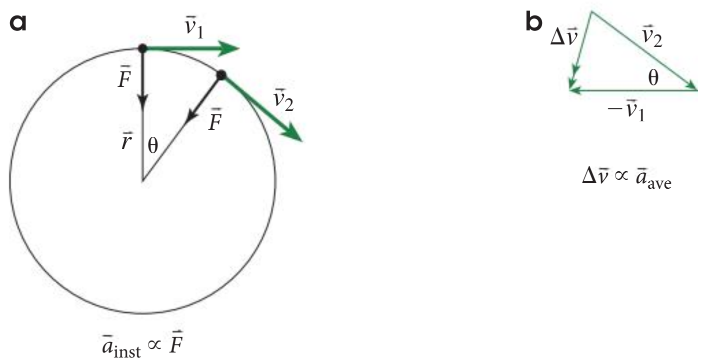


Figure 1.18 ▶ a) The net force, $\vec{F}$, acting on a particle in uniform circular motion points radially inwards towards the centre of the circle, for any position. The instantaneous acceleration, $\vec{a}_{inst}$, is in the same direction as the force. The particle's velocity, which is radial, is perpendicular to the force and acceleration. b) The average acceleration is the change in velocity divided by the time taken for that change to occur:

$$\vec{a}_{ave} = \frac{\vec{v}_2 - \vec{v}_1}{\Delta t} = \frac{\Delta \vec{v}}{\Delta t}.$$

In the limit that the time interval Δt is very small, $\vec{a}_{ave}$ approaches $\vec{a}_{inst}$ and $\Delta \vec{v}$ points towards the centre of the circle.

The centre-seeking, or **centripetal** acceleration, a_c, is found using similar triangles (Figure 1.19).

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

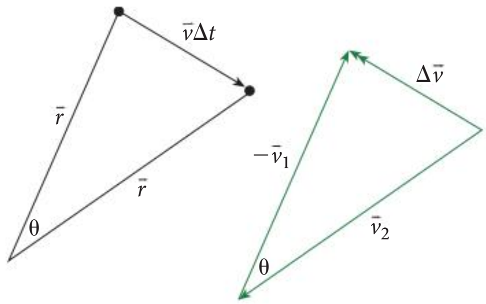
$$\Rightarrow \frac{\Delta v}{\Delta t} = \frac{v^2}{r}$$

$$\Rightarrow a_c = \frac{v^2}{r}$$

Substituting for v, we get $a_c = \dfrac{4\pi^2 r}{T^2}$

Centripetal acceleration is always directed radially in towards the centre of the circle: $a_c = \frac{v^2}{r} = \frac{4\pi^2 r}{T^2}$

This allows us to calculate the magnitude of the acceleration when we know the speed and radius data.



◀ Figure 1.19
The vector diagrams for distance covered in a time interval and change in velocity are similar. In similar triangles the ratio of corresponding sides is the same. Note that $\bar{v}_1$ and $\bar{v}_2$ are perpendicular to r at the relevant points on the circle. When $\bar{v}$ is translated to the middle of the time interval, it can be seen to point radially in towards the centre.

Force and uniform circular motion

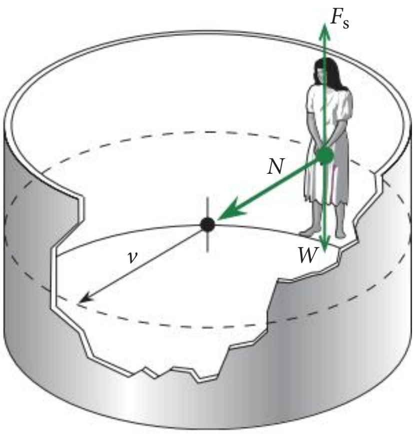
As the acceleration of an object undergoing **uniform circular motion** is directed towards the centre of the circle, the net force F_{net} is also applied towards the centre of the circle.

The magnitude of the net force can be found using Newton's second law:

$$\Sigma F(\text{on object}) = ma$$
$$\Rightarrow \Sigma F(\text{on object}) = m\frac{v^2}{r} = m\frac{4\pi^2 r}{T^2}$$

The net force applied to an object to keep it moving around a circle is called the centripetal force. There is no real force called centripetal force. It is a net force, the sum of real forces. A person in a rotor (Figure 1.20) is forced around the circle by the normal force exerted by the wall of the rotor on the person and towards the centre . This is written as

$$F_c = \frac{mv^2}{r} = m\frac{4\pi^2 r}{T^2}$$



▲ Figure 1.20
A person in a 'rotor' maintains their circular motion by the normal force applied by the wall to the person. This is the centripetal force.

'Centrifugal' force – a fictitious force

Being in an accelerating frame of reference can lead to incorrect conclusions. A person in a 'rotor' may perceive themselves as being thrown outwards by a 'centrifugal' force. In reality, the wall of the rotor is applying the centripetal force necessary to keep the rider moving in a circle, as viewed by a spectator.

Cases of apparent 'centrifugal force' can always be explained by using a valid, non-accelerating frame of reference to analyse the motion in question.

In the rotor ride shown in Figure 1.20, it is the normal force that provides the net, centre-directed force to keep the person in uniform circular motion. As we shall see in the next section, the other component of the contact force, the friction force, can also be part of the net force that causes circular motion.

A third force, the tension in a rope or cable, may also be used to supply the net force for circular motion. Remember that tension always acts in the direction of the cable. If the only forces acting on an object swung in a circular path are the tension force in the cable and the gravitational force, then the sum of the two gives the net force. This net force is what is termed 'centripetal force'. Notice that there is no such force, called centripetal force, in the sense that tension and gravitational force are forces. Centripetal force is a net force, the vector sum of real forces, in this case tension and gravitational forces.

Consider an object being whirled in a horizontal circle, as in Figure 1.21(a). You can see that the string makes an angle θ with the horizontal so that the tension has both a horizontal and a vertical component.

The vertical component of the tension is equal to the gravitational force acting on the object, as it has zero vertical acceleration.

Vertically:

$$F(\text{net}) = 0$$
$$\Rightarrow T(\text{vertical}) - mg = 0$$
$$\Rightarrow T\sin\theta = mg$$

T(horizontal) is the only component of the force that can act on the object to cause it to change direction, hence accelerate:

Horizontally:

$$F(\text{net}) = m\frac{v^2}{r}$$
$$\Rightarrow T(\text{horizontal}) = T\cos\theta$$
$$\Rightarrow T\cos\theta = m\frac{v^2}{r}$$

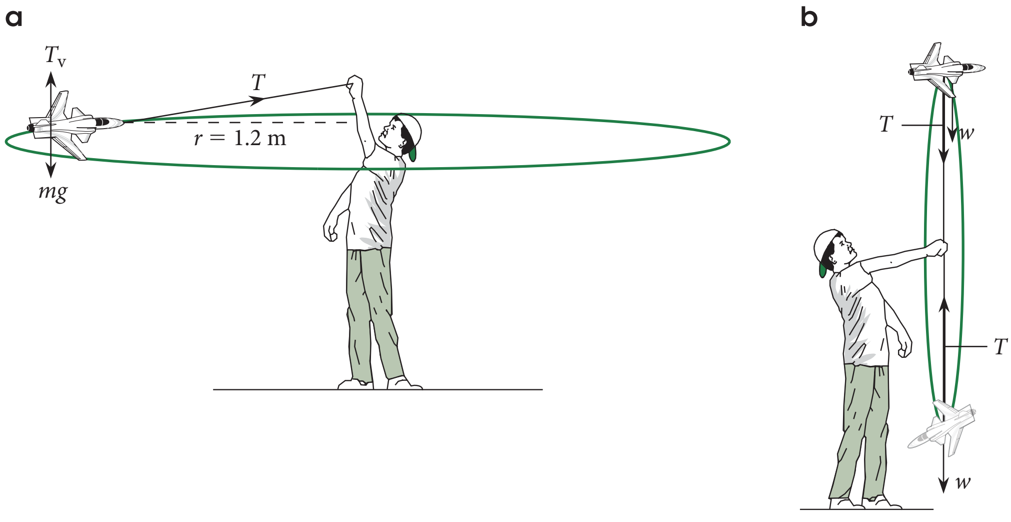


Figure 1.21 ▶ An object being whirled in a) a horizontal and b) a vertical circle

The tension in this case has a constant magnitude, but its direction varies.

When an object is whirled in a vertical circle, as in Figure 1.21(b), it is the sum of the gravitational force (w) and the vertical component of the tension that add to give the net force towards the centre. Take the simple case of an object being whirled in a vertical circle with the cable vertical also. At any point on the circle the net force is:

$$\vec{F}(\text{net}) = \vec{w} + \vec{T}$$

At the bottom of the circle, the net force is upwards. Hence, at this point the tension must be larger than the weight, as they act in opposite directions. At the top of the circle, where w is acting in the same direction as T, the total force must still be the same, so the tension must be less and may even drop to zero. In this case the tension not only varies in direction but in magnitude also.

WORKED EXAMPLE 1.8

1 A 50 kg person is in a rotor moving with a speed of 6.0 m s^{-1}. The radius of the rotor is 3.0 m. What is the centripetal force applied to the person? (2 marks)

2 A 250 g aeroglider on the end of a string is swung in a horizontal circle with a radius of 1.2 m, as shown in Figure 1.21. The aeroglider makes a revolution every 2.0 s.

- **a** What is the speed of the aeroglider? (2 marks)
- **b** What is the horizontal component of the tension (force) in the string? (2 marks)
- **c** What is the acceleration of the aeroglider? (2 marks)

Answers	Logic	
1 $F_c = \frac{mv^2}{r}$	Correct equation and substitution.	1 mark
$= \frac{50 \text{ kg} \times (6 \text{ m s}^{-1})^2}{3 \text{ m}}$		
$= 600$ N towards the centre	Correct answer given with direction.	1 mark
2 a $v = \frac{2\pi r}{T}$	Correct equation used.	1 mark
$= \frac{2\pi \times 1.2 \text{ m}}{2.0 \text{ s}}$		
$= 3.8$ m s^{-1}	Correct answer given.	1 mark
b $F_C = T_H = \frac{mv^2}{r}$	Correct equation used.	1 mark
$= \frac{0.250 \text{ kg} \times (3.8 \text{ m s}^{-1})^2}{1.2 \text{ m}}$		
$= 3.0$ N towards the centre	Correct answer with direction.	1 mark
c $a_c = \frac{v^2}{r}$	Correct equation used.	1 mark
$= \frac{(3.8 \text{ m s}^{-1})^2}{1.2 \text{ m}}$		
$= 12$ m s^{-2} towards the centre	Correct answer given with direction.	1 mark

Try these yourself

1 What is the maximum speed allowed for a rotor ride of radius 2.5 m if the maximum net (centripetal) force exerted on a 70 kg person is not allowed to exceed 1000 N? (3 marks)

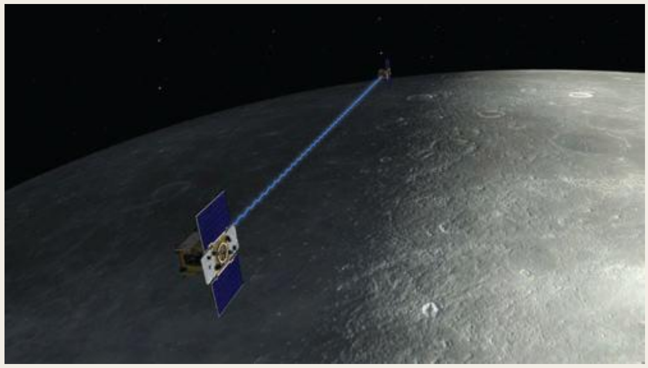
2 What is the radius of a rotor ride that is exerting a 400 N net force on a 30 kg child when the speed of the child is 4.5 m s^{-1}? (3 marks)

Scientific literacy: NASA's GRAIL mission to map the Moon's gravitational field

In 2011 NASA's Gravity Recovery And Interior Laboratory (GRAIL) mission placed two spacecraft into the same orbit around the Moon. As they flew over areas of greater and lesser gravity, caused both by visible features such as mountains and craters and by masses hidden beneath the lunar surface, they moved slightly towards and away from each other. An instrument aboard each spacecraft measured the changes in their relative velocity very precisely. Scientists translated this information into a high-resolution map of the Moon's gravitational field.

This gravity-measuring technique is essentially the same as that of the Gravity Recovery And Climate Experiment (GRACE), which has been mapping Earth's gravity since 2002.

GRAIL's engineering objectives were to map lunar gravity and to use that information to increase understanding of the Moon's interior and thermal history. Spacecraft have been observed to change orbit unexpectedly as a result of unobserved regions of mass concentrations or mascons. Accurate gravitational maps of the Moon will enable future Moon missions to be safer and more precise.


NASA/JPL-Caltech/MIT

▲ **Figure 1.22** An artist's impression of the twin spacecraft that comprise the GRAIL mission

Getting the two spacecraft where they needed to be, when they needed to be there, was extremely challenging. It required a set of manoeuvres never before carried out in solar system exploration missions.

The two GRAIL spacecraft were near-twins, each about the size of a washing machine, with minor differences resulting from the need for one specific spacecraft (GRAIL-A) to follow the other (GRAIL-B) as they circled the Moon at a height of 55 km. They each described the same polar orbit lasting 113 minutes.

The on-board Lunar Gravity Ranging Systems measured changes in the distance between the two spacecraft down to a few microns – about the diameter of a red blood cell. That's not bad, considering they were flying 150 km apart.

Each spacecraft carried a set of cameras for MoonKAM (Moon Knowledge Acquired by Middle school students). This was the first time a NASA planetary mission had carried instruments expressly for an education and public outreach project. Among other things, these cameras recorded the final crash site at the end of the mission.

The two GRAIL spacecraft communicated with each other and Earth through antennas. In this way they could record any changes in the distance between them. One of the antennas on each craft was mounted on the sunny side of the spacecraft and another on the dark side. The sunny-side antennas pointed to Earth during the full moon and the dark-side antennas pointed to Earth during new moons. This system avoided the need to rotate the antennas mechanically during the mission. Any changes would have altered the spacecraft's centre of mass and disturbed the measurements.

The mission lasted for about a year, terminating at the crash site, Sally Ride, in December 2012.

Source: NASA: GRAIL Mission overview, July 12, 2011 (http://www.nasa.gov/mission_pages/grail/overview/index.html#)

Questions

1 What do the acronyms, GRAIL and MoonKAM, stand for?

2 What was the purpose of the GRAIL mission?

3 Why were two spacecraft used for this mission?

4 Describe how measuring the distance between the two craft enables the local value of g to be calculated.

5 The moon has a mean radius of 1.74×10^3 km. Calculate:

 a the orbital speed of the GRAIL twins

 b the orbital acceleration of the GRAIL twins

6 Do some more research and find out why NASA launched this mission in the first place. Why would it be useful to have a map of the moon's gravitational field?

WORKED EXAMPLE 1.9

A 0.20 kg object is whirled in a vertical circle on the end of a string of length 0.60 m. The speed of the object is a constant 3.0 m s^{-1}.

a What is the tension in the string at the top of the circle? (3 marks)

b Calculate the tension in the string at the bottom of the circle. (3 marks)

Answers	Logic	
a At the top: $\vec{F}_{net} = T + mg$	Correct analysis of forces.	1 mark
$T + mg = m\frac{v^2}{r}$		
$T = m\frac{v^2}{r} + mg$	Rearrange for T.	1 mark
$\Rightarrow T = 0.20\text{kg} \times \frac{(3.0\text{ m s}^{-1})^2}{0.60\text{m}} - 0.20\text{kg} \times 9.8\text{m s}^{-2}$	Correct substitution and answer.	1 mark
$T = 1.0\text{N}$		
b $\vec{F}_{net} = T - mg$	Correct analysis of forces.	1 mark
$T - mg = m\frac{v^2}{r}$		
$T = m\frac{v^2}{r} + mg$	Rearrange for T.	1 mark
$\Rightarrow T = 0.2\text{ kg}\frac{(3.0\text{ m s}^{-1})^2}{0.60\text{ m}} + 0.20\text{ kg} \times 9.8\text{ m s}^{-2}$		
$T = 5.0\text{ N}$	Correct substitution and answer.	1 mark

Try these yourself

A rock with mass 1.5 kg is being whirled in a vertical circle on a string 0.80 m long. The speed of the rock is constant. At the top of the circle, the tension T in the string is zero.

a Find the speed of the rock. (3 marks)

b Find the tension in the string when the rock is at the bottom of the circle. (3 marks)

Going around corners

In order to travel around a circular path such as running around the bend in a 200 m race or driving a car around a corner, you must push outwards on the ground. The ground will push back on you (Newton's third law) towards the centre of the circular path you are taking. Remember that the contact force has two components: the parallel friction component and the perpendicular normal component. When you drive a car, it is the friction force that acts on the tyres in the direction of the car's motion that makes the car go forwards. When you apply the brakes, it is again the friction force, this time acting in the opposite direction, which slows the car down. Without friction, the tyres would slip against the road, the wheels would spin, and you wouldn't be able to start or stop.

When you want to drive around a corner on a flat road, the tyres must push down at an angle to the horizontal. The ground then pushes back with an equal and opposite force. It is friction, or the parallel component of this reaction force, that supplies the centripetal force needed to make the car go around the corner. Hence when a car corners on a flat road, for the part of the corner that is approximately circular we can say that:

$$F_c = F_{friction} = m\frac{v^2}{r}$$

Remember that friction and the normal force are two components of the same force – the contact force that one surface exerts on another surface.

Think carefully about the direction of friction forces. Friction acts to oppose the relative motion of surfaces in contact. This does not always mean that it 'acts to oppose motion' as you may have heard.

We can see from this that the ability of a car (and its driver) to successfully negotiate a corner depends on how sharp the corner is and how fast the car is going. The faster the car is going, the greater the frictional force required, and the force increases with the square of the velocity. There is a maximum friction force that the road can exert on the tyres. Therefore, slowing down for a corner can make a big different between staying on the road and having an accident. The maximum friction force is significantly reduced by water, mud or oil on the road, and by worn tyres.

Cornering on a banked road

If a vehicle travels around a corner on a banked road and it maintains the same **radius of curvature**, it travels horizontally. In this case, there must be an inwardly directed net force, $\vec{F}_{\text{net}}$ acting on the vehicle.

There are at least two forces acting on the vehicle. These are its weight, due to gravity, and the contact force due to the road acting on the tyres. Air resistance or drag may also be important if the car is going fast. Again, consider the contact force as two components; the normal force and the fiction force. For this analysis we shall make the simplifying assumption that only the weight and normal forces are significant.

The forces applied to the vehicle are therefore:

- normal force, $\vec{N}$
- weight force, $\vec{w}$

The vector sum is:

$$\vec{F}_{\text{net}} = \vec{N} + \vec{w}$$

From Figure 1.23:

$$\frac{F_{\text{net}}}{w} = \frac{F_{\text{net}}}{mg} = \tan\theta$$

$$\Rightarrow F_{\text{net}} = mg\tan\theta$$

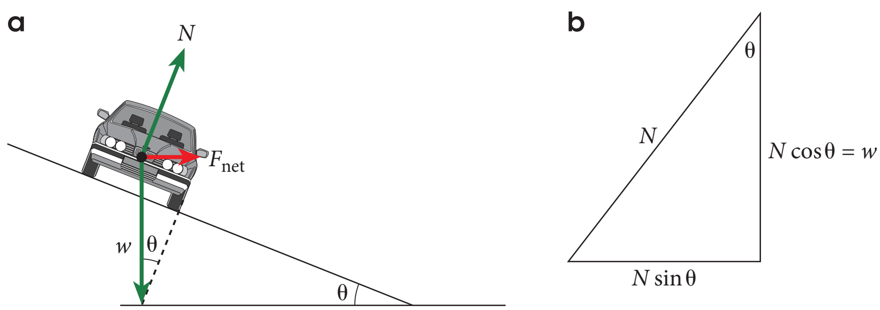


Figure 1.23 ▶ Forces on a car cornering on a frictionless banked road. The net force is the horizontal component of the normal force.

The weight force acts vertically downwards, and so *cannot* contribute to the centripetal force, which is horizontal. It is therefore the horizontal component of the normal force that provides the centripetal force.

We know that the vertical component of the normal force must be equal to the weight:

$$N_{\text{V}} = N\cos\theta = w$$

so:

$$N = \frac{w}{\cos\theta}$$

The centripetal force is then:

$$F_c = N_H = N\sin\theta = \frac{w\sin\theta}{\cos\theta} = w\tan\theta$$

Recalling that $w = mg$, we have:

$$F_c = mg\tan\theta = m\frac{v^2}{r}$$

This analysis assumes no friction force. The friction force is of course never really zero. In this example, the friction force may act parallel to and up the slope, or down the slope, depending on the speed of the car. If the car is going very slowly, the friction force acts up the slope, preventing (or slowing) the car from sliding down the slope. If the car is going very quickly, the friction force may act in the opposite direction, pushing the car up the slope.

WORKED EXAMPLE 1.10

The car in Figure 1.23 has a mass of 1500 kg. It is travelling horizontally at 20 m s^{-1} around a bend that is banked at 10° to the horizontal.

a What is the normal force acting on the car? (3 marks)

b What is the net force acting on the car? (2 marks)

c What radius of curvature must the road have so that the car can turn the corner with no friction force acting? (3 marks)

Answers	Logic	
a $N\cos\theta = mg$	Correct equation used.	1 mark
$N = \frac{mg}{\cos\theta}$	Rearrange for *N*.	1 mark
$N = \frac{1500\text{ kg} \times 9.8\text{ m s}^{-2}}{\cos 10°}$		
$N = 1.49 \times 10^4\text{ N}$	Correct substitution and answer.	1 mark
b $F_{net} = mg\tan\theta$	Correct equation used.	1 mark
$F_{net} = 1500\text{ kg} \times 9.8\text{ m s}^{-2} \times \tan 10°$		
$= 2.59 \times 10^3$ N towards centre of curve	Correct substitution and answer.	1 mark
c $F_{net} = m\frac{v^2}{r}$	Correct equation used.	1 mark
$r = \frac{mv^2}{F_{net}}$	Rearrange for *r*.	1 mark
$\Rightarrow r = \frac{1500\text{ kg} \times (20\text{ m s}^{-1})^2}{2.59 \times 10^3\text{ N}}$ $= 232$ m	Correct substitution and answer.	1 mark

Try this yourself

A car moving at 17 m s^{-1} enters a curve with a radius of 150 m. What is the ideal angle of banking for this road so that the horizontal component of the normal force on the car is equal to the centripetal force required to maintain the car's circular motion? Why is the mass of the car not needed for this calculation? (6 marks)

Vertical circular motion

A car of mass m, travelling at speed v, goes over the crest of a hill, as shown in Figure 1.24, which we shall approximate as a part of a circular surface with a radius r.

At the top, the net force is downwards to the centre of the circle:

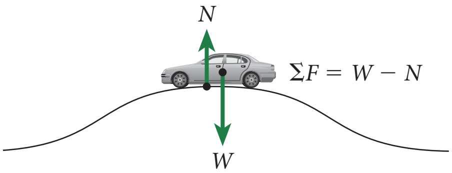


Figure 1.24 ▲ The net force on a car on the crest of a hill is the vector sum of the normal force and the weight force. The normal force can be zero if the speed is great enough.

$$\vec{F}_{\text{net}} = mg - N$$

$$mg - N = m\frac{v^2}{r}$$

$$\Rightarrow N = mg - \frac{mv^2}{r}$$

This means that there is a speed at which $N = 0$, above which the car will leave the ground. The speed at which this occurs depends only on the **radius of curvature** of the crest:

$$m\frac{v^2}{r} - mg = 0$$

$$\frac{v^2}{r} - g = 0$$

$$\Rightarrow v^2 = gr$$

$$v = \sqrt{gr}$$

At this speed the occupants would feel a moment of weightlessness (even though they are not) as the car is in free-fall.

If the car goes through a dip, which can be approximated as part of a circular surface, at the bottom, the net force is upwards towards the centre of the circle (Figure 1.25).

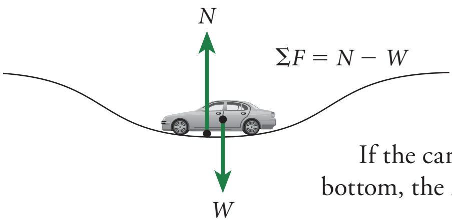


Figure 1.25 ▲ The net force on a car in a dip in the road is upwards. It is the vector sum of the normal force and the weight force.

$$\Sigma\vec{F}_{net} = N - mg$$

$$N - mg = m\frac{v^2}{r}$$

$$\Rightarrow N = mg + m\frac{v^2}{r}$$

Hence the normal force is much greater than mg and you feel 'heavier' because the car seat pushes up on you harder.

We have seen in this section that the centripetal force is the net force acting on an object in circular motion. This force may be due to a single force, such as the normal force on a rotor ride, or one component of a single force, such the horizontal component of the tension for an object whirled in a horizontal circle. It may also be due to the sum of two (or more) forces, such as the gravitational force and tension for an object whirled in a vertical circle on a cable. In each of these cases, and in all others, the centripetal force is the *net* force acting. When you analyse circular motion you should be able to identify which forces add to give this net force. The centripetal force is the *result* of forces acting, not a separate force, so it should never be added to other 'physical' forces.

Remember that the centripetal force is the net force acting on an object undergoing circular motion. You should never add the centripetal force to other forces to get a total force – the centripetal force *is* the total force.

The centripetal force may be due to a single force, or the sum of two or more forces.

QUESTION SET 1.4

Remembering

1 Describe uniform circular motion in your own words.

2 Why is an object that is undergoing uniform circular motion said to be accelerating?

Understanding

3 A person in a 'rotor' ride experiences a force directed towards the centre of the ride. Why does this occur?

4 Give an example of when a non-physics student may cite 'centrifugal force' to explain what happened to them.

Applying

5 A car undergoing uniform circular motion going around a corner has a centripetal force *F* acting on it.

- a In which direction is this force?
- b In terms of *F*, what is the magnitude of the centripetal force if the car enters the same curve moving at twice the speed? Show your working.

6 A track for a very fast train cannot have curves with a small radius. Use an appropriate equation to show why this is so.

7 What net force will need to be applied to a 70-tonne aeroplane that is turning with a radius of curvature of 5.0 km and flying at 100 m s^{-1}? What is the origin of this force?

Analysing

8 Show with a diagram why the tension in a string is greatest at the bottom of the circle when a rock is being whirled in a vertical circle.

Reflecting

9 List examples of motion other than those mentioned in this chapter in which uniform circular motion occurs or is occurring.

10 A stunt car goes through a loop, so that it follows a vertical circular path. Assume that the car has approximately constant speed.

- a Sketch a graph of:
  - i the weight force acting on the car as a function of time as it performs the loop.
  - ii the magnitude of the normal force acting on the car as a function of time as it performs the loop.
  - iii the magnitude of the total force acting on the car as a function of time as it performs the loop.
- b Explain how your graphs are possible, given that the weight and normal must add at all times to give the total (centripetal) force.

WOW

Simple machines

Renaissance builders identified six types of simple machine: lever, wheel and axle, inclined plane, wedge, pulley and screw. Simple machines change the direction of the effect of a force or change the effort needed to achieve a result. The wheel and axle enables a force at the perimeter of a circular object to rotate the object about its central axis. The brilliant Florentine master builder, Filippo Brunelleschi, used these to advantage in the building of the dome of the great church of Florence, Santa Maria del Fiore di Fiorenza, known as the Duomo.

Figure 1.26 ▶ The dome of the Duomo in Florence, Italy

DUOMO

Learn more about the Duomo and Brunelleschi, the man who designed it.

Torque and rotation

We have seen that when a force acts on an object the object accelerates in a straight line in the direction of the force. This is a referred to as translational motion. We have also looked at uniform circular motion, which is type of rotational motion.

Imagine using a spanner to undo a bolt. You want the bolt to turn, but you don't want to move it sideways or up or down. So you want it to twist, but not to translate. Similarly, when you open a door, you want to make the door rotate about its hinges. You don't want the door to come off its hinges and move up or down or go forwards or backwards. In these cases in which we want rotation but not translation, we need to apply a **torque** to the object.

Torque is the rotational equivalent of force. To get a stationary object to rotate, we must apply a torque. To get a rotating object to stop rotating, we must also apply a torque. The greater the torque that is acting on an object, the greater the change in the rotational motion of the object.

A torque causes a change in rotational motion, and is due to a force acting on an object at some distance from a pivot point or axis of rotation.

The torque, τ, acting on an object due to a force, F, is given by

$$\tau = rF\sin\theta$$

where r is the distance from the pivot point to the point at which the force acts. The angle, θ, is the angle between the line joining the pivot point to the point of application of the force, and the force itself.

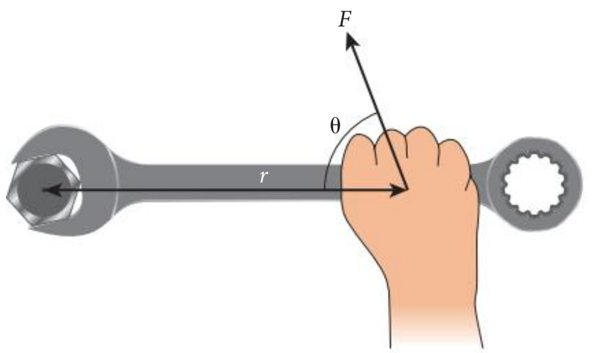


Figure 1.27 ▼ The torque applied to the spanner by the hand is $\tau = rF\sin\theta$.

This is shown in Figure 1.27. Torque is also called **moment**, particularly in engineering.

Torque has units of N m. These are the same units as work, which we often write in units of J because work is an energy. However torque *is not* a type energy, even though it has the same units, so we never write the units of torque as J. Energy is a scalar quantity, but torque is a vector quantity. The direction of the torque vector is important because it determines the direction of the resulting rotation.

To find the direction of torque, point the fingers of your right hand in the direction of the line joining the pivot to point of application of the force, as shown in Figure 1.28. Now curl your fingers in the direction of the force. Your thumb points in the direction of the torque vector, and your fingers show the direction of the resulting rotation. It may seem odd that the torque vector is perpendicular to the direction of rotation. However, every point on a rotating object is changing the direction of its velocity constantly, as we saw in circular motion. So it makes sense to define the torque vector as perpendicular to the plane in which rotation occurs. This is the only way that we can assign it a unique direction for the whole rotating object.

From the equation for torque we can see that the greater the force, the greater the torque. The greater the distance between the pivot and the point of application of the force, the greater the torque. When the force acts right at the pivot point the torque is zero.

The angle is also important. Think about using a spanner to undo a bolt again. To get the maximum torque, you apply the largest force you can, at the end of the handle furthest from the bolt. You apply the force perpendicular to the handle of the spanner, as shown in Figure 1.29. The angle in this case is $\theta = 90°$ and the torque has its maximum value, $\tau_{max} = rF$. If you pull or push on the handle of the spanner along the line of the handle, then the angle is $\theta = 0$, and the torque is also zero, and no rotation is achieved.

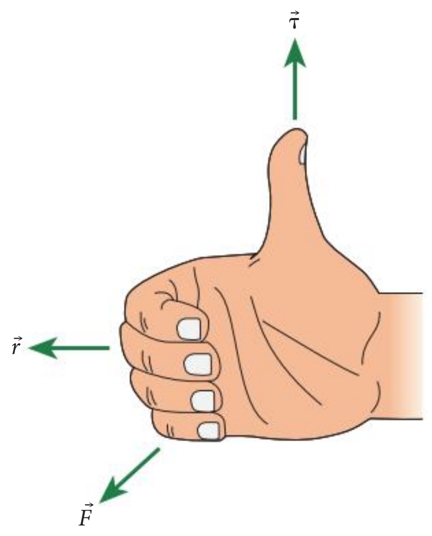
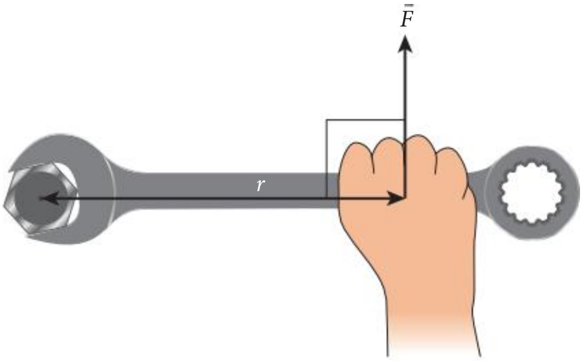
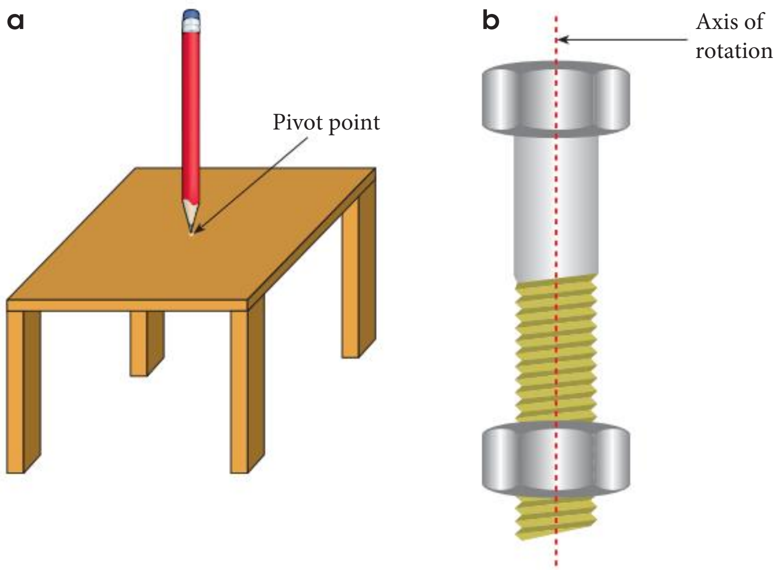


Figure 1.28 ▲ The right-hand rule for finding the direction of a torque

The pivot point is the point about which the object rotates. We always talk about the torque *about a point*, as the torque can be calculated about any arbitrary point on an object. The point we choose determines the distance *r*. Often there is an obvious choice, such as the hinges of a door, or some other point at which a restraining force acts. If you try to balance a pencil on its end on your desk (Figure 1.30a), it will rotate about its tip and fall over. The tip is the pivot point. Sometimes we talk about an axis of rotation rather than a pivot point. This just means we are thinking about a three-dimensional object rotating about a particular axis. For example, a bolt or screw (Figure 1.30b) rotates about a line along its own long axis when it is turned with a spanner.



▲ **Figure 1.29**
Torque is at a maximum when the lever arm (*r*) and the force (*F*) are perpendicular.



◀ **Figure 1.30**
a) The tip of the pencil is the pivot about which the pencil will rotate when it falls. b) The bolt rotates about a line along its own long axis.

Torque is more correctly given as the vector cross product of the vector r *and the force* F*:*

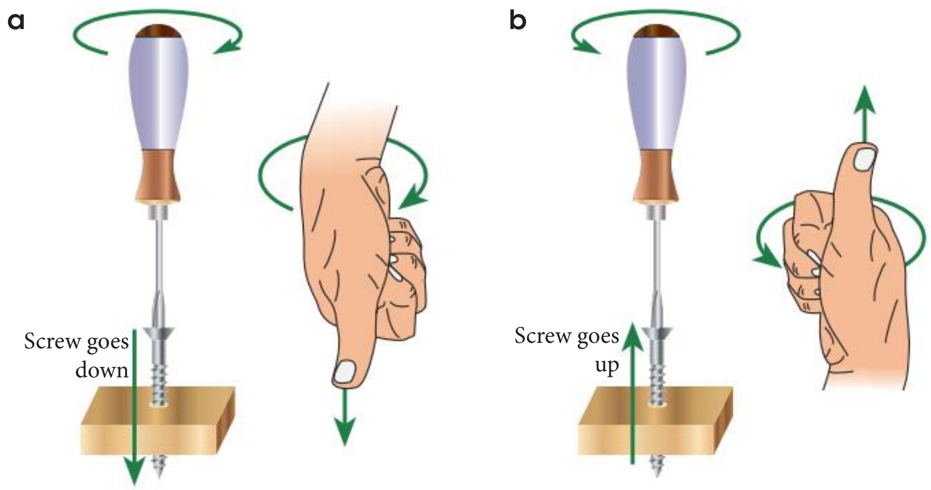
$$\vec{\tau} = \vec{r} \times \vec{F}$$

Vector cross products are used often in physics, and are discussed in more detail in Chapter 4.

WOW

Right-handed thread

Most screws and bolts have 'right-handed' thread as shown in Figure 1.31. To remember which way to turn them, point your right thumb in the direction in which you want the screw to move – either into or out of the material. Your fingers then naturally curve in the direction you should turn the screw: clockwise to tighten and anticlockwise to loosen. Note that a very small fraction of screws are made with left-handed thread. These are generally used in applications where the screw is part of a device that rotates, for example the rotating handle on an apple slinky machine. This makes it less likely that they will come undone during normal operation of the device. There are many right-hand rules in physics, you will meet other useful ones in Chapters 4 and 5.



▲ **Figure 1.31** You turn a screw a) clockwise to tighten it and b) anticlockwise to loosen it.

WORKED EXAMPLE 1.11

A mechanic uses a spanner to tighten a bolt in a car. She uses a spanner with a handle 35 cm long, and applies a force of 95 N perpendicular to the handle.

a What torque does she apply? (3 marks)

b If she applied the force at an angle of 45°, how much force would she need to apply to produce the same torque? (4 marks)

Answers	Logic	
a $\tau = rF\sin\theta$	Relate torque to given data.	1 mark
$\tau = 0.35\text{ m} \times 95\text{ N} \times \sin(90°)$	Substitute values with correct units.	1 mark
$\tau = 33\text{ N m}$	Calculate final value and round to correct significant figures.	1 mark
b $\tau = rF\sin\theta$	Relate torque to given data.	1 mark
$F = \dfrac{\tau}{r\sin\theta}$	Rearrange for force.	1 mark
$F = \dfrac{33\text{ N m}}{0.35\text{ m} \times \sin 45°}$	Substitute values with correct units.	1 mark
$F = 130\text{ N}$	Calculate final value and round to correct significant figures.	1 mark

Try this yourself

If 50 N m is the torque required to turn the bolt, and the mechanic can exert a maximum force of 150 N, what is the minimum length handle she needs on her spanner? (3 marks)

The distance r is often called the **lever arm**, and the longer the lever arm is, the greater the applied torque for a given applied force.

A lever is any device that allows you to apply a force at one point to produce a larger or smaller force at some other point. A seesaw is a simple example of a lever. When you sit on one end, the seesaw rotates about the pivot in the middle because of the torque due to your weight. This produces a force at the other end of the seesaw. The torque produced at any point along the seesaw is the same. Hence, an object close to the pivot point experiences a greater force because the lever arm is smaller here.

A lever is a device for amplifying forces. The longer the lever arm, the more the force is amplified.

Most tools we use are actually levers or combinations of levers. When we wish to exert a large force, we use a long lever arm between the pivot and the applied force, and a short lever arm between the pivot and the load.

The torque applied at the long end is equal to the torque at the load:

$$\tau_{\text{applied}} = r_1 F_{\text{applied}} = r_2 F_{\text{load}} = \tau_{\text{load}}$$

Tools that are designed to take a small applied force and amplify it usually have a long handle, and a short distance from the pivot to the load. Long-handled pruning secateurs are an example of this.

WORKED EXAMPLE 1.12

A small boy sits on a seesaw and asks his father to sit on the other end to balance it. The boy has a mass of 22 kg and the seesaw is 4.4 m long with the pivot at its centre. The father has a mass of 95 kg.

a If the father sits down on the other end of the seesaw and lifts his feet from the ground, how much force is applied to the boy? What will happen? (5 marks)

b Where should the father sit to just balance the boy's weight? (4 marks)

Answers	Logic	
a $r_{\text{father}}F(\text{by father}) = r_{\text{boy}}F(\text{by boy})$	Relate the force applied by the father to the force on the boy.	1 mark
$F(\text{by father}) = F(\text{by boy})$	Recognise that $r_{\text{father}} = r_{\text{boy}}$	1 mark
$F_{\text{f}}(\text{by father}) = w_{\text{father}} = m_{\text{father}}g$	Recognise that the force exerted by the father on the seesaw is equal to his weight.	1 mark
$F(\text{by boy}) = m_{\text{father}}g = 95\,\text{kg} \times 9.8\,\text{m s}^{-2} = 930\,\text{N}$	Substitute values with units and calculate final answer.	1 mark
This force pushes upwards on the boy as the seesaw pivots. The boy's weight (the gravitational force downwards on him) is much less than this, so he accelerates upwards.		1 mark
b $\tau_{\text{father}} = \tau_{\text{boy}}$ $r_{\text{father}}F(\text{by father}) = r_{\text{boy}}F(\text{by boy})$	The torques must be equal in this case if the seesaw is not to rotate.	1 mark
$r_{\text{father}} = \dfrac{(r_{\text{boy}}F(\text{by boy}))}{F(\text{by father})}$	Rearrange for r_{father}.	1 mark
$r_{\text{father}} = \dfrac{(r_{\text{boy}}m_{\text{boy}}g)}{(m_{\text{father}}g)} = \dfrac{(r_{\text{boy}}m_{\text{boy}})}{m_{\text{father}}}$	Recognise that the force each applies is equal to their weight.	1 mark
$r_{\text{father}} = \dfrac{(4.4\,\text{m} \times 22\,\text{kg})}{95\,\text{kg}} = 1.0\,\text{m}$	Substitute values with units and calculate final answer.	1 mark

Try these yourself

1. If the father sits closer to the pivot than 1.0 m, what will happen? Draw a diagram showing the forces acting. (3 marks)
2. If the boy's mother (62 kg) now sits on the seesaw right at the opposite end from her son, is there any point on the seesaw where the father can sit to balance it? If so, find the point and show it on a diagram. (3 marks)

In Worked example 1.12, when the torques applied at each end of the seesaw are equal, the seesaw balances. It does not rotate in either direction. We say that the seesaw is in **equilibrium** because it is not accelerating in a straight line nor is it rotating.

Work and torque

Remember that when we apply a force, *F*, and move an object through a distance, *s*, we are doing work:

$$W = Fs$$

The work done is the energy transferred to the object from the agent doing the work.

When we apply a torque, for example to one end of a lever, if the lever moves and pushes on a load at some other point, we are also doing work. The work done is again equal to the force applied multiplied by the distance through which the force is applied. The lever acts to amplify the force, so that the force acting on the load is greater than the force we apply. So it may at first seem as though a lever can violate conservation of energy by allowing us to do more work at one end of the lever than is done at the other end. However this is not the case.

Think again about the seesaw in Worked example 1.12. Imagine the father sits down at a distance closer to the pivot than that needed to balance his son's weight, before his son gets on the seesaw. His son then gets on at the far end. The force applied by the son to his father (the load) via the seesaw is

$$F_{load} = \frac{(F_{applied}\ r_{applied})}{r_{load}}$$

This force is greater than the weight of the father if he is close enough to the pivot point, and he experiences a net force and hence accelerates upwards. Figure 1.32 shows this. In some time, the boy moves down a distance h_{boy} and the father moves up a distance h_{father}. The triangles shown in Figure 1.32(b) are similar triangles, so

$$\frac{r_{applied}}{r_{load}} = \frac{h_{applied}}{h_{load}}$$

and therefore

$$F_{load} = \frac{(F_{applied}\ h_{applied})}{h_{load}} \quad \text{or} \quad F_{load}\ h_{load} = F_{applied} h_{applied}$$

The work done *on* the load, is equal to the work done *by* the applied force.

This means that although a lever allows us to apply a greater force to the load than that which we apply to our end of the lever, we cannot actually move the load as far as we have to move our end of the lever.

A lever allows a force to be amplified, but the force must be applied through a greater distance to do the same amount of work as the larger force acting through a shorter distance at the other end of the lever.

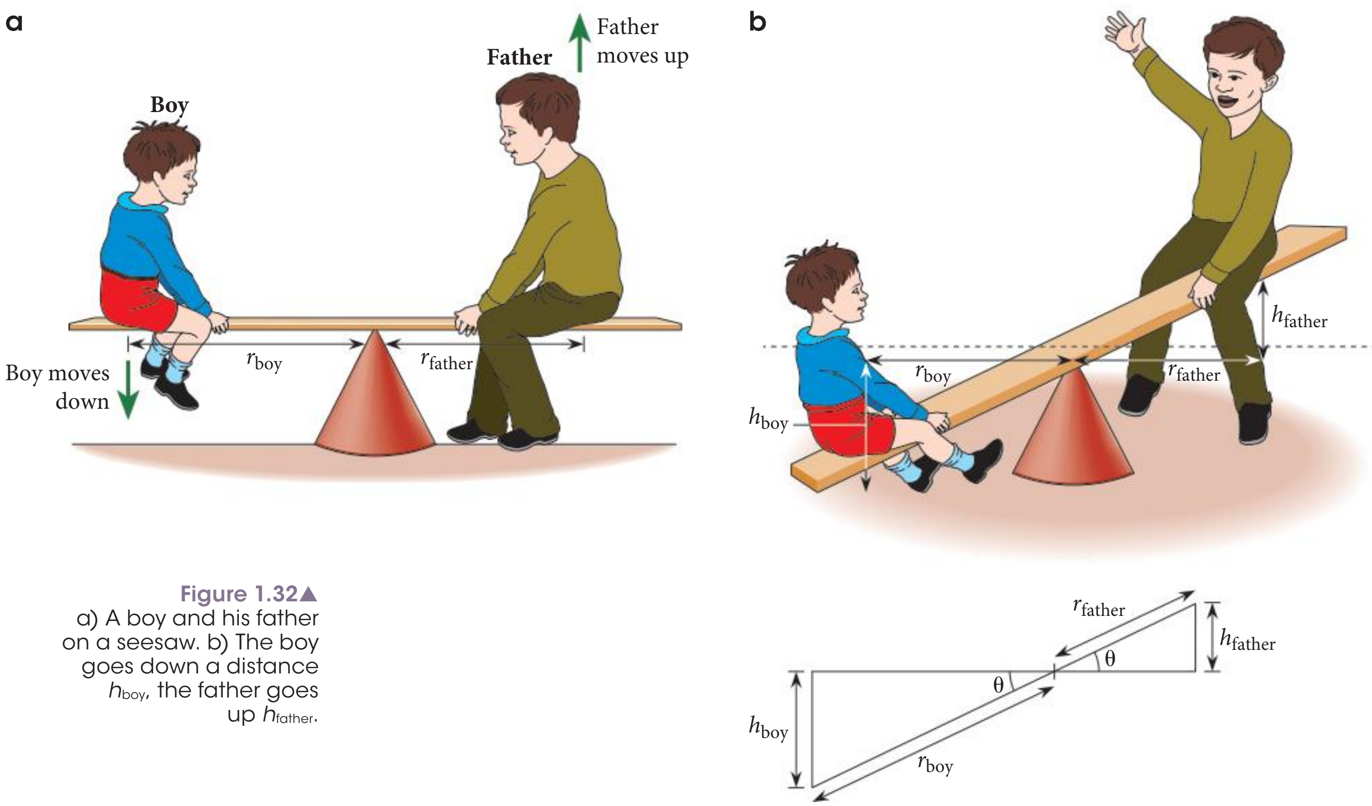


Figure 1.32▲
a) A boy and his father on a seesaw. b) The boy goes down a distance h_{boy}, the father goes up h_{father}.

WORKED EXAMPLE 1.13

A man is using a lever to move a small water tank into position. He uses a rock as the pivot under a long length of wood, and pushes the end of the wood under the tank so that its centre of mass is 0.80 m from the pivot point. The tank has a mass of 250 kg and the man lifts it through a height of 25 cm. The lever arm between the pivot and the man has a length of 3.5 m.

a How much work does the man do on the water tank? (3 marks)

b What force does he apply to do this work, assuming the tank moves slowly at constant speed as it lifts? (4 marks)

c Through what distance must he apply this force? (3 marks)

Answers	Logic	
a $W_{\text{on tank}} = Fs = W_{\text{tank}}h_{\text{tank}} = m_{\text{tank}}gh_{\text{tank}}$	Relate work to data given.	1 mark
$W_{\text{on tank}} = 250\,\text{kg} \times 9.8\,\text{m s}^{-2} \times 0.25\,\text{m}$	Substitute values with correct units.	1 mark
$W_{\text{on tank}} = 610\,\text{J}$	Correct answer given.	1 mark
Note that $W_{\text{on tank}}$ is the work done on the tank, it is an energy and has units of J, W_{tank} is the weight of the tank, it is a force and has units of N.		1 mark
b $F_{\text{man}}r_{\text{man}} = W_{\text{tank}}r_{\text{tank}}$	The torque applied by the man must be equal to the torque applied by the weight of the tank.	1 mark
$F_{\text{man}} = \dfrac{(W_{\text{tank}}\ r_{\text{tank}})}{r_{\text{man}}} = \dfrac{(m_{\text{tank}}\ gr_{\text{tank}})}{r_{\text{man}}}$	Rearrange for F_{man}.	1 mark
$F_{\text{man}} = \dfrac{250\,\text{kg} \times 9.8\,\text{m s}^{-2} \times 0.8\,\text{m}}{3.5\,\text{m}}$	Substitute values with correct units.	1 mark
$F_{\text{man}} = 560\,\text{N}$	Correct answer given.	1 mark
This force is much less than the weight of the tank, which is 2500 N!		
c $W_{\text{by man}} = W_{\text{on tank}}$	Apply law of conservation of energy.	1 mark
$F_{\text{man}}h_{\text{man}} = W_{\text{on tank}}$		1 mark
$h_{\text{man}} = \dfrac{W_{\text{on tank}}}{F_{\text{man}}}$	Rearrange for h_{man}.	1 mark
$h_{\text{man}} = \dfrac{610\,\text{J}}{560\,\text{N}} = 1.1\,\text{m}$	Substitute and calculate final answer	1 mark
You could also use the ratios of the lever arms and the heights to solve part c.		

Try these yourself

A girl (25 kg) sits on one side of a seesaw, 1.8 m from the pivot. Her brother pushes on the other end so that she is slowly lifted 0.75 m up before the seesaw bumps against a stop and she stops moving. (3 marks)

a How much work is done on the girl?

b If the brother applies a constant force of 180 N, at what distance from the pivot must he be pushing down on the seesaw?

c How far down does he push the seesaw at his end?

Equilibrium

When no net force *and* no net torque acts on an object we say that the object is in equilibrium.

No net force does not mean that no forces are acting. It means that all forces, when correctly added as vectors, sum to give a zero total force. There are no **unbalanced forces** acting.

From Newton's second law, $F = ma$, this means that the object has zero acceleration. Hence it has constant velocity.

An object in equilibrium has constant velocity.

The vector sum of all forces acting on the object is zero: $\Sigma\vec{F} = \vec{F}_{net} = 0$

This equation is extremely useful for analysing the forces acting on objects in equilibrium. However we need to remember that forces are vectors, and add them as such. Often it is easiest to break the forces into components and apply the equation of equilibrium in each direction separately. When we do this, we use:

$$\Sigma F_H = 0 \quad \text{and} \quad \Sigma F_V = 0$$

where F_H are the horizontal components of all forces acting, and F_V are the vertical components.

Consider for example the block shown in Figure 1.33(a). When the block is stationary, the only forces acting are the weight of the block and the normal force acting on the block, both in the vertical direction. So we can write:

$$\Sigma F_V = 0, \quad \text{so} \quad N - w = 0, \text{ or } N = w$$

Remember that the normal force is the perpendicular component of the contact force between surfaces. It is only vertical when the surfaces are horizontal. The parallel component is the friction force.

Hence, when no other forces are acting and an object is not moving, its weight is equal to the normal force acting on it.

Now consider the situation in Figure 1.33(b). The block is pulled along at a constant velocity. Constant velocity means the block is again in equilibrium. In the horizontal direction, the forces acting are the horizontal component of the applied force, $F_{applied,H} = F_{applied}\cos\theta$, and the friction force, $F_{friction}$ (the parallel component of the contact force). Hence:

$$\Sigma F_H = 0, \quad \text{so} \quad F_{applied}\cos\theta - F_{friction} = 0, \text{ or } F_{applied}\cos\theta = F_{friction}$$

In the vertical direction, we have three forces: the normal force, N, the weight of the block, w, and the vertical component of the applied force, $F_{applied,V} = F_{applied}\sin\theta$. Applying the equilibrium equation in the vertical direction gives:

$$\Sigma F_V = 0, \quad \text{so} \quad F_{applied}\sin\theta + N - w = 0, \text{ or } F_{applied}\sin\theta + N = w$$

We can see that in this case the normal force is no longer equal to the weight of the block. The normal force is reduced because of the applied force.

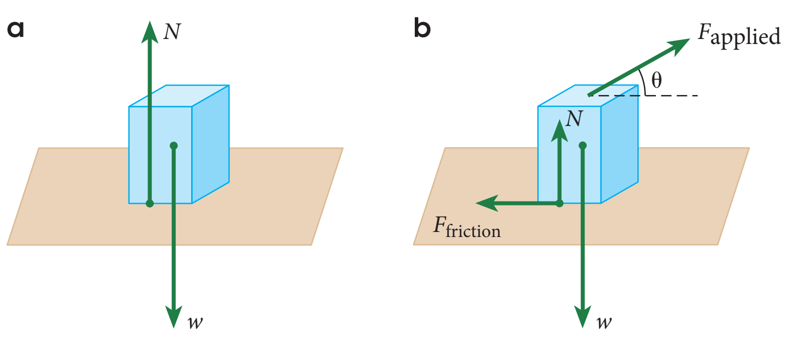


Figure 1.33 ▶
a) $N = w$ for a block at rest on a table.
b) $N + F_{applied}\sin\theta = w$ for a block being pulled at constant velocity.

Some students think that the normal force is always equal to the weight of an object. This is not true. It is only the case when no other vertical forces are acting.

The condition that $\Sigma F = 0$ is a necessary condition for equilibrium, but it is not sufficient. It is possible that there is zero net force acting on an object and that it is not in equilibrium. This can happen when the forces acting on an object are balanced, but they do not act along the same line.

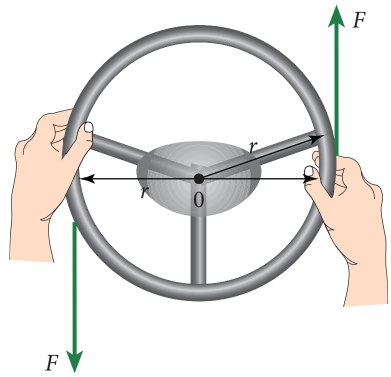
Consider the steering wheel shown in Figure 1.34. The hands exert equal and opposite forces on the wheel, so there is no net force. However each hand also exerts a torque on the wheel, and these torques add to give a total torque twice that of the torque due to one hand.

$$\tau = rF + rF = 2rF$$

A pair of equal, opposite and parallel forces such as those shown in Figure 1.34 are called a 'couple'. They result in a torque called a **couple moment** on the object, which causes the object to rotate but not translate.

Hence the steering wheel is not in equilibrium because its rotational state of motion is changing.

The second requirement for equilibrium is that there is no net torque: $\Sigma\tau = \tau_{net} = 0$.



▲ **Figure 1.34**
Each hand exerts a force, F, on the wheel. The total force is zero. Each hand also exerts a torque, $\tau = rF$. The total torque $\tau_{total} = 2rf$, so the wheel rotates.

We shall use couples to calculate the torque on motors in Chapter 5.

WORKED EXAMPLE 1.14

A 35 kg boy sits on a seesaw at a distance 1.6 m from the pivot. The seesaw has a mass of 45 kg, and is uniform, with the pivot supporting it at its centre.

a Use the equations of equilibrium to find the torque that must be applied to prevent the seesaw rotating. (3 marks)

b If this torque is to be applied by a 25 kg boy sitting on the seesaw, what normal force must be applied by the pivot of the seesaw? (3 marks)

Answers	Logic	
a $\Sigma\tau = 0,\ \tau_1 + \tau_2 = 0$	Apply the equation of equilibrium for torque.	1 mark
$\tau_2 = -\tau_1 = r_1F_1 = r_1m_1g$	Find an expression for torque in terms of data given.	1 mark
$\tau_2 = 1.6\,\text{m} \times 35\,\text{kg} \times 9.8\,\text{m s}^{-2}$ $\tau_2 = 550\,\text{Nm}$	Substitute values including units and calculate final answer.	1 mark
b $\Sigma F = 0,\ F_{pivot} - F_1 - F_2 - w_{seesaw} = 0$	Apply the equation of equilibrium for force, noting that all the forces are in the vertical direction, so we can add them algebraically (taking care with signs).	1 mark
$F_{pivot} = F_1 + F_2 + w_{seesaw} = m_1g + m_2g + m_{seesaw}g$ $= (m_1 + m_2 + m_{seesaw})g$	Find an expression for the force in terms of data given.	1 mark
$F_{pivot} = (35\,\text{kg} + 25\,\text{kg} + 45\,\text{kg}) \times 9.8\,\text{m s}^{-2}$ $F_{pivot} = 1.0\,\text{kN}$	Substitute values including units and calculate final answer.	1 mark

Try these yourself

1 At what point on the seesaw must the 25 kg boy sit to balance the torque due to the 35 kg boy? (3 marks)

2 Use the equations of equilibrium to show that the normal force acting on a table is equal to the weight of the table when there is nothing on the table, but larger than the table's weight when any object is placed on top of the table. Use force diagrams. (3 marks)

QUESTION SET 1.5

Remembering

1 Write the equation for torque and define all the terms used.

2 Write the equations for equilibrium.

Understanding

3 You are having difficulty undoing a screw. Should you use a screwdriver with a longer handle or a thicker handle? Draw a diagram and explain your answer.

4 A person stands next to a table. Explain how they can apply forces to the table to:

  a increase the normal force of the floor acting on them.

  b decrease the normal force of the floor acting on them.

5 Why is it so much harder to hold a heavy bag with your arm outstretched horizontally than to hold it close to your body?

Applying

6 How much force must be applied at a distance of 35 cm from a bolt to produce a torque of 250 Nm on the bolt?

7 If a torque of 250 Nm is applied to a 1.0 cm thick bolt that is embedded in wood and the bolt does not turn, what frictional force is acting on the bolt due to the wood?

8 A car is moving at constant speed down an inclined road at an angle of 12° to the horizontal. The car has a mass of 1500 kg and experiences a drag force of 250 N.

  a Calculate the normal force acting on the car's tyres due to the road.

  b Calculate the friction force acting on the car's tyres due to the road.

  c Draw a diagram showing all the forces acting on the car.

Analysing

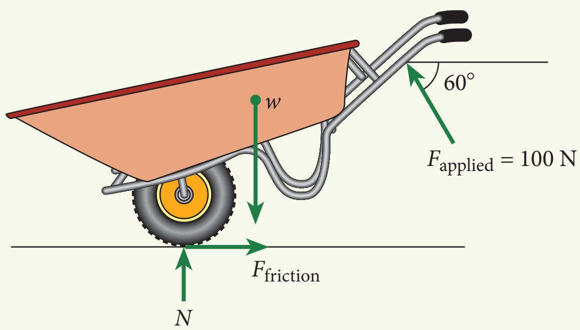
9 A hammer uses a torque to effectively hammer a nail into a piece of wood. Explain, with the aid of diagrams, why this is the case. Hint: Consider what you do with your wrist and arm when you use a hammer.

10 The 15 kg wheelbarrow shown in Figure 1.35 is in equilibrium. The weight acts through the centre of mass, a distance 65 cm from the pivot as shown. A force of 100 N is applied to the handle at 60° to the horizontal. Use the equations of equilibrium to calculate:

a the normal force acting on the wheelbarrow due to the ground.

b the friction force acting on the wheelbarrow due to the ground.

c the distance between the pivot (the wheel) and the applied force.



▲ Figure 1.35

CHAPTER SUMMARY

- A force is applied by one object on another: $\vec{F}$(by A on B)
- Forces are vectors. In vector diagrams forces are drawn pointing away from the point of application. Vectors are added and subtracted geometrically.
- Newton 1: If the vector sum of all forces acting on an object is zero, then its velocity remains constant, including zero velocity.
- Newton 2: $\vec{a} = \dfrac{\Sigma F(\text{on object})}{m}$
- Newton 3: $\vec{F}$(by A on B) and $\vec{F}$(by B on A) are equal in magnitude, opposite in direction, have the same fundamental nature AND each force acts on a different object.
- A gravitational field surrounds each mass and affects other masses.
- The gravitational field near the surface of Earth is $\vec{g} = 9.8\,\text{N kg}^{-1} = 9.8\,\text{m s}^{-1}$.
- Near Earth, the weight force, $\vec{w}$, applied by Earth's mass on a mass, m, is $\vec{w} = mg$.
- Projectile motion can be resolved into independent vertical and horizontal components that can later be combined.
  For initial launch velocity, $\vec{u}$:
  - Horizontally: $v_x = u_x = u\cos\theta$ = constant; thus, $x = u_x t$
  - Vertically: $v_y = u_y = u\sin\theta$

    $v_y = u_y + gt$

    $v_y^2 = u_y^2 + 2gy$

    $y = u_y t + \frac{1}{2}gt^2$
- For motion along an inclined plane, components are taken parallel and perpendicular to the plane,
  - Parallel to the plane, the component of the weight is $F_{//} = mg\sin\theta$.
  - Perpendicular to the plane, the component of the weight is $F_{\perp} = mg\cos\theta$. The normal force has the same magnitude but opposite direction.
- For an object travelling in a circle of radius, r, at constant speed, v, in a time interval, T:

  $v = \dfrac{2\pi r}{T}$

  $a_c = \dfrac{2\pi v}{T} = \dfrac{4\pi^2 r}{T^2} = \dfrac{v^2}{r}$

  - net force, F_{net}, is the sum of real forces such as weight, normal force, tension, friction:

    $F_{net} = ma_c = m\dfrac{2\pi v}{T} = m\dfrac{4\pi^2 r}{T^2} = m\dfrac{v^2}{r}$
- Torque is the equivalent of force for rotational motion. Torque relates to forces that are applied at positions other than the centre of mass:

  $\tau = r_{\perp}F = rF\sin\theta$
- A non-zero net torque causes rotational motion.
- Many tools are levers that use torque to change the magnitude and direction of a force.
- For an object to be in equilibrium two conditions must be met:

  $F_{net} = 0$ and $\tau_{net} = 0$

CHAPTER GLOSSARY

action-at-a-distance non-contact force; one object experiences a force due to the presence of another object that is not touching it

centripetal centre-seeking; applied to force and acceleration on and of objects moving with uniform circular motion

component the part of a vector pointing in a particular direction (usually horizontal or vertical)

contact force a force applied to an object by another object having physical contact

couple moment the torque or moment resulting from two equal but opposite forces acting on an object

equilibrium having a constant state of motion; acted on by no unbalanced forces or torques

friction the component of the contact force between surfaces which is parallel to the surfaces, and resists relative motion of the surfaces

lever arm the distance between the point of application of a force and a pivot point

moment *see* torque

net force the resultant of all the forces acting on an object

normal force a reaction force perpendicular to the surface; a component of the contact force

radius of curvature the radius of a curve that is an arc (part of a circle in shape)

tension the pulling force applied to an object by a string or cable

torque the turning effect of a force; the product of force and distance from the axis of rotation or pivot to the point of application of the force; also called moment

unbalanced force when the resultant of two or more forces acting is not zero, i.e. $\Sigma F \neq 0$

uniform circular motion circular motion with constant speed

vector a quantity that has magnitude and direction

CHAPTER REVIEW QUESTIONS

Remembering

1 What is the horizontal acceleration of a projectile?

2 Define equilibrium.

3 Why are long-handled tools generally more effective than short-handled tools?

4 What general expression can be used to find the horizontal component, u_x of a projectile's initial velocity?

5 What is meant by the 'tension' in a rope?

6 Define 'centripetal acceleration'?

Understanding

7 At the bottom of a rollercoaster ride, the people on the ride feel much heavier than usual. Draw a diagram to show the forces acting and then explain the reason for this sensation.

8 Explain why it is accepted that an object undergoing uniform circular motion is accelerating.

9 Why are the speed limits on windy, curved roads reduced?

10 Why do equal but opposite forces that act along the same line, even if that line does not pass through the pivot point, not result in a torque? Draw a diagram to explain your answer.

11 Explain why curved railway tracks are usually banked towards the inside of the curve. Use a force vector diagram to illustrate your answer.

Applying

12 A ball is thrown vertically upwards at 35 m s^{-1}.

a What is the maximum height reached by the ball?

b How long will the ball take to reach this height?

13 A golf ball is struck at 45 m s^{-1} and an angle of 15° above the horizontal.

a What is the vertical component of the ball's initial velocity?

b Will the ball pass over the head of a 2.0 m high golfer standing 50 m away?

14 The bottom of a rollercoaster ride has a track with a radius of curvature of 28 m. The ride passes this point with a speed of 12 m s^{-1}. What is the normal force on a 50 kg person by the ride's carriage?

15 A car goes over a crest in the road that has a radius of curvature of 34 m. At what speed will the car lose contact with the road?

16 A spanner is used to apply a torque of 500 N m to a nut.

a What minimum force must be applied to achieve this torque if the handle of the spanner is 45 cm long?

b What minimum force must be applied if it is applied at an angle of 25° to the handle?

17 A projectile is fired from ground level at an angle of 45° above the horizontal with a speed of 38 m s^{-1}.

a Find the horizontal and vertical components of the projectile's initial velocity.

b Find the time of flight for the projectile, assuming it is travelling over level ground.

c Find the horizontal distance (the range) of this projectile.

18 A 2.0 kg rock is being whirled around in a horizontal circle held by a 0.70 m long string that will break if the tension in the string exceeds 38 N. Find the shortest period, *T*, for the motion of the rock just before the string breaks. Model the string as approximately horizontal.

19 A train goes around a bend of radius 350 m that is banked at an angle of 8.0°. At what speed should the train travel around the bend?

Analysing

20 A girl is swinging on a maypole in a playground as shown in Figure 1.36.

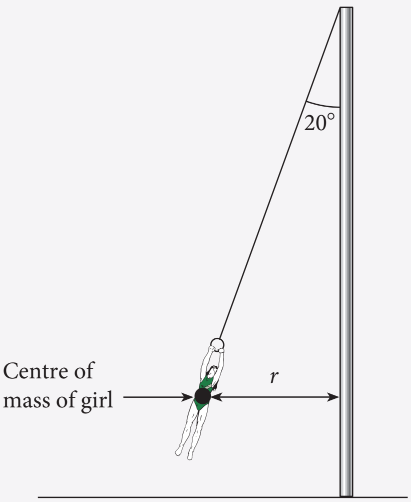


Figure 1.36

The girl has a mass of 36 kg and when she is moving with a speed of 2.0 m s^{-1} the light rope makes an angle of 20° with the vertical. Consider the motion of the centre of mass of the girl, which moves in a horizontal circle of radius *r*.

a What is the vertical component of the tension in the rope?

b What is the horizontal component of the tension in the rope?

c What is the tension in the rope?

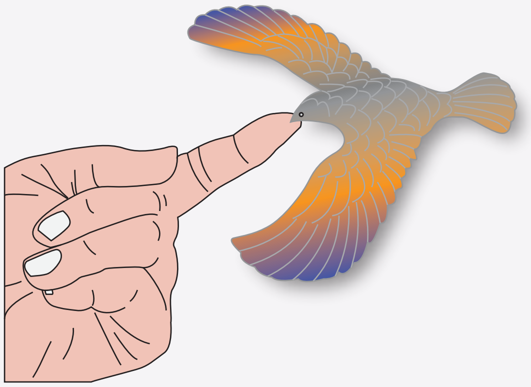
d What is the net force acting on the girl?

e What is the radius of the circle?

21 How many times more centripetal force is required to act on a vehicle moving around a curve with a radius of curvature of 200 m compared with a curve with a radius of curvature of 400 m if the vehicle is travelling safely at the same speed?

22 A projectile is fired at an angle of 20° above the horizontal. Another projectile is fired at the same speed at an angle of 70° above the horizontal. Show that the horizontal distances travelled by both projectiles is the same.

23 A balancing bird toy (as shown in Figure 1.37) can balance on a finger. Model the bird as being made of three parts: the body and two wings. Draw a diagram showing the forces acting on each part of the bird, and the torques resulting from these forces. Write the equations of equilibrium for the bird based on the forces and torques shown in your diagram. Explain how the bird balances.



▲ Figure 1.37

Reflecting

24 How has the content of this chapter assisted you in your understanding of the causes of car crashes that occur on a bend on a country road?

25 Make a spider diagram for all the physics you have learnt in this chapter which has helped you understand how a car works. Include all the relevant forces you can think of, as well as torques. Draw a diagram showing where the different forces and torques act.

CHAPTER 2
GRAVITY

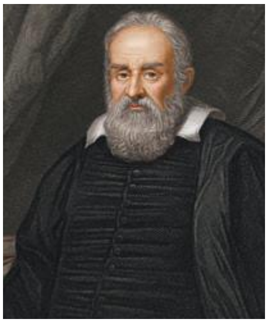
Shutterstock.com/Brian A. Jackson

By the end of this chapter you will have covered the following material.

Science Understanding

- The movement of free-falling bodies in Earth's gravitational field is predictable (ACSPH093)
- All objects with mass attract one another with a gravitational force; the magnitude of this force can be calculated using Newton's Law of Universal Gravitation (ACSPH094)
- Objects with mass produce a gravitational field in the space that surrounds them; field theory attributes the gravitational force on an object to the presence of a gravitational field (ACSPH095)
- When a mass moves or is moved from one point to another in a gravitational field and its potential energy changes, work is done on or by the field (ACSPH096)
- Gravitational field strength is defined as the net force per unit mass at a particular point in the field (ACSPH097)
- Newton's Law of Universal Gravitation is used to explain Kepler's laws of planetary motion and to describe the motion of planets and other satellites, modelled as uniform circular motion (ACSPH101)

Introduction



Alamy/Stock Montage Inc.

Figure 2.1 ▲ Galileo Galilei

Kinematics was studied in Nelson Physics Units 1 & 2 for for the Australian Curriculum.

Why do things fall to the ground? To us, this seems obvious: gravity pulls things down towards Earth. For Aristotle (384–322 BCE) and his contemporaries, the answer was also obvious: things fell to the ground because they were made of earth, and earth naturally moved towards Earth. They had a kind of heaviness or **gravitas** that enabled them to fall straight down. In Aristotle's world, there was no need for our modern concept of gravity.

A long appraisal of Aristotle's ideas led to the emergence of a way of understanding motion that we would now recognise as kinematics – the relationship between measurements of distance and time. It was not until the fifteenth and sixteenth centuries that experiments on projectiles and the motion of falling objects were carried out. The most significant of these experiments were undertaken by Galileo (1564–1642 CE). He showed experimentally that falling objects accelerated more or less uniformly towards Earth.

In 1687, Newton (1643–1727 CE) finally showed how an **inverse-square law** of gravitational force could account for this acceleration. This law was 'universal' because it incorporated all motion, from Earth to the ends of the universe.

Newton's gravity was built on the impressive measurements of Tycho Brahe (1546–1601 CE) and the mathematical interpretation of these data by Johannes Kepler (1571–1630 CE).

Earlier work by Nicolaus Copernicus (1473–1543 CE), itself indebted to the accuracy of Muslim astronomers, such as Muhammad al-Battani (c. 868–929 CE), and Galileo on the motion of the planets around the Sun, as well as data from the Royal Observatory at Greenwich (1675 onwards) contributed to Newton's confidence in the universality of his gravitational theory.

Newton's universal gravitational law remained undisputed until Albert Einstein (1879–1955 CE) made significant modifications in his 1915 paper on general relativity. These changes, and the upsurge in high-quality astronomical observations – Earth-based and from space – during the last hundred years have hugely enhanced our knowledge of the universe. Theories of dark matter and dark energy are significantly based on our understanding of gravity.

PROJECTILE MOTION

Read about the history of the analysis of projectile motion by Galileo.

Alamy/Hemis

Figure 2.2 ▲ Monument to Tycho Brahe and Johannes Kepler in Prague

Universal gravitation

Edmund Halley (1656–1742 CE), Robert Hooke (1635–1703 CE) and Newton all recognised that the elliptical orbits of planets, first described by Kepler, could be explained by a force that depended on distance from the Sun. Newton coined the word 'gravity' to name this force. He wrote to Halley: 'It is now established that this force is *gravitas*, and therefore we shall call it gravitas from now on.' We see here that Newton used the Aristotelian idea of 'heaviness' to describe what we now refer to in English as gravity.

In 1687, Newton published *Philosophiae Naturalis Principia Mathematica* in which he described the law of universal gravitation.

We can imagine that every mass has a **gravitational field**, g, surrounding it. This field reaches to infinity. The gravitational field of a mass, M, exerts a force on another mass, m, as shown in Figure 2.5. When the masses are not in contact, the force is an action-at-a-distance force. We say that the force is mediated by the field.

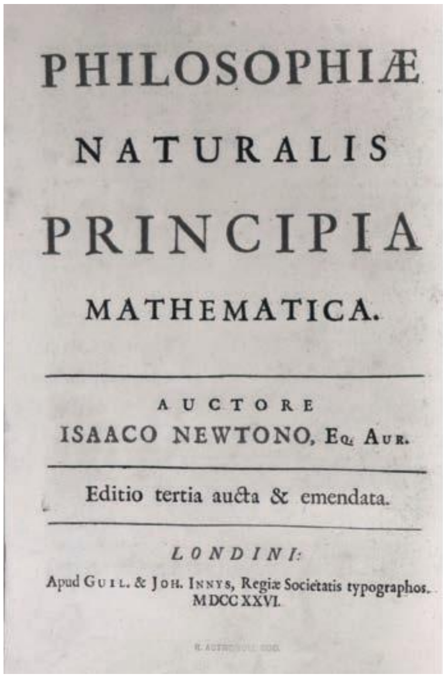
When measuring the local value of the gravitational field due to M, we try not to disturb the field too much by a large mass, so we use a small test mass. The force on the small test mass, m, due to the gravitational field of M causes m to accelerate. It is this acceleration that is the field strength:

$$g_M = \frac{F(\text{by } M \text{ on } m)}{m}$$

As force is a vector, field is a vector that points in the same direction as the gravitational force.

Corbis/The Gallery Collection

▲ Figure 2.3
Sir Isaac Newton



PHILOSOPHIÆ
NATURALIS
PRINCIPIA
MATHEMATICA.

AUCTORE
ISAACO NEWTONO, Eq. Aur.

Editio tertia aucta & emendata.

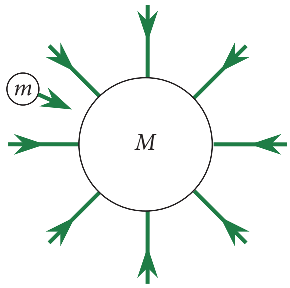
LONDINI:
Apud Guil. & Joh. Innys, Regiæ Societatis typographos.
MDCCXXVI.

Alamy/The Art Archive

▲ Figure 2.4
Philosophiae Naturalis Principia Mathematica, 1686, by Isaac Newton

HISTORY OF GRAVITY

Read about the history of gravity and the people behind the advances in our understanding.



▲ Figure 2.5
The gravitational field surrounding a mass, *M*, is the acceleration of *m*

Gravitational force and gravitational field are models that explain action-at-a-distance.

- Gravitational force is modelled as acting at the centre of mass of an object.
- Gravitational field is modelled as the force per unit mass, directed towards the centre of mass of an object.
- Gravitational force and gravitational field are vectors.

The gravitational field has the units of N kg^{-1}, which, as we have seen previously, is the same as the units of acceleration, m s^{-2}. Thus, if we can measure the acceleration of a mass placed near M, we can find the value of the gravitational field.

Newton showed that the gravitational force, which would keep the planets in their orbits, was an inverse-square law.

The force by M on m was dependent on both masses, as well as the inverse of the square of the distance, r, between m and M. The distance, r, is measured from **centre of mass** to centre of mass of objects. Thus:

$$F(\text{by } M \text{ on } m) = G\frac{Mm}{r^2}$$

It follows from the definition of gravitational field that the field of M at distance r from M is given by:

$$g = G\frac{M}{r^2}$$

Notice that the field associated with M does not depend on the mass m in the field. The gravitational field due to M exists whether we put another mass nearby or not.

You saw in Nelson Physics Units 1 & 2 for the Australian Curriculum Chapter 11 that an inverse-square $\left(\frac{1}{r^2}\right)$ *distribution is characteristic of any point-like or spherical source because of the symmetry of the source.*

The gravitational field near Earth is $9.8\,\text{N kg}^{-1}$*. This means that all objects, independent of their masses, fall at the same acceleration,* $9.8\,\text{m s}^{-2}$*. This result was used in Unit 2 and in the previous chapter.*

Field models are powerful tools that have excellent explanatory and predictive power. The gravitational field model will be joined by field models of electromagnetism and the strong and weak nuclear forces.

Newton's law of universal gravitation:

$$F(\text{by } M \text{ on } m) = G\frac{Mm}{r^2}$$

Gravitational field surrounding a mass, M:

$$g = G\frac{M}{r^2}$$

Gravitational field strength is measured by finding the acceleration of a small test mass in the field.

Units of gravitational field: N kg^{-1} or m s^{-2}

GRAVITY

Watch this short explanation of gravity.

WOW

Isaac Newton and the Black Death

In 1665, while Newton was studying at Cambridge University near London, an outbreak of the Black Death, or bubonic plague, swept through London and Cambridge. The university was closed, sending Newton back to his mother's apple orchard, where he spent many hours contemplating, among other things, the path of the Moon in its circle around Earth. He concluded that the force that made an apple fall to the ground is the same force that keeps the Moon in its orbit. Newton returned to Cambridge, but it was to be another 22 years before he finally published his law of universal gravitation.

STRENGTH OF GRAVITY

The Cavendish experiment was the first to measure the strength of the gravitational constant *G*.

Newton's third law and universal gravitation

We have seen that the field due to M acts on a mass, m (Figure 2.6), with a force:

$$F(\text{by } M \text{ on } m) = G\frac{Mm}{r^2}$$

But, what about the force applied by m on M? This force has the same magnitude:

$$F(\text{by } m \text{ on } M) = G\frac{mM}{r^2}$$

This is just what Newton's third law describes. Recall that for two masses, the gravitational forces:

- are equal in magnitude
- are opposite in direction
- have the same fundamental nature

AND

- each force acts on a different object.

The acceleration of m is the field strength of M at distance r: $g_M = G\frac{M}{r^2}$

The acceleration of M is the field strength of m at distance r: $g_m = G\frac{m}{r^2}$

Thus, m and M accelerate at different rates, but the magnitude of the force applied to each is the same. For example, the gravitational field of Earth acts on a 100 g apple with a force of 0.98 N. The apple accelerates at $9.8\ \text{m s}^{-2}$. The apple acts on the Earth with the same force, 0.98 N, but Earth's mass, approximately 5.97×10^{24} kg, will not accelerate very much!

No matter how much you may be tempted to add these two gravitational forces to make zero, this is not possible. Adding forces is what Newton's second law describes. But Newton's second law refers to forces acting on a single object. These forces act on different objects, m and M respectively; thus, they cannot be added.

The universal gravitational constant, *G*

In 1798, seventy-one years after Newton's death, Henry Cavendish (1731–1810) measured the value of the constant of proportionality, G, in Newton's law of universal gravitation. He placed very massive lead balls near two very much smaller balls at the end of a long rod, as in Figure 2.7.

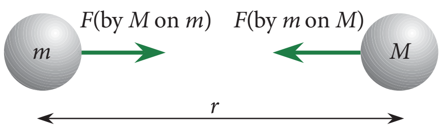
The forces applied to each of the smaller balls caused a rotation of the rod. This rotation was opposed by the metal suspension line. The amount the suspension line rotated was used in the calculations to find G.

Cavendish measured the value of G to within about 1% of the currently accepted value:

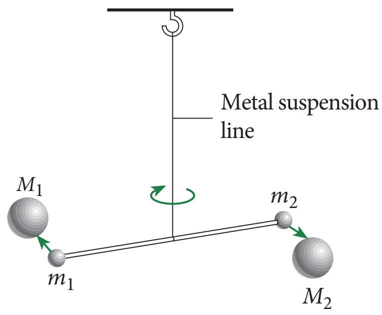
$$G = 6.673\,84 \times 10^{-11}\,\mathrm{N\,m^2\,kg^{-2}}$$

CAVENDISH AND THE VALUE OF *G*

Learn about the Cavendish experiment.



▲ **Figure 2.6**
The force by one mass, M, on the other mass, m, is the same magnitude, but oppositely directed, to the force applied by m on M.



▲ **Figure 2.7**
Cavendish's measurement of the universal gravitation constant, G.

WORKED EXAMPLE 2.1

What is the gravitational force of attraction between Earth and the Sun? (2 marks)

- Mass of Earth: 5.97×10^{24} kg
- Mass of the Sun: 2.0×10^{30} kg
- Radius of Earth's orbit around the Sun: 1.5×10^{11} m

Answer	Logic	
$F = G\dfrac{Mm}{r^2}$	Force of attraction between Earth and the Sun is given by an equation.	1 mark
$= 6.67 \times 10^{-11}\,\mathrm{N kg^{-2} m^2} \times \dfrac{(2.0 \times 10^{30}\,\mathrm{kg})(5.97 \times 10^{24}\,\mathrm{kg})}{(1.5 \times 10^{11}\,\mathrm{m})^2}$	Substitute the correct values with units.	$\frac{1}{2}$ mark
$= 3.6 \times 10^{22}\,\mathrm{N}$	Calculate the answer.	$\frac{1}{2}$ mark

Try this yourself

Given the mass of the Moon is 7.35×10^{22} kg and its mean orbital radius about Earth is 3.84×10^{8} m, find the gravitational force of attraction between Earth and the Moon. (2 marks)

WOW

The hunt for gravity waves

Einstein's theory of general relativity describes gravity as an artefact of the warping of the fabric of space–time by the presence of a mass. This can be visualised by thinking of space–time as the surface of a trampoline and a mass as being a person standing on the trampoline. The stretching of the trampoline means that a ball will roll around the person in the same way that an object moves around a mass in space. When events such as two black holes colliding occur in the universe, it is believed that ripples in space–time will spread out, rather like ripples on the surface of a trampoline. Australian physicists are looking for variations in the precise signals from pulsars received by the Parkes radiotelescope in an effort to detect gravity waves and test the theory of general relativity.

QUESTION SET 2.1

Remembering

1 Define 'gravitational field'. How does gravitational field differ from gravitational force?

2 a Write an algebraic expression for:
   i gravitational field.
   ii gravitational force.
  b What two variables determine the acceleration due to gravity at the surface of a planet?

Understanding

3 'Gravitational field does not depend on the mass, *m*, in the field.' Explain this statement.

4 What is the difference between Aristotelian *gravitas* and Newtonian *gravitas*?

5 Two masses, *m* and *M*, are in an isolated system. The gravitational forces, *F*(by *M* on m) and *F*(by *m* on *M*), are equal and opposite. Why do they not add to a zero net force?

Applying

6 Two masses, *m* and *M*, have a gravitational force of attraction *F* when they are a distance *r* apart. What is the force when this distance is increased to 4*r*?

7 What is the gravitational force of attraction between a 500 kg satellite with an orbital radius of 7000 km and Earth? (Mass of Earth = 5.97×10^{24} kg)

8 What is the acceleration of a rock dropped onto the Moon's surface? The radius of the Moon is 1700 km, and its mass is 7.3×10^{22} kg.

Analysing

9 Why could Cavendish claim that his experiment was 'weighing the Earth'? Support your answer with quantitative calculations. (Gravitational field strength = 9.8 N kg^{-1}, radius of Earth = 6370 km)

Reflecting

10 Are you satisfied that masses really do attract other masses by fields that stretch out to infinity? Explain.

Forces such as the normal force and the friction force are manifestations of the fundamental forces. The friction force and the normal force are due to the interaction of electrons on the surfaces of objects.

Field theory does not explain why interactions are not instantaneous. This limitation of field theory was one factor that led to the development of a different model of forces – the exchange particle model. This model is discussed in detail in Chapters 9 and 10.

Gravity and fields

Each fundamental force (gravitational, electromagnetic, strong and weak) can be described as acting via a **field**. These fundamental forces are all action-at-a-distance forces. They all allow us to explain how one object is able to exert a force on a second object without being in contact with it.

The gravitational field model allows us to explain how objects can exert forces without being in contact. It also allows us to:

- predict the acceleration of an object in any gravitational field.
- calculate the mass of an object from the observed force it exerts on another object.
- calculate the mass of distant objects such as planets by observing their orbits about the Sun.

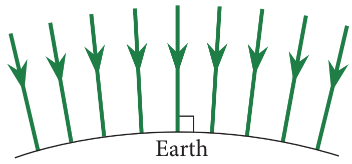
Measuring the acceleration of objects dropped on the surface of the Moon tells us about the mass of the Moon. The radius of the Moon can be measured from astronomical observations, and then combining the size and mass information tells us that the Moon has a density very similar to that of Earth's crust. This information was important in the development of modern theories of the formation of the Moon. These theories say that the Moon was actually formed when a massive object collided with Earth and a chunk broke off – that chunk stayed in orbit and became the Moon.

Gravity 'near Earth'

At Earth's surface, we are always subject to the force of attraction applied to masses by the mass of Earth. Unless otherwise constrained, all objects fall from a height to the surface with an acceleration of $9.8\,m\,s^{-2}$. This is the effect of Earth's gravitational field on the masses.

'Near Earth' is an approximation. The field lines in a local area are very nearly parallel to each other and very nearly strike the surface at right angles (Figure 2.8). Up to several kilometres – well above the tallest buildings – the field strength varies by very little. We say there is **negligible** variation in field strength over any local region.

The approximation that the gravitational field is constant is reasonable when close to the surface of Earth. The field decreases as $\frac{1}{r^2}$ above Earth's surface. Once an object is at the height of the International Space Station (about 400 km), the gravitational field is approximately 90% of that at Earth's surface. For satellites that may be several thousand kilometres above Earth's surface, the near Earth approximation cannot be used because g is substantially less than $9.8\,m\,s^{-2}$.



▲ **Figure 2.8**
The gravitational field is perpendicular to the surface and directed to the centre of Earth.

Always think carefully about your choice of model and approximations. The uniform field approximation for near Earth is reasonably close to the surface of Earth. However, it is not a good choice for modelling the behaviour of satellites.

WORKED EXAMPLE 2.2

What is the gravitational field strength on the surface of the Moon? (3 marks)

Use:

- M: $mass_{Moon} = 7.3 \times 10^{22}\,kg$
- r: $radius_{Moon} = 1740\,km$

Answer	Logic	
$g = G\dfrac{m}{r^2}$	Use the correct formula.	1 mark
$= 6.67 \times 10^{-11}\,Nm^2\,kg^{-2}\,\dfrac{7.3 \times 10^{22}\,kg}{(1740 \times 10^3\,m)^2}$	Substitute the correct values with units.	1 mark
$= 1.6\,m\,s^{-2}$	Calculate the answer.	1 mark

Try these yourself

1 What is the gravitational field strength for a satellite orbiting Earth at an altitude of 36 000 km? (3 marks)
Use $mass_{Earth} = 6.0 \times 10^{24}\,kg$; $radius_{Earth} = 6370\,km$.

2 Earth's gravitational field strength is $4.9\,m\,s^{-2}$ at a satellite. How far away from Earth's centre must the satellite be? (3 marks)

ACTIVITY 2.1

THE INVERSE-SQUARE LAW

Aim

To model the inverse-square law

You will need

Access to a computer with Excel or another spreadsheet application installed

What to do

1. In column A, list the numbers 1–10.
2. In column B, insert a formula so that the result is the inverse of the corresponding row in column A.
3. In column C, insert a formula to square the value of the corresponding row in column B.
4. Insert a scatter graph for the values of columns C (the *y* axis) and column A (the *x* axis).

What did you discover?

1. What shape is the resulting line of best fit?
2. What does this relationship show for the values of the cells in columns A and C?
3. How does this model the relationship shown in the inverse-square law?
4. What is the mathematical relationship in an inverse-square law?

Weight, apparent weight and weightlessness

Weighing scales such as kitchen and bathroom scales do not display weight; they display mass. They measure the weight force and then convert it into a mass, which is displayed in kilograms.

Mass and **weight** are different. Mass is an intrinsic property of an object, which depends on how many and what sort of atoms make up the object.

Weight is the gravitational force applied to a mass:

$$w = mg$$

It follows that weight depends on where the mass is placed in relation to other masses.

Our common experience of weight is related to the mass of Earth. The only significant mass affecting our mass is the approximately 6.0×10^{24} kg of mass below our feet. Weighing machines measure the force applied by this mass, then convert it to a mass reading on the scale. For example, a person with a mass of 50 kg will usually be measured by the weighing machine as having a weight, $w = 50\,\text{kg} \times 9.8\,\text{N}\,\text{kg}^{-1} = 490\,\text{N}$; yet the scale will show 50 kg.

Mass and weight are different.

Mass is a measure of how much matter (the number and type of atoms) in an object.
- Unit: kilogram, kg
- It *does not* vary with the position of the object.

Weight is the gravitational force acting on an object.
- Unit: newton, N
- It *does* vary with the position of the object.

The confusion that many people have between weight and mass comes from the way in which we typically measure mass. When you measure your mass, you usually do so by standing on a set of scales. If you stand still on the scales, on a still, level floor, Earth applies a weight force on you: $w = mg$. You exert a force on the scales downwards: *F*(by you on scales). According to Newton's third law, the scales exert a force on you upwards: *F*(by scales on you). This force is equal in magnitude to the force you applied downwards.

The forces acting *on you* can be used to produce an equation for the net force (Newton's second law).

If you and the scales are in equilibrium, the net force on you is zero:

$$\Sigma F(\text{on you}) = 0$$

$$\Rightarrow N(\text{by scales on you}) + w(\text{by Earth on you}) = 0$$

$$\Rightarrow N(\text{by scales on you}) = -w(\text{by Earth on you})$$

Hence we can deduce that, at equilibrium, the normal force exerted on you by the scales is equal and opposite to the gravitational force by Earth on you: your *weight*. The normal force by the scales on you, N(by scales on you) is the way we measure the force you applied to the scales, N(by you on scales). As has been indicated, this is a consequence of Newton's third law.

That is, the scales really measure the normal force that you exert on them: N(by you on scales). Usually this is done by measuring the compression of a spring, or a piezoelectric crystal. The scales are calibrated to convert this compression, which depends on the force, into a *mass* that is displayed in kilograms. The actual calibration process of going from distance compressed to mass can be quite complicated, but it takes into account the factor of $g = 9.8\,\text{N kg}^{-1}$ in the conversion. So the quantity that is displayed on the screen or pointed to on the scale is your mass in kilograms. *It is not your weight.*

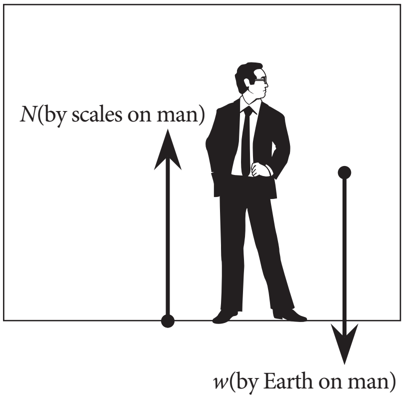
We use the normal force on an object to measure the weight force on the object. The mass of the object is then inferred from this measure.

The normal force is only equal to the weight if the following three conditions all apply:
- **The surface exerting the normal force is horizontal.**
- **No vertical forces other than weight and the normal force are acting on the object.**
- **The object being weighed is in equilibrium.**

The normal force you exert on the scales may not be the same as your weight. You can, for example, push up on a nearby table while standing on the scales. The table will then push down on you, which will increase your normal force on the scales. Consequently, the scales will push with a greater normal force upwards and you will apparently have a larger mass, hence weight. Similarly, you can apparently decrease your weight by pushing down on the table.

Figures 2.9–2.11 show a man in a lift. Imagine that the man is standing on a set of scales. The reading on the scales will change, depending on whether he is accelerating up or down. His mass will not change. Nor will his weight – the gravitational force in Earth's field – change (except for the negligible change associated with change in distance from Earth's centre).

We shall discuss each figure in terms of the forces applied to the man; that is, we shall infer from changes to the mass reading that a change to the normal force applied to the man by the scales, hence by the man to the scales, has occurred. In each case, the force measured will be his **apparent weight**, as it is what the scales suggest. In the static case, of course, his apparent weight is his true weight in Earth's gravitational field. In the following we take down as the negative direction and consider the magnitudes of the force.



▲ Figure 2.9
In a stationary lift the weight force and the normal force are equal: $N - mg = 0$.

Figure 2.9 shows the man in equilibrium, neither accelerating up nor down.

$$N(\text{by scales on man}) - w(\text{by Earth on man}) = 0 \text{ (Newton's second law)}$$

$$\Rightarrow N(\text{by scales on man}) = w(\text{by Earth on man})$$

$\Rightarrow$ The scales measure the weight of the man in Earth's gravitational field.

Figure 2.10 shows the man accelerating upwards.

$$N(\text{by scales on man}) - w(\text{by Earth on man}) = ma \text{ (Newton's second law)}$$

$$\Rightarrow N(\text{by scales on man}) = w(\text{by Earth on man}) + ma$$

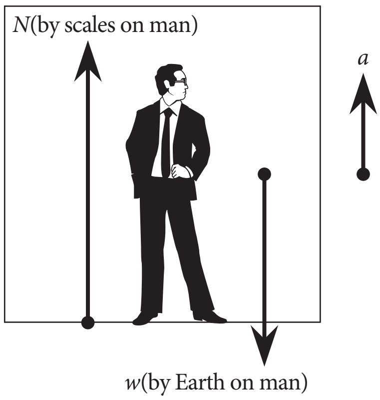


Figure 2.10 ▲
In a lift accelerating upwards, the normal force is greater than the weight: $N = w + ma$.

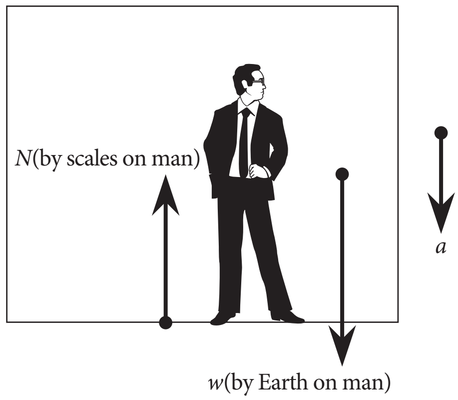


Figure 2.11 ▲
In a lift accelerating downwards, the normal force is less than the weight force: $N = w - ma$.

⇒ The scales measure an apparent weight that is greater than the weight of the man in Earth's gravitational field.

Figure 2.11 shows the man accelerating downwards.

$$w(\text{by Earth on man}) - N(\text{by scales on man}) = ma \text{ (Newton's second law)}$$

$$\Rightarrow N(\text{by scales on man}) = w(\text{by Earth on man}) - ma$$

⇒ The scales measure an apparent weight that is less than the weight of the man in Earth's gravitational field.

Imagine now that the lift is in **free fall**. The man and the lift will be accelerating at the same rate. The man cannot apply a normal force to the scales and the scales cannot apply a normal force to the man. The scales will register zero: the man's apparent weight will be zero and he will experience **apparent weightlessness**. However, his weight in the field is still $w = mg$.

Your apparent weight is related to how you *feel* in an accelerating reference frame. Although your weight in the field is still $w = mg$, you feel 'heavier' when you are being pushed upwards, and 'lighter' when you are accelerating downwards. In the limit that there is no normal force acting on you, you feel 'weightless'. This is the case when you are in free fall. In this case, you are falling with the acceleration due to gravity because the only significant force acting on you is your weight. This is why astronauts 'float' in space in the International Space Station. They are actually falling with the local

Figure 2.12 ▶
The crew of the International Space Station is apparently weightless in their environment.

Alamy/Universal Images Group Limited

Scientific literacy: Testing the gravitational inverse-square law

Henry Cavendish (1731–1810) was the first person to measure the universal gravitational constant, G. He used a system of very large masses to attract other, smaller masses. By a cunning arrangement involving a torsion balance, he was able to deduce a value of G to 1% accuracy (Figure 2.13(a)). This was an extraordinary undertaking in the 18th century. Cavendish understood the importance of extraneous variables on the results. He took action to ensure these effects were minimised. His accuracy was not bettered for more than 100 years!

In order to achieve this result, Cavendish calibrated the torsion balance using known forces to twist the wire. He developed an empirical relationship between the angle of twist and the force applied by the wire in reaction to the force applied to the wire. This is the force per unit twist angle. A torsion balance makes use of torque. Torque is the rotational equivalent of translational force. A force, F, applied at a perpendicular distance, $r_\perp$, from a point of suspension or axis of rotation causes a rotation or torque, τ: $\tau = r_\perp F$. For the torsion pendulum, the torques from both sides cause a twist. When the torsion wire is twisted enough its restoring force per twist angle is just equal to the combined torques of the two masses. The forces measured in this way are of the order of 10^{-10} N.

Very much more precise equipment can now be used to measure the universal gravitation constant. The current value of the universal gravitational constant is $G = (6.673\,84 \pm 80) \times 10^{-11}\,\mathrm{m^3\,kg^{-1}\,s^{-2}}$.

> Will the law of universal gravitation always be true? Nothing seems more certain than the 'fact' that there are three dimensions of space. But can we be sure that there are only three dimensions? Imagine a tightrope walker balancing on a cable high above the ground. To the tightrope walker the cable is effectively a 1D object. But an ant sees the cable as a 2D object, because it can crawl along and also around the cable. Today, increasing numbers of physicists are seriously questioning whether we are like tightrope walkers, unaware of the true number of dimensions in space. They suggest that the best way to discover the dimensions of space is to study how the gravitational attraction between two objects depends on the distance between them. If the universe contains more than three spatial dimensions our current laws of gravity should break down at small distances.
>
> One of the outstanding challenges in physics is to finish what Newton started and achieve the ultimate 'grand unification' of gravity with the other three fundamental forces (electromagnetic force, strong force, weak nuclear forces) into a single theory. String theorists, for example, use 1D strings and higher-dimensional 'branes' rather than familiar point-like particles. String theorists seriously entertain the idea that there are actually six or seven additional spatial dimensions, which are needed to make the theory both mathematically consistent and capable of describing gravity.
>
> Adelberger, E., Heckel, B. & Hoyle, C.D. (2005) 'Testing the gravitational inverse-square law', *Physics World*, April 2005, http://physicsworld.com/cws/article/print/2005/apr/03/testing-the-gravitational-inverse-square-law

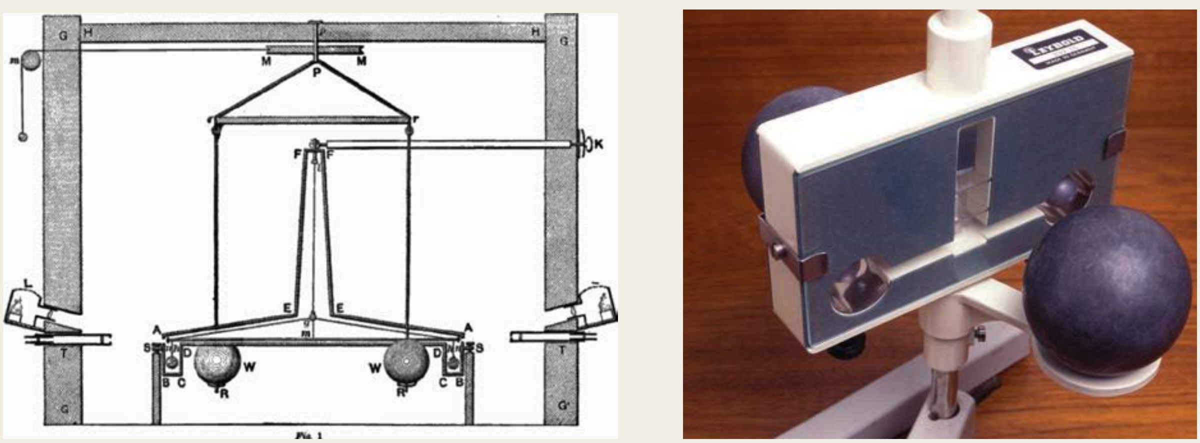
HENRY CAVENDISH

Find out about the first person to measure the universal gravitational constant.

EXPLANATION OF CAVENDISH'S EXPERIMENT

The apparatus Cavendish used to measure the value of the universal gravitational constant G is shown in this video.

a b



Wikipedia/Creative Commons

Alamy/sciencephotos

Figure 2.13 ▲
a) Cavendish's sketch of his equipment and b) a modern precision instrument used to measure the universal gravitational constant $6.67384 \times 10^{-11}\,\mathrm{m^3\,kg^{-1}\,s^{-2}}$

Questions

1 Draw a simplified diagram of Cavendish's experiment to show the position of the large masses, the smaller masses and the forces that lead to the twisting of the torsion wire.

2 List the quantities that Cavendish measured. Show how these quantities relate to Newton's equation:

$$F = G\frac{m_1 m_2}{r^2}$$

3 What is the percentage uncertainty in the currently accepted value of G? If you were to undertake this experiment in a school laboratory would you be justified in using Cavendish's result or the current value? Give reasons.

4 Draw and label a diagram to show distances, forces and torques associated with a torsion pendulum.

5 How is the analogy between a person and an ant on a tightrope relevant to the development of understanding about gravitational force? (Conduct further research if necessary.)

WEIGHTLESSNESS

View videos of astronauts enjoying weightlessness.

Apparent 'weightlessness' occurs when the only force acting on objects is weight!

acceleration due to gravity (about 90% that of Earth's surface *g*), but so is the spaceship in which they are travelling. The spaceship can only exert a normal force on astronauts (or anything else) when they push against it.

True weightlessness occurs when the gravitational field at a point is zero. For example, the Moon exerts a gravitational force towards Earth and the Earth exerts a gravitational force towards the Moon. Objects between the Moon and Earth and close enough to the Moon and far enough from Earth will experience a zero net force. As the local value of '*g*' is zero, a mass at this point is weightless (but not massless).

WOW

Lagrangian points

Lagrangian points are points in space at which peculiar things happen with the gravitational fields of Earth, Moon and Sun. One such point about 1.5 million km towards the Sun from Earth is known as L1. A spacecraft can be positioned here so that it orbits the Sun with the same period as Earth. Earth's gravitational field opposes that of the Sun's sufficiently for this to occur. The solar observatory SOHO is positioned at L1, among a few other observatories. SOHO monitors the solar wind and can give us a few hours warning of the approach of dangerously high levels of energetic particles before they hit Earth's magnetic field.

Other forces at a distance

Gravitational force is not the only force that acts at a distance. The other such forces are the electrostatic forces between charged particles, the magnetic force that is easily observed using magnets and compasses, and the strong and weak nuclear forces that act over very small distances within the nuclei of atoms. Without these two forces atoms would not be stable, but they are not detectable over distances usually encountered in everyday life.

When forces act at a distance, the force is a result of the objects involved interacting with the surrounding field. It takes a finite time for the interaction to be transmitted from one object to the other. The effect is not instantaneous, as is sometimes thought. **Field theory** does not explain why interactions are not instantaneous. This limitation of field theory was one factor that led to the development of a different model of forces – the exchange particle model. This model is discussed in detail in Chapters 9 and 10.

Case study

Dr George Hobbs

George Hobbs is a research scientist at the CSIRO Astronomy and Space Sciences division based in Marsfield, Sydney. He has received the Young Tall Poppy, NSW Scientist of the Year award as well as the CSIRO Julius Career Award. Included in his busy schedule is the Pulse@Parkes student program, by which either school excursions can come to CSIRO or CSIRO can come to the school. The Pulse@Parkes experience allows students to take control of the world famous 64 m Parkes radiotelescope for a 90-minute session. During this time, signals from pulsars are received and the information stored, ready for analysis. Pulsars are believed to be collapsed remnants of large stars that have reached the end of their lives as stars. Gravity is so strong in these dense remains that matter as we know it collapses, forcing electrons and protons to combine. The resulting material, made of neutrons, is millions of times more dense than normal matter. Pulsars rotate rapidly on their axes and emit a narrow beam of electromagnetic radiation in the radio frequency part of the spectrum. When these beams sweep over Earth, radiotelescopes detect a short burst of a radio signal. The time between of these signals is very regular.

George Hobbs and a team of astronomers and astrophysicists are looking for variations in these signals that may be caused by a warping of the fabric of space–time. Einstein's theory of general relativity predicts that this warping may occur when very massive objects, such as black holes, combine. Then 'ripples', or gravity waves, are sent out through space at the speed of light.

Pulsars, which are often located at the centres of nebulae – gas and dust left over from exploding stars – are also used by Dr Hobbs as distance indicators in space. The various frequencies of radio waves being detected from a pulsar travel at slightly different speeds through space. This causes a small delay in their arrival. The speed difference is due to the presence of interstellar electrons at very low densities.

◀ **Figure 2.14**
Dr George Hobbs helping students control the Parkes radiotelescope during a Pulse@Parkes session.

Questions

1 What type of matter do astronomers believe are the source of the radio signals being used to look for gravity waves?
2 Why would the variation on a signal from a pulsar be considered abnormal?
3 Why are pulsars often found at the centres of nebulae?
4 Why do you think the discovery of gravity waves will be considered such a big breakthrough?
5 Why would Dr Hobbs consider it important to spend time with school students on programs such as Pulse@Parkes?

QUESTION SET 2.2

Remembering

1 How can a gravitational field model enable us to explain action-at-a-distance?

2 a Distinguish between mass and weight.
  b Explain how weight is measured using a set of weighing scales.

3 For a non-zero gravitational field, define:
  a weight.
  b apparent weight.
  c apparent weightlessness.

Understanding

4 Why is the gravitational field of a planet considered an 'acting-at-a-distance' force even though it may be applying forces to objects that are in contact with it.

5 Explain why the Moon's gravitational field exerts a greater force on Earth than the Sun's gravitational field, even though the Sun's mass is many millions of times greater.

Applying

6 What is the gravitational field strength on the surface of a planet with mass 3.0×10^{24} kg and a radius of 4000 km?

7 What is the apparent weight of a 40 kg child in a lift that is descending with an acceleration of 1.8 m s^{-2}?

Analysing

8 An astronaut on the Moon drops a ball from a height of 1.50 m. It takes 1.36 s for the ball to fall to the surface of the Moon. Given that the Moon has a radius of 1.74×10^6 m, calculate its mass.

9 At what distance from Earth's centre is the net gravitational field of Earth and Moon zero? Take the distance between Earth and Moon to be 3.84×10^5 km. ($M_{\text{Earth}} = 5.97 \times 10^{24}$ kg, $M_{\text{Moon}} = 7 \times 10^{22}$ kg)

Reflecting

10 Draw a concept map or spider diagram to show the relationships between mass, weight, gravitational field, gravitational force, measurement of mass and weight, normal force, apparent weight, weightlessness.

Gravitational potential energy

See Chapter 8 of Nelson Physics Units 1 & 2 for the Australian Curriculum.

There are two types of energy: kinetic energy and **potential energy**. Energy is transferred when a force acts over a distance. When a force, F, acts on an object and moves it through some displacement, s, in the direction of the force, **work** is done:

Recall from Chapter 1 of Nelson Physics Units 1 & 2 for the Australian Curriculum that the definition of work is the energy transferred due to the action of a force.

$$W = Fs$$

Both force and displacement are vectors; however, work is done only when the force and the component of the displacement are parallel to each other.

When work is done on an object its kinetic energy changes. When the component of the displacement is in the same direction as the force, the work done causes the kinetic energy of the object to increase. At the same time, the potential energy of the system must decrease so that energy is conserved. Conversely, when the component of the displacement is in the opposite direction to the force, kinetic energy decreases and the potential energy increases.

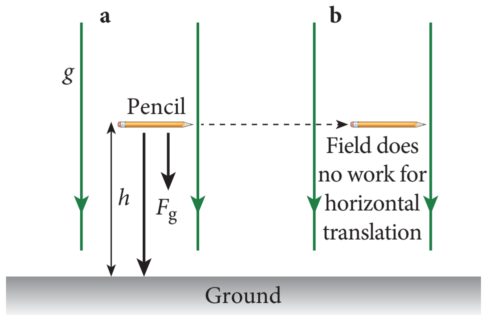
Thus, the change in kinetic energy is the opposite of the change in potential energy:

Remember that potential energy belongs to a system of interacting objects. It is not meaningful to refer to the potential energy of a single object.

$$\Delta E_k = -\Delta E_p$$

Gravitational potential energy is due to the interaction of objects via their gravitational fields. It is the gravitational field that mediates or exerts the force on one object due to the mass of another. Hence we say that the gravitational field does work when an object falls in Earth's gravitational field. In this case, the work is done by the field on the object because the object moves in the direction of the field. The kinetic energy increases and the potential energy decreases.

Consider a pencil allowed to fall in Earth's gravitational field, as in Figure 2.15(a). If we define our system as the pencil and Earth, the gravitational field of Earth does work on the pencil, increasing its kinetic energy. The potential energy of the Earth–pencil system has decreased. If we then pick up the pencil and lift it through some height, we must apply a force in the direction opposite to the gravitational field (Figure 2.15(a)). We do work equal to Fs on the pencil, where F is the force we apply through a distance s. If the pencil begins and ends at rest, there is no change in the pencil's kinetic energy. Yet we have done work, so energy must have been transferred. The field was also doing work at the same time. In this case the work done by the field was negative. The total work done *on the pencil* is zero. The total work done *on the Earth–pencil system* is the amount of work that we have done, and is equal to the change in potential energy of the Earth–pencil system. We say that this work was done *on* the field. So the energy transferred by the application of the force appears as an increase in the potential energy of the system.



▲ Figure 2.15
a) When we lift the pencil at constant speed, we do work $F_{applied}h$ and the gravitational field does work F_gh. Work is done on the field and the kinetic energy is unchanged, but the potential energy of the pencil–Earth system is increased. When the pencil falls through a distance of h, the field does work F_gh on the pencil, increasing its kinetic energy. b) When we move the pencil horizontally, no work is done on or by the field and there is no change in potential energy.

When you hold the pencil stationary, both you and the gravitational field are exerting a force on it. However the displacement is zero, so no work is done by either force. This is also the case when the displacement is perpendicular to the field, that is in the horizontal direction. In this case there is no component of displacement in the direction of the field, so the work done by or on the field is zero (Figure 2.15(b)).

The gravitational potential energy belongs to the *system*, which is both the object creating the field *and* the object experiencing a force due to the field. But, you might wonder, where is the energy stored? A pencil does not contain gravitational potential energy. Gravitational potential energy is *not* stored in the object.

We need to think about the relationship between force and potential energy again. The field is able to do work because it exerts a force. So we can model the potential energy as being stored *in the field*. The energy is distributed throughout all space where the field exists. The energy in a given volume (the energy density) depends on the field strength.

In Chapter 3 we shall see that potential energy is also stored in an electric field. We use a quantity called potential, which you met in Chapter 3 of Nelson Physics Units 1 & 2 for the Australian Curriculum, to represent the energy changes as a charge moves in an electric field.

In field theory, we model the action-at-a-distance forces, including gravity, as being mediated by a field. The field applies a force to objects in the field.

Because the field is able to do work, we say that the potential energy of the system is stored in the field.

Choosing a zero of gravitational potential energy

To be able to say how much potential energy a system has, we need to be able to define a zero energy position or configuration for the system. Note that we are again talking about *systems*, not isolated objects.

Energy transfers and systems are described in Chapters 1 and 2 of Nelson Physics Units 1 & 2 for the Australian Curriculum.

Both kinetic and potential energy cannot be defined for an isolated object. An object only has potential energy because a force is exerted on it, and the force must have some agent. Kinetic energy must be measured against some reference frame.

Choosing a zero for the potential energy of the Earth–pencil system may seem obvious. If we take the zero as being when the pencil is on the ground, then the potential energy of the system when the pencil is at any height, h, above the ground is simply $mg\Delta h = mg(h - 0) = mgh$. Recall from *Nelson Physics Units 1 & 2 for the Australian Curriculum* that we can arrive at this conclusion by considering the work done to lift the pencil to this height. The work done must be equal and opposite to the work done by the gravitational field if the pencil is to begin and end at rest, so:

$$W = \Delta E_P = F_g s$$

$$\Rightarrow \Delta E_P = mg\Delta h$$

The potential energy of an object is always dependent on other objects, which generate the force field. Even kinetic energy is not truly the property of a single object because it is due to motion, which is always relative to other objects in a frame of reference.

This is a useful working definition when considering forces and motion close to Earth's surface. However, you must be careful to define exactly what you mean by the surface or ground level, as this may vary according to the situation. It also means that the potential energy of any Earth–object system becomes negative whenever the object falls below the defined zero level. There is nothing wrong with a negative potential energy. The negative sign means that the potential energy at the end is less than the potential energy at the beginning. We only ever measure changes in potential energy. These changes can be positive or negative.

This 'ground level' definition for zero potential energy is not very useful as soon as we are talking about other planets, or the solar system. Nor is it useful if we want to be able to look at the behaviour of things that move a long way above the surface of Earth. The simplest way to define a zero that everyone can agree on, and that is not based on any single particular object, is to take the zero as being when all objects in a system are infinitely separated.

Consider a system of massive objects very far apart from each other. When all the objects in the system are infinitely separated, the forces acting on them are zero. We define the potential energy of this configuration as zero. If the objects are not moving there is no kinetic energy, so the total energy of the system is zero.

The same convention for choosing the zero of potential energy for a system is used in electrostatics, as we shall see in Chapter 3. However the electrostatic force can be attractive or repulsive, so the potential energy of a system of charges can be either positive or negative.

The gravitational force is always attractive. Any change from this zero configuration lowers the potential energy of the system to a negative value. The gravitational field due to each object does work on the other objects, bringing them closer together. The work done by the fields decreases the potential energy of the system, as it attracts the objects closer together. This means that the kinetic energy must increase to compensate. As they accelerate closer together, their kinetic energy increases.

The law of conservation of energy means that no change occurs to the total energy in a system:

$$\Delta E_{\mathrm{T}} = 0$$
$$\Rightarrow \Delta E_{\mathrm{k}} + \Delta E_{\mathrm{p}} = 0$$
$$\Rightarrow \Delta E_{\mathrm{k}} = -\Delta E_{\mathrm{p}}$$

This is the general result, which we saw previously in the particular case of an object moving in Earth's gravitational field.

The zero of gravitational potential energy is defined to be when the components of the system are infinitely separated.

All other configurations have a negative potential energy.

Changes in energy in a gravitational field

When an object moves in the direction of the gravitational field, the gravitational field does positive work, the potential energy of the system decreases. If no other forces are acting, the kinetic energy of the object increases.

When an object moves against the gravitational field in an isolated system (where no other force is exerted on it) the gravitational field does negative work and the kinetic energy of the object decreases. The potential energy of the system increases.

In an open system, work can be done on the objects in the system by an external agent, for example you lifting a pencil in the Earth–pencil system. In this case, if an object is moved against the field, the potential energy of the system again increases, but there may or may not be a change in kinetic energy.

We can summarise these changes in Table 2.1.

Table 2.1 Summary of energy changes

System	Object moves	Work is done	Potential energy	Kinetic energy
Closed	With the field	by the field	decreases	increases
	Against the field	on the field	increases	decreases
Open	With the field	by external agent	decreases	increases
	Against the field	by external agent	increases	either

Note that when work is done by an external agent to move an object in a field, the field still does work also. The work done by the field is positive if the object moves with the field and negative if it moves against the field.

WORKED EXAMPLE 2.3

A 10 kg school bag is lifted onto a shelf 2.0 m above the ground.

a How much work was done on the bag? (2 marks)

b By how much did the bag's gravitational potential energy change? (1 mark)

c If the bag fell off the shelf, with what kinetic energy would the bag land on the ground? (1 mark)

Answers	Logic	
a $W = mg \times s$	Using work done, $W = F \times s$, and $F = mg$ we get this equation.	1 mark
$= 10\text{ kg} \times 9.8\text{ m s}^{-2} \times 2.0\text{ m}$ $= 196\text{ J}$	Substitute the correct values with units.	1 mark
b $\Delta E_p = 196\text{ J}$	The bag's E_p changed by the same amount as the work done on the bag.	1 mark
c $E_k = 196\text{ J}$	Assuming there is no air friction, $\Delta E_p = -\Delta E_k$.	1 mark

Try these yourself

1 **a** How much work is done on the gravitational field when a 60 kg diver falls through a vertical height of 3.0 m? (2 marks)

b Using energy conservation, with what speed will the diver enter the water? (3 marks)

2 A 400 kg rocket is launched from ground level. When it is at an altitude of 100 m its speed is 50 m s^{-1}. (5 marks)

a What is the E_k of the rocket when the rocket is at 100 m altitude?

b How much work was done on the rocket to change its gravitational potential energy?

c How much work in total was done on the rocket?

QUESTION SET 2.3

Remembering

1 **a** Write the conservation of energy law in terms of energy changes.

b When an object falls in a gravitational field, potential is reduced. Where does this energy go to?

2 Define:

a kinetic energy.

b potential energy.

c gravitational potential energy.

Understanding

3 When an object near the Earth moves horizontally the gravitational potential energy in the field does not change. Explain.

4 Explain why a falling object is having work done on it by the gravitational field.

Applying

5 How much work is done in raising a 200 kg mass through a vertical height of 30 m?

6 The gravitational potential energy associated with an object is -6.0×10^3 J. The object is not moving.

- **a** What is its kinetic energy? Explain.
- **b** What is the minimum amount of work that must be done to ensure it does not reach Earth? Give reasons.

Analysing

7 Explain why the gravitational field of an asteroid is small compared to Earth's.

8 400 J of work is done on a stationary 5.0 kg mass to raise it from a position 100 m above the ground.

- **a** How far above the ground is it raised?
- **b** When dropped from its new height, what is its speed as it passes its original position?
- **c** How fast is it travelling when it strikes the ground? Assume all other forces are negligible.

Reflecting

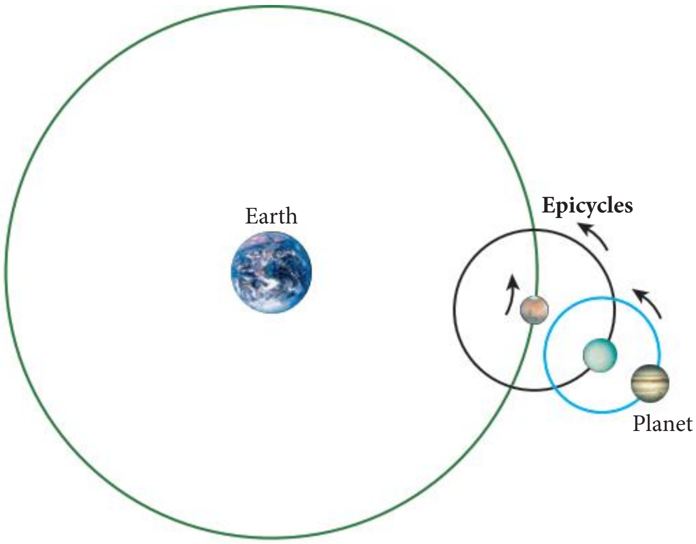
9 Discuss the reasons why defining zero gravitational potential energy as being an infinite distance away avoids problems that other definitions might have.

Models of planetary motion

Models of planetary motion have been proposed for centuries. The ancient Greek philosopher Plato had the heavenly bodies fixed in **concentric** crystalline spheres (Figure 2.16). The stars were fixed in one sphere while the Moon and Sun had different spheres. The word 'planet' is Greek for 'wanderer'. The planets were observed to wander across the sky. This example of a model made to explain observations is one of many developed over the centuries as thinking changed and the models improved.

Ptolemy's later models retained the circles and added **epicycles**, circles on circles (Figure 2.17). These were required in the models so that the motion of the planets in the sky could be explained.

Figure 2.16 ▶ Plato's model of the universe had concentric crystalline spheres containing the heavenly bodies



◀ Figure 2.17
Epicycles, circles on circles, were added to Plato's model of the universe.

Better observations required better models. However, the models quickly became complex, with no explanation of how the epicycles were maintained.

Kepler's laws of planetary motion

Kepler inherited the volumes of Tycho Brahe's meticulous observations of the motions of the planets, the Moon and the stars. Built up over many years, Brahe's measurements and recordings, all made before the invention of the telescope, enabled the mathematically minded Kepler to propose a new model for the motion of the planets.

GREEK ASTRONOMY

Read a detailed account of the development of models of the universe from the early Greeks to modern times.

Kepler's first law: the law of orbits

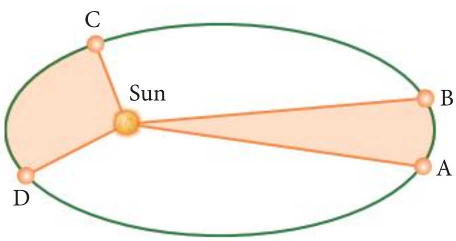
It had always been assumed that the planets orbited Earth, and in later models, the Sun, in perfectly circular orbits. This was, in part, due to the belief that the heavens were perfect and that circles were considered to be a perfect shape. Moving in anything but a circle had not been proposed previously. Kepler found that, if the planets were considered as moving in elliptical orbits, then their observed positions in the sky could be predicted almost perfectly.

Kepler's analysis involved angular momentum. Angular momentum is the rotational equivalent of momentum in linear motion. We will use angular momentum in Chapter 10.

Kepler's first law: All planets move in elliptical orbits with the Sun at one focus.

An ellipse is a curved shape, such as that shown in Figure 2.18. It has a major or long axis, which is the longest line between two points on the edge drawn through the geometric centre. The minor axis of an ellipse is the shortest line joining two points on the edge drawn through the geometric centre. An ellipse has two foci. These are special points such that any line drawn from one focal point to any point on the edge, and then to the other focal point has the same length.

A circle is a special case of an ellipse in which the two axes are equal and the two foci are at the same position. The orbits of most planets, moons and satellites are very close to circular, with some exceptions.



▲ Figure 2.18
Segments AB and CD are swept out in an equal time interval.

Kepler's second law: the law of areas

Kepler noticed that the speeds of the planets changed during their orbits. Nearer to the Sun their speeds increased; further away their speeds decreased. He was able to conclude that the areas covered in equal time intervals were the same.

Kepler's second law: A line that connects a planet to the Sun sweeps out equal areas in equal times.

Kepler's third law: the law of periods

By doing further work on Tycho Brahe's data and using his own observations, Kepler showed that there is a relationship between the average radius of orbit of the planets and their period of revolution around the Sun.

Kepler's third law: The square of the period of a planet's orbit is proportional to the cube of the mean radius of its orbit:

$$T^2 \propto r^3$$

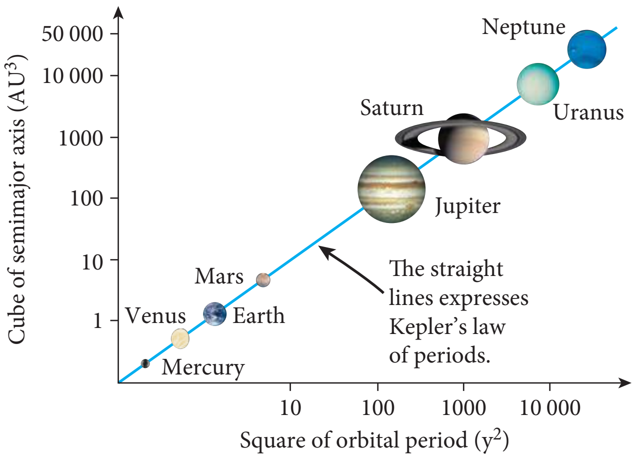


Figure 2.19 ▲ A graph of r^3 versus T^2 for the planets in our solar system

Kepler arrived at his three laws empirically. He based the laws on an analysis of the data Brahe had given him, and his own observations. They were not based on any underlying models or theories. Kepler's laws gave excellent predictive power, but as they had no theoretical basis; they did not give any explanation of the observed behaviour of the planets. Newton's model for gravity and the principle of conservation of momentum provided the theoretical framework needed to explain *why* planets and other orbiting bodies move as described by Kepler's laws.

Newton's law of universal gravitation and Kepler's third law

We can deduce Kepler's third law from Newton's universal gravitation law as follows. For simplicity, we assume that the planet follows a circular orbit, although a more complex geometrical analysis of elliptical orbits gives the same result. We use the equations we derived for circular motion in Chapter 1.

A planet (mass m) orbits the Sun (mass M) at a distance, r, from the Sun. The only force applied to the planet is the force mediated by the gravitational field of the Sun. The circular motion equations can be used:

$$F(\text{by Sun on planet}) = G\frac{mM}{r^2}$$

$$F(\text{net on planet}) = m\frac{4\pi^2 r}{T^2}$$

$$\Rightarrow m\frac{4\pi^2 r}{T^2} = G\frac{mM}{r^2}$$

$$\Rightarrow \frac{T^2}{r^3} = \frac{4\pi^2}{GM}$$

This ratio is a constant for all objects that revolve around the Sun. This is what was expected from Kepler's third law, namely:

$$T^2 \propto r^3$$

$$\Rightarrow T^2 = kr^3 \ (k = \text{constant})$$

$$\frac{T^2}{r^3} = k$$

Figure 2.19 shows this relationship. Notice that the radius is given in units of AU, which is the Earth's orbital radius (approximately 1.5×10^8 km). The period is given in years, y.

The analysis does not apply just to planetary motion around the Sun. It applies to all satellites orbiting much larger masses, for example, the moon and satellites orbiting the Earth or moons orbiting Jupiter. The value of the constant will change from central mass to central mass because the constant $\frac{4\pi^2}{GM}$ depends on the mass, M, around which the bodies are circulating. This means that k is not a universal constant, but a constant that is particular to a given system.

WORKED EXAMPLE 2.4

1 Find the period T for a satellite of Earth with an orbital radius of 42 000 km. Take Earth's mass as 6.0×10^{24} kg. (4 marks)

2 A small planet is observed to orbit a star every 30 days. A second planet orbits the same star at a distance that is nine times the orbital radius of the first planet. What is the period of the second planet? (4 marks)

Answers	Logic	
1 $\frac{T^2}{r^3} = \frac{4\pi^2}{GM}$	Use the correct formula.	1 mark
$T^2 = \frac{4\pi^2}{GM} \times r^3$	Rearrange for T.	1 mark
$T = \sqrt{\frac{4\pi^2}{(6.67 \times 10^{-11}\ \text{N kg}^{-2}\ \text{m}^2) \times (6.0 \times 10^{24}\ \text{kg})} \times (42000 \times 10^3\ \text{m})^3}$	Substitute the correct values with units.	1 mark
$= 8.5 \times 10^4$ s (approx. 1 day)	Calculate the answer.	1 mark
2 Here, the mass M of the star is unknown. The term $\frac{4\pi^2}{GM}$ is the same for both planets as they are orbiting the same star, so we can use: $\frac{T^2}{r^3} = \text{constant}$	Use the correct strategy.	1 mark
$\frac{T_1^2}{r_1^3} = \frac{T_2^2}{r_2^3}$ and $r_2 = 9r_1$, $T_1 = 30$ days	Use the correct equation.	1 mark
$T_2^2 = \frac{T_1^2 \times r_2^3}{r_1^3}$		
$T_2 = \sqrt{\frac{(30\ \text{days})^2 \times (9\,r_1)^3}{r_1^3}}$	Rearrange for T_2.	1 mark
$T_2 = 810$ days	Calculate the answer.	1 mark

Try these yourself

1 A spacecraft orbits the Moon with a period of 4.0 hours. It is then lowered into an orbit with half the radius. What is its new orbital period? (4 marks)

2 What is the mass of a star if its planet has an orbital period of 95 Earth days and an orbital radius of 5.0×10^{10} m? (4 marks)

Satellite motion

The motion of a satellite can be modelled as uniform circular motion. Most satellites have circular or very nearly circular orbits around Earth. They are in a constant state of free-fall; the only force acting on them is gravity or their weight. The gravitational force by Earth's mass on a satellite is directed towards Earth's centre, which is also the centre of the satellite's circular orbit. Therefore, the net force acting on the satellite is perpendicular to the velocity of the satellite.

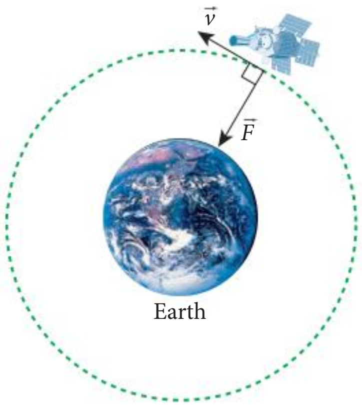


Figure 2.20 ▲ The net force on a satellite is perpendicular to its velocity and directed towards the centre of its circular orbit.

SATELLITES

Watch the video clip of this satellite launch. Observe the angle of the launch vehicle shortly after lift-off.

Orbital speed

The gravitational force applied to a satellite is the centripetal force keeping the satellite in its circular orbit (Figure 2.20).

The gravitational force can be found from $F_G = G\dfrac{Mm}{r^2}$, which provides the centripetal force $F_c = \dfrac{mv^2}{r}$.

Equating these two forces gives:

$$F_c = F_G$$

$$\frac{mv^2}{r} = G\frac{Mm}{r^2}$$

$$v^2 = \frac{GM}{r}$$

$$v = \sqrt{\frac{GM}{r}}$$

Orbital velocity enables a satellite to fall at just the rate necessary to return to its original position:

$$v = \sqrt{\frac{GM}{r}}$$

WORKED EXAMPLE 2.5

What is the speed of any satellite orbiting Earth at an altitude of 630 km? (3 marks)

Mass of Earth = 5.97×10^{24} kg

Radius of Earth = 6370 km

Answer	Logic	
$r = 6370 + 630 = 7000$ km	Calculate the radius of the satellite's orbit.	1 mark
$v = \sqrt{\dfrac{GM}{r}}$	Select the appropriate equation.	1 mark
$= \sqrt{\dfrac{(6.67 \times 10^{-11}\,\text{N kg}^{-2}\,\text{m}^2)(5.97 \times 10^{24}\,\text{kg})}{7000 \times 10^3\ \text{m}}}$	Substitute known values into the equation with units.	1 mark
$= 7.5 \times 10^3\,\text{m s}^{-1}$ or $7.5\,\text{km s}^{-1}$	Calculate the answer.	1 mark

Try these yourself

1 Describe what would happen to the orbital speed of a satellite if the satellite was moved to an orbit with a higher altitude. (1 mark)

2 Find the altitude of a satellite in orbit around Earth given that its orbital speed is $5.0\,\text{km s}^{-1}$. (4 marks)

Satellite orbits

An orbiting satellite is still very much within Earth's gravitational field. It is falling to Earth with an acceleration equal to the gravitational acceleration g at that distance from Earth. An orbiting spacecraft or satellite is given a horizontal velocity so that, as it falls, it is also moving horizontally with such a speed that its path is circular. If no force other than that due to the gravitational field acts on the satellite, the satellite will continue in its orbit forever.

In order to escape the effect of the Earth's gravitational field, extra energy must be expended.

Escape velocity is the speed needed for an object to leave the gravitational field of a planet or other large mass so that it no longer experiences a gravitational force due to that large mass. It can be derived by considering the initial and final kinetic energies of the object. This change in kinetic energy is equal to the change in potential energy of the system. The initial potential energy is negative, and the final potential energy must be zero if the object is to escape the gravitational field.

The escape velocity can be shown to be related to the distance of the object from the source of the field, and the mass of the body from which it is to escape:

$$v_{\text{escape}} = \sqrt{\frac{2GM}{r}}$$

For objects launched from the surface of Earth, this becomes

$$v_{\text{escape}} = \sqrt{\frac{2GM_{\text{Earth}}}{R_{\text{Earth}}}} = 1.12 \times 10^4\,\text{m s}^{-1}$$

Note that this is the escape speed ignoring the rotation of Earth. The rotation of Earth can be used to assist in achieving escape velocity by launching the object at the appropriate angle and close to the equator where the speed due to rotation is greatest.

Escape velocity is the minimum velocity required by an object to escape the gravitational field of a large mass:

$$v_{\text{escape}} = \sqrt{\frac{2GM}{r}}$$

Satellites in **low Earth orbit** are high enough to be modestly affected by atmospheric friction but low enough to be relatively easily serviced from Earth. This region extends from about 250 km to 1000 km above Earth's surface.

A **geostationary** satellite remains above one place on Earth. It must travel directly above a point on the equator. **Geosynchronous** satellites travel above any great circle. A great circle is any circle on Earth whose radius extends from Earth's centre. Both geostationary and geosynchronous orbits have approximately 24 hour periods (23 h 56 min 4 s or 86 164 s).

WOW

Low Earth orbit satellites

The International Space Station (ISS) has been in low earth orbit since November 1998. Its altitude averages 370 km. Even at this height, its rocket motors must be fired every few weeks to boost it into a higher orbit. The friction of the very low density atmosphere found at this altitude is sufficient to slow the ISS down. As this happens, the orbital radius decreases. At lower altitudes there is more friction and the whole process would result in the ISS eventually crashing back to Earth in a fiery re-entry.

sts-114 Crew/NASA

▲ **Figure 2.21**
The International Space Station in low Earth orbit

QUESTION SET 2.4

Remembering

1 a State Kepler's three laws.
  b Why are Kepler's laws referred to as empirical laws?

2 Distinguish between orbital velocity, orbital acceleration and escape velocity.

Understanding

3 Draw sketches to distinguish between geostationary, geosynchronous and low Earth orbits.

Applying

4 Explain why the acceleration of a satellite with an orbital radius of 42 000 km is less than that for a satellite with 7000 km orbital radius.

5 Titan, a moon of Saturn, has an orbital radius of 1.22×10^6 km. It takes Titan 15 days and 22 hours to revolve around Saturn. Find the mass of Saturn.

6 A satellite is orbiting the Sun with a radius of 1.5×10^{10} m. What is its orbital speed, given the mass of the Sun is 2.0×10^{30} kg?

7 Find the orbital period of an asteroid around the Sun, given that its mean orbital radius is twice Earth's orbital radius. Do not use the mass of the Sun in this calculation.

Analysing

8 The law of parsimony states that the simpler explanation is to be preferred. How does this law apply to the explanations of planetary motion provided by Ptolemy, Kepler and Newton?

Reflecting

9 Create a series of annotated diagrams to illustrate what you have learnt about the orbits of planets and satellites.

CHAPTER SUMMARY

- G is the gravitational constant: $G = 6.67 \times 10^{-11}\,\text{N}\,\text{m}^2\,\text{kg}^{-2}$
- Gravity is an action-at-a-distance force. Action-at-a-distance forces can be modelled as being mediated by a field.
- All objects with mass are surrounded by a gravitational field.
- All objects with mass experience a force in a gravitational field. This force is called the weight force.
- The gravitational field at a particular point is defined as the force per unit mass experienced by an object at that point:
  $$g = \frac{F}{m} = \frac{GM}{r^2}$$
- The gravitational field has units of $\text{N}\,\text{kg}^{-1}$ or $\text{m}\,\text{s}^{-2}$
- In the absence of other forces, the gravitational field is the acceleration of an object in the field.
- Close to the surface of Earth the gravitational field is approximately constant and has the value $9.8\,\text{m}\,\text{s}^{-2}$.
- Gravitational potential energy is stored in any system of masses because the gravitational force acts on the masses.
- The zero of gravitational potential energy is, by convention, when all objects in the system are infinitely separated.
- In field theory, we model the energy as being stored in the field.
- When an object moves in a gravitational field, work is done on or by the field.
- When work is done by the field, the potential energy of the system decreases. When work is done on the field, the potential energy of the system increases.

- Early models of the universe had the heavenly bodies fixed on spheres in space, including crystalline spheres and later epicycles, to explain the observed motion of the planets.
- Galileo observed that falling bodies had constant horizontal motion and accelerated vertical motion.
- Tycho Brahe's meticulous observations enabled Kepler to propose his three laws of planetary motion:
  - Kepler's first law: Planets orbit the Sun in elliptical orbits
  - Kepler's second law: Planets orbits sweep out equal areas in equal time
  - Kepler's third law – the law of periods: $T^2 \propto r^3$ for planets orbiting the same central mass, and:

$$\frac{T^2}{r^3} = \frac{4\pi^2}{GM}$$

- Kepler's laws of planetary motion were derived empirically but can be explained as resulting from Newton's law of universal gravitation and conservation of angular momentum.
- Newton's law of universal gravitation gives the gravitational force that any mass, M, exerts on any other mass, m:

$$F = G\frac{Mm}{r^2}$$

- A satellite's motion can be modelled as uniform circular motion.
- The gravitational force acting on a satellite is the net (centripetal) force.
- The orbital speed of a satellite is given by $v = \sqrt{\frac{GM}{r}}$.

CHAPTER GLOSSARY

apparent weight subjective experience of weight, which is generally measured as the normal force exerted on an object that is not in equilibrium

apparent weightlessness the experience of having no normal force exerted on you; this occurs during free fall

concentric having the same centre

centre of mass the average position of the mass in an object or group of objects. It is the point at which the gravitational force can be modelled as acting when the object is in a gravitational field

epicycles circles on circles, as used to describe the orbits of the planets

field the means by which action-at-a-distance forces are exerted

field theory the theory that describes forces as being mediated by fields and potential energy as being stored in fields

free fall falling with the acceleration g, the local gravitational field strength

geostationary orbit of period 24 h approx. directly above a single point on Earth's equator

geosynchronous orbit of period 24 h approx. above a great circle on Earth

gravitational potential energy the potential energy associated with the interaction of objects via the gravitational force. The potential energy is stored in the gravitational field

gravitas Aristotelian idea about the 'heaviness' of objects made of earth that allowed them to fall in straight lines towards Earth

gravitational field the field that mediates the gravitational force between all objects with mass; the field surrounding all objects with mass, $g = \frac{GM}{r^2}$

inverse-square law describes a relationship in which the dependent variable is proportional to the square of the inverse of the independent variable

low Earth orbit an orbit between 250 km and 1000 km above Earth

negligible any value or variation in a value that is too small to be taken into account

potential energy energy stored in a system due to the interaction of components in the system via forces; energy stored in a field. Potential energy gives a system the ability to do work

weight the gravitational force that acts on an object, $w = mg = \frac{GmM}{r^2}$

work energy transferred due to the action of a force, $W = Fs$

CHAPTER REVIEW QUESTIONS

Remembering

1 Write down Kepler's third law.

2 Why was Tycho Brahe's work so important for Kepler?

3 Write down the gravitational field strength according to Newton's law of universal gravitation.

4 a What are epicycles?
   b Why were they needed in early planetary models?

5 Why is the net force acting on a satellite considered to be a centripetal force?

6 Define these terms.
   a Weight
   b Apparent weight
   c Weightlessness

Understanding

7 a How are work and gravitational force connected?
   b Explain why no work is being done when you hold a brick above your head.

8 How is gravitational potential energy defined? Why must a zero point be stated before the potential energy of a system can be calculated?

9 Under what circumstances is your weight equal to the normal force acting on you? If you carry a heavy school bag, is your weight equal to the normal force exerted on you by the ground? Draw a diagram showing all the forces acting on you, to explain your answer.

10 Why is an orbiting satellite considered to be in free-fall?

11 What happens to the gravitational force between two objects when they are moved 10 times further apart?

Applying

12 What minimum work must be done to move a 1500 kg car up a 300 m high hill? What work is done by the gravitational field as the car is moved up the hill?

13 What is the orbital speed of the moon of a planet, given that the planet has a mass of 4.5×10^{24} kg and the Moon's mean orbital radius is 3.8×10^{8} m?

14 A 5.0 kg rock is dropped from 10 m above the surface of Ganymede, a moon of Jupiter. Ganymede has a mass of 1.5×10^{23} kg and radius of 2600 km.
   a What is the gravitational field near Ganymede?
   b What is the acceleration of the rock?
   c How much gravitational potential energy was transformed as the rock fell to the surface?
   d What was the maximum amount of kinetic energy gained by the rock?

15 What is the centripetal force acting on Earth by the Sun? The mass of the Sun is 2.0×10^{30} kg, the mass of Earth is 5.97×10^{24} kg and the distance between the two is 1.5×10^{11} km.

16 What is the orbital period in Earth years for an asteroid that is orbiting the Sun with a mean orbital radius that is three times the orbital radius of Earth?

17 What is the mass of the central body being orbited by a star when the star's orbital period is 8.0 hours and its distance from the central body is 4.5×10^{10} m?

18 A geostationary satellite orbits Earth once every 24 hours. (Mass of Earth = 5.97×10^{24} kg; radius of Earth = 6.38×10^{6} m).
   a What is the value of the Kepler ratio $\frac{R^3}{T^2}$?
   b How high above Earth's surface is it placed?

Analysing

19 To what speed must a satellite be propelled if it is to maintain an orbit with a radius of 10 000 km around Earth?

20 Explain how Kepler's third law can be used by astronomers to deduce the mass of a star.

21 When watching a satellite launch, the launch vehicle is seen to ascend at an angle rather than vertically. Explain this observation.

22 A space station is made in the shape of a doughnut, 160 m across. It rotates about an axis so that the workers experience down to be towards the outer wall.
   a Sketch the space station and indicate the force that simulates the gravitational force applied to a worker standing upright.
   b At what speed must the space station rotate in order to simulate Earth's gravitational field?

Reflecting

23 Models are simplifications of real systems used to help us understand the universe. Draw a concept map summarising gravitational field theory.

24 How has your understanding of forces and energies changed while studying this chapter? Write a short summary and compare your summary with those of other students.

CHAPTER 3
ELECTRIC FIELDS

Shutterstock.com/Igor Zh

By the end of this chapter you will have covered the following material.

Science Understanding

- Electrostatically charged objects exert a force upon one another; the magnitude of this force can be calculated using Coulomb's Law (ACSPH102)
- Point charges and charged objects produce an electric field in the space that surrounds them; field theory attributes the electrostatic force on a point charge or charged body to the presence of an electric field (ACSPH103)
- A positively charged body placed in an electric field will experience a force in the direction of the field; the strength of the electric field is defined as the force per unit charge (ACSPH104)
- When a charged body moves or is moved from one point to another in an electric field and its potential energy changes, work is done on or by the field (ACSPH105)

Introduction

There are four fundamental forces. These are the strong force, the weak force, the gravitational force and the **electromagnetic force**. The gravitational force was described in the previous chapter. The strong and weak forces will be described in Chapters 9 and 10. The electromagnetic force is a combination of the electrostatic and magnetic forces. This chapter is about the **electrostatic forces** that are due to charges.

You have seen that there is a gravitational field around all objects with mass. It is this gravitational field that allows objects with mass to exert forces on each other at a distance. This field, and hence the force, depends on the mass of the objects and on the distance between them. It increases linearly with the masses and decreases with the square of the distance between the objects. The gravitational force is always attractive, acting to pull two objects with mass closer together. The **electric field** is in many ways similar to the gravitational field. The electric field is created by charged particles.

You have already seen some of the applications of electricity in *Nelson Physics Units 1 & 2 for the Australian Curriculum*, Chapters 5 and 6. You have probably already used electricity at least a dozen times today. An understanding of electrostatics is the first step in understanding how electricity is produced, transferred and used in the many devices that we take for granted. We shall look at some of these applications in the next chapter.

If the objects are moving very fast we need to take into account relativistic effects and use a relativistic field model. Relativity is described in Chapter 6. If the objects are very tiny we may need to use a quantum mechanical model. Quantum mechanics is discussed in Chapters 7 and 8.

For almost any situation there is a choice of models which can be used. In general, we try to use the simplest one that captures all the features of the system and which has good predictive and explanatory power.

The electric field

The first model that we shall be using in this chapter is the **electrostatic field model**. This model was developed from Michael Faraday's idea of 'lines of force'. These lines of force enable an object to exert a force on a second object some distance away. Faraday described both electric and magnetic fields. His field model was refined and extended by others, particularly James Clerk Maxwell.

For the electrostatic field model we assume that the charges are not moving, and we ignore any quantum or relativistic effects. This model is extremely good at predicting the behaviour of most interacting charged objects. You will see that this model is very similar to Newton's model of universal gravitation.

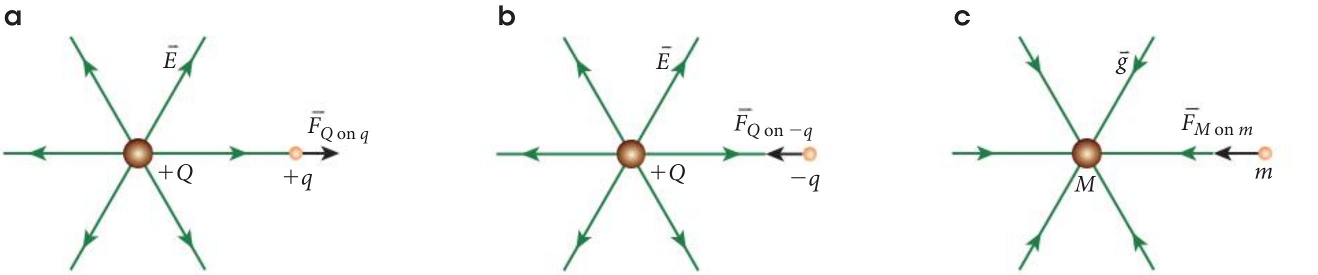
You can do experiments to show that a charged object exerts a force on another charged object some distance away. For example, if you rub a balloon on your hair and then hold it close to your head you will see or feel your hair being pulled towards it. The charge on the balloon attracts the charge on your hair, even though the two are not in contact. In the same way, the mass of Earth attracts the mass of any other object without having to touch it.

All objects with electric charge create an electric field around themselves. This field depends on the size of the charge and the distance from the charged object. It increases with the size of the charge and decreases proportionally with the square of the distance. Unlike the gravitational field, the electric field can exert an attractive or repulsive force (pull or push) because charge comes in both positive and negative types. You already know that like charges repel each other and unlike charges attract. Hence a positive charge pulls a negative charge towards itself and pushes another positive charge away.

The electric field at a point is defined as the force per unit charge that acts on a small positive test charge at that point.

$$\vec{E} = \frac{\vec{F}_{\mathrm{E}}}{q}$$

$\vec{E}$ is the electric field and $\vec{F}_{\mathrm{E}}$ is the force acting on an object with charge q. Force has units of newtons (N) and charge has units of **coulombs** (C), so electric field has units of $\mathrm{N\,C^{-1}}$.



▲**Figure 3.1**
a) A large positive charge repels a small positive charge. b) A large positive charge attracts a small negative charge. c) A large mass attracts a small mass.

Force is a vector – it has both magnitude and direction – so field is also a vector. Its direction at any point is the same as that experienced by a small positive test charge at that point. The force on a negative charge is in the opposite direction to the field. In Figure 3.1 the small positive test charge experiences a force due to the electric field created by the large charge. The field is the magnitude and direction of this force, divided by the charge on the small test charge.

This is similar to the definition of the gravitational field, which is force per unit mass acting on a small test mass. However with electric fields we have to think about direction a bit more carefully.

The units of electric field are N C^{-1}. 1 N = 1 kg m s^{-2} = 1 J m^{-1}, and 1 V = 1 J C^{-1}, so the units for electric field can also be written as V m^{-1}.

An electric field exerts a force in the direction of the field on a positive charge, and in the opposite direction on a negative charge.

Table 3.1 Some electric field strengths

Electric field due to...	Approximate field strength, N C^{-1}
Hairdryer, 20 cm away	4
Thunderstorm	50, upwards
Earth's fair weather field	100, downwards
High voltage overhead power lines, 30 m away	10 to 1000
Old electric blanket, 10 cm away	2000

In Nelson Physics Units 1 & 2 for the Australian Curriculum you studied electric circuits. The potential difference supplied by the battery in a circuit creates an electric field that applies a force to the free electrons, creating a current.

WORKED EXAMPLE 3.1

A battery is connected across a piece of copper wire giving an electric field of 3.0 N C^{-1} in the wire. What is the force on an electron in this wire? (5 marks)

Answer	Logic	
$\vec{E} = \dfrac{\vec{F}_E}{q}$	Use the definition of field.	1 mark
$\vec{F}_E = \vec{E}q$	Rearrange to find the force.	1 mark
$\vec{F}_E = (3.0\,\text{N C}^{-1})(-1.6 \times 10^{-19}\,\text{C})$	Substitute known values, remembering to include units.	1 mark
$\vec{F}_E = -4.8 \times 10^{-19}\,\text{N}$	Do the calculation.	1 mark
The force is in the direction opposite that of the field.	State the direction of the force.	1 mark

Try this yourself

What is the force on a copper nucleus in this wire? (6 marks)

In *Nelson Physics Units 1 & 2 for the Australian Curriculum* you studied circuits and saw that when a battery is connected across a conductor it causes a current to flow. The battery creates an electric field in the wire because one terminal is positive and the other is negative. This polarity is due to charge separation in the battery. This field in the conductor exerts a force on all the charged particles in the wire. The conduction electrons, which are free to move, are accelerated by this electric field. The movement of these electrons along the wire due to their drift velocity is what we observe as a current. Electric fields can be set up whenever there is a means of separating charges.

Recall from your studies of radiation and nuclear physics in *Nelson Physics Units 1 & 2 for the Australian Curriculum* that electrons have a charge of -1.6×10^{-19} C. This is called the **electron charge** or the elementary charge, *e*, because all particles have an integer multiple of this charge, $q = ne$ where $n = \pm 0, 1, 2, \ldots$ Protons and positrons (β^+ particles) have charge $+1.6 \times 10^{-19}$ C. Ions that have extra electrons have a charge equal to the number of electrons times the electron charge. Ions with missing electrons have a charge equal to some multiple of $+1.6 \times 10^{19}$ C. For example, an O^{2-} ion has charge $2 \times -1.6 \times 10^{-19}\text{ C} = -3.2 \times 10^{-19}$ C. An Fe^{3+} ion has charge $3 \times 1.6 \times 10^{-19}\text{ C} = 4.8 \times 10^{-19}$ C.

WOW

Electroreception in platypuses

Platypuses can detect electric fields with special cells called electroreceptors on their bills. This sense is called *electroreception*. Platypuses and echidnas are the only mammals that have this ability. The platypus's electroreception was first demonstrated in an experiment in 1986, in which platypuses found and attacked hidden batteries. Previously, their hunting ability with their nostrils, ears and eyes closed had been a mystery. Their electroreceptors can detect electric fields as small as 0.002 N C^{-1}. Platypuses combine information from electroreceptors and pressure sensors on their bills to determine the direction of their prey and their distance from it.

Nature Picture Library/Georgette Douwma

▲**Figure 3.2** A platypus has very sensitive electroreceptors in its bill to help it hunt for food.

To understand how a charged object behaves in a field we need to use another model: Newton's laws. Newton's first law tells us that an object experiencing a net force will accelerate. Newton's second law quantifies the acceleration. Newton's second law says that $a = \frac{\vec{F}}{m}$. Using the definition of electric field $\vec{E} = \frac{\vec{F}}{q}$, we can see that $\vec{F} = \vec{E}q$, so that:

$$\vec{a} = \frac{\vec{E}q}{m}$$

WORKED EXAMPLE 3.2

A Ca^{2+} ion with a mass of 6.7×10^{-26} kg experiences an acceleration of 4.8×10^{13} m s^{-2} as it moves through a channel in a cell membrane. What is the electric field in the membrane? (6 marks)

Answer	Logic	
$\vec{a} = \frac{\vec{E}q}{m}$	Relate the acceleration to the field.	1 mark
$\vec{E} = \frac{\vec{a}m}{q}$	Rearrange for the electric field.	1 mark

▶

$a = 4.8 \times 10^{13}\,\mathrm{m\,s^{-2}}$ $m = 6.7 \times 10^{-26}\,\mathrm{kg}$ $q = 2 \times 1.6 \times 10^{-19}\,\mathrm{C} = 3.2 \times 10^{-19}\,\mathrm{C}$	Identify numerical values for all parameters.	1 mark
$E = \dfrac{(4.8 \times 10^{13}\ \mathrm{m\,s^{-2}})(6.7 \times 10^{-26}\ \mathrm{kg})}{3.2 \times 10^{-19}\ \mathrm{C}}$	Substitute the correct values with units.	1 mark
$E = 1.0 \times 10^{7}\,\mathrm{m\,kg\,s^{-2}C} = 1.0 \times 10^{7}\,\mathrm{N\,C^{-1}}$.	Calculate final answer.	1 mark
The field is in the direction of the acceleration.		1 mark

Try this yourself

What is the acceleration of a sodium ion, Na^+, in this channel? (5 marks)

WOW

Electric fields in cell membranes

The cells in your body have an electric field in their membranes. This field is maintained by pumps in the cell wall which move ions through the membrane. The outside of the cell is positive compared to the inside, so the electric field points inwards and is approximately constant in the cell membrane. The fields are about 10 000 000 N C^{-1}! The main ions involved are K^+, Na^+ and Cl^+. Large molecular anions (A^{4-}) are also involved in binding the positive ions. Other ions, such as calcium, Ca^{2+}, are important in nerve cells. When a nerve cell sends a message, the electric field in its membrane collapses and then acts in the opposite direction as ions flow into and out of the cell.

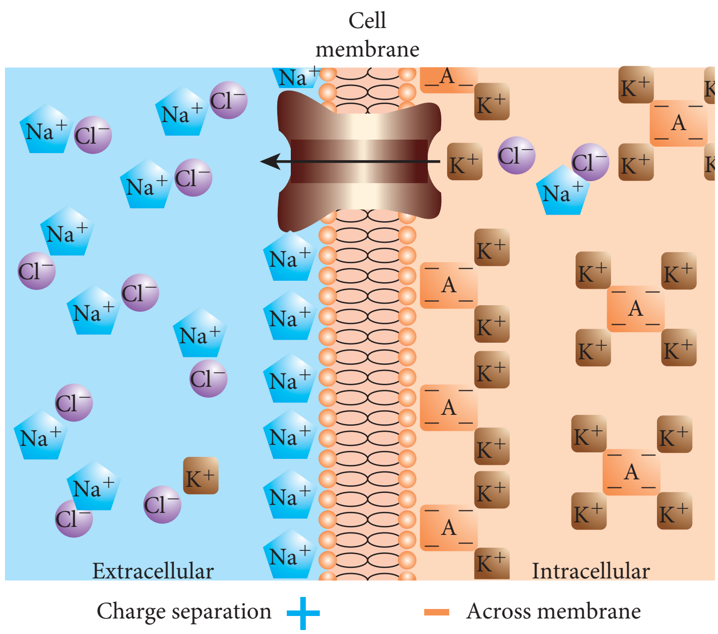


Figure 3.3 A cell maintains a large electric field across its membrane by pumping ions into and out of the cell.

Electric field lines

Fields can be represented by field diagrams. An electric field diagram uses lines with arrows to show the direction of the force on a small positively charged test particle.

To draw a field diagram, start by considering the force on a test charge at various points. Let's start with a single positive charge and think about what happens when we put a small positive test charge close to it. We know that like charges repel, so the test charge will be accelerated away from the positive charge, as shown in Figure 3.4(a). Anywhere close to the positive charge the direction of the electric field is radially away from the charge. If we join up the arrows in Figure 3.4(b) we get field lines, as shown in Figure 3.4(c).

Figure 3.4(c) shows a field line diagram for a single positive point charge. If the point charge was negative instead of positive, the field lines would all point inwards. A small positive test charge would be attracted to a negative charge (see Figure 3.5). This looks just like the gravitational field around a spherical object such as Earth, as you saw in Chapter 2.

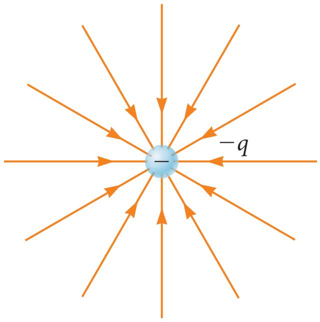
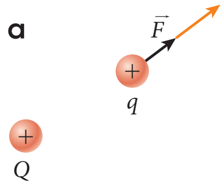


Figure 3.5 ▲
Field line diagram for a negative point charge



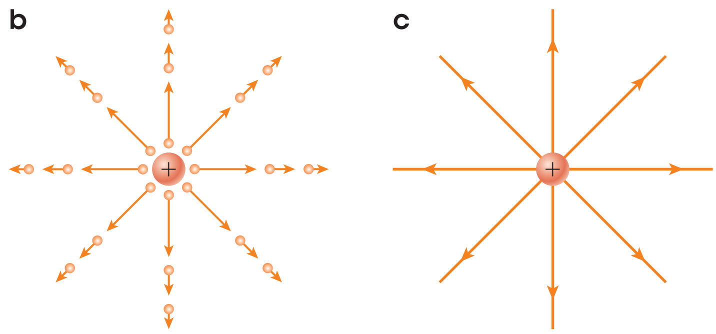


Figure 3.4 ▲
Constructing a field diagram. Start by drawing arrows to show the acceleration at various points, then join them up to form field lines.

In Nelson Physics Units 1 & 2 for the Australian Curriculum you studied forces. Recall that forces need to be added as vectors. As fields are a force per unit charge (or mass), they also need to be added as vectors.

Electric field lines have the following characteristics:

- **They point in the direction of the force acting on a positively charged particle due to the field.**
- **They never cross.**
- **They begin on positive charges and end on negative charges.**
- **The field strength is proportional to the density of field lines.**

When you draw a field line diagram there is an infinite number of possible field lines that you can draw. However, you only have finite time, so choose a sensible number! In general, make sure you draw enough lines so that you can see what the field looks like around the charge or charges. Typically *at least* eight field lines are needed. They should be evenly spaced around a point charge. Make the ratio of field lines coming out from or entering any charge proportional to the magnitude of the charge. This will ensure your field lines have the properties listed above.

WORKED EXAMPLE 3.3

Draw a field line diagram for a pair of equal positive charges. (5 marks)

Answer

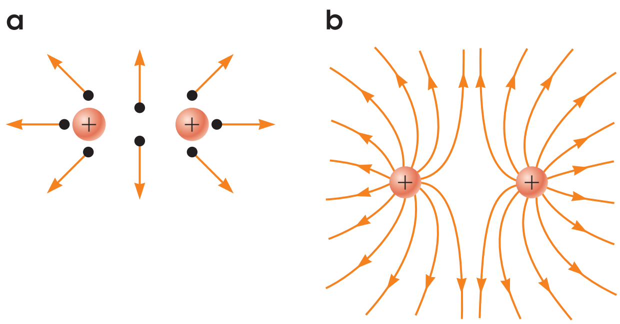


Figure 3.6 ▲
Construction of a field diagram for a pair of positive charges. a) Show the force acting on a small positive test charge at points around the charges.
b) Connect the lines of force to show field lines.

Logic

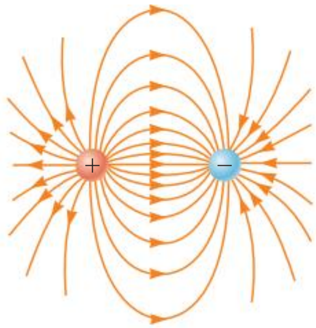
Think about putting a small positively charged test object at points close to the charges as shown in Figure 3.6(a).	1 mark
Draw arrows showing the forces due to the two charges.	1 mark
The vector sum of these two forces gives the total force, and the direction is the same as the direction of the field at that point.	2 marks
Do this at lots of points (Figure 3.6(a)) and then join up your arrows to form field lines as in Figure 3.6(b).	1 mark

Try this yourself

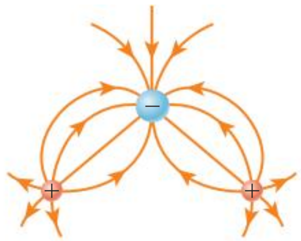
Draw a field diagram for a pair of negative charges. (5 marks)

A common configuration of charges is a **dipole**. This is a combination of a positively charged object and a negatively charged object close together. Figure 3.7 shows the electric field in the region around a dipole.

You have probably met the idea of **polar** molecules in chemistry. Polar molecules such as water (Figure 3.8) are loosely bound by **van der Waals forces** because of the attraction between the negative part of one molecule and the positive part of a nearby molecule. This is also called hydrogen bonding because it commonly happens with hydrogen.



▲Figure 3.7
Electric field around a dipole



▲Figure 3.8
Electric field due to a polar water molecule

A charged particle *starting at rest* in an electric field will accelerate along a path parallel to the field lines. It will accelerate in the direction of the field lines if it has a positive charge. If it has a negative charge it will accelerate in the opposite direction. If a charged particle has some initial velocity in an electric field, then the field lines give the direction of acceleration, but they *do not* show the path of the particle. You have already seen this as projectile motion in a gravitational field. If an object is released from rest its path is that of a gravitational field line down to the surface of Earth. If it has some initial velocity then the gravitational field lines give the direction of acceleration, but *do not* give the path of the object.

CHARGES AND FIELDS

This simulation lets you move charges around and see their electric field. Try it with two positive charges, two negative charges and a positive and a negative charge.

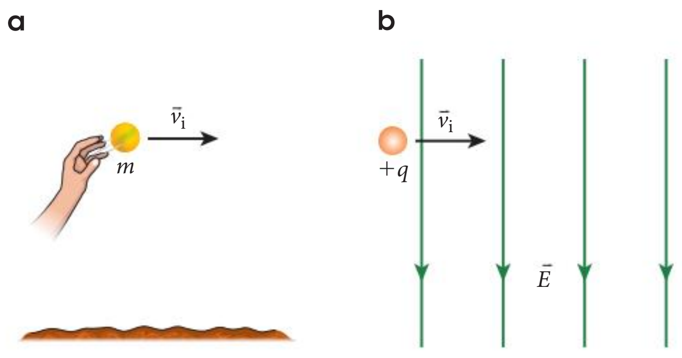
WORKED EXAMPLE 3.4

1 Figure 3.9(a) shows a ball thrown horizontally close to the surface of Earth. (5 marks)

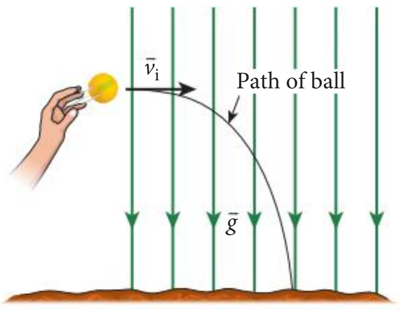
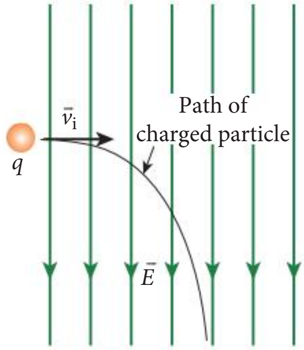
- a Draw a diagram showing the gravitational field lines.
- b Draw the path taken by the ball.
- c Describe in words the shape of the path.

2 Figure 3.9(b) shows a positively charged particle entering a uniform electric field, travelling perpendicular to the field. (3 marks)

- a Draw a diagram showing the path taken by the particle.
- b Describe in words the shape of the path.



◀Figure 3.9 a) A ball thrown horizontally close to the surface of Earth. b) A positively charged particle enters a uniform electric field and travels in a direction perpendicular to the field.

Answers	Logic	
1 a  **Figure 3.10 ▲** Gravitational field lines and path of ball	The field lines are directed straight down and are parallel because the field is constant.	2 marks
b See Figure 3.10.	The ball is pulled down by the gravitational force, but the horizontal velocity is unchanged.	2 marks
c The ball follows a parabolic path.	Give a correct description of path.	1 mark
2 a  **Figure 3.11▲** Electric field lines and path of particle	The particle is pulled in the direction of the electric field, but the horizontal velocity is unchanged.	2 marks
b The particle follows a parabolic path. This is the same shape as the path followed by the ball in the gravitational field.		1 mark

Try this yourself

Repeat question 2 for a negatively charged particle in a uniform electric field. (4 marks)

In Nelson Physics Units 1 & 2 for the Australian Curriculum you studied the kinematics equations for constant acceleration. The acceleration due to a uniform field is constant. This is true for both the gravitational field close to Earth and any uniform electric field. Hence we use the same equations in both cases.

Just as in the gravitational case, we use kinematics to calculate the trajectory. The acceleration depends on the electric field. In the case of a uniform field, as described next, the acceleration is constant.

Uniform electric fields

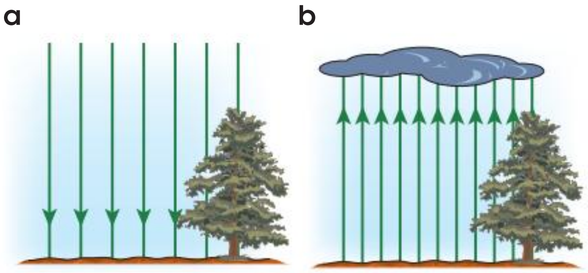
A **uniform field** is one that has the same magnitude and direction at all points.

A uniform electric field exists close to any large, flat, uniform distribution of charge. If we look at the field very close to the surface of a large charged sphere, such as the dome of a van de Graaf generator, the field is approximately uniform. You have seen and used this useful approximation many times before for the gravitational field of Earth. Every time you write $F = mg$ you are making the approximation that Earth is flat. Most of the time this is perfectly reasonable. Close to the surface of Earth the gravitational field is effectively constant.

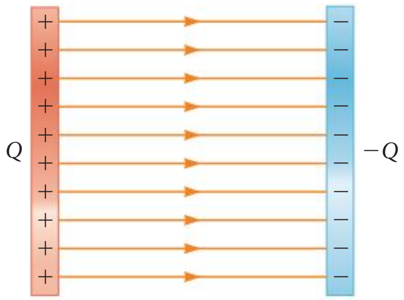
Earth also has an electric field; you will be familiar with the effects of this field in stormy weather, such as that shown in Figure 3.12 The size of the field varies with weather conditions (Figure 3.13). In sunny weather Earth has an electric field of about $100\,\mathrm{N\,C^{-1}}$, pointing downwards. It is larger in fog or snow, and reverses direction in heavy rain. The field varies over the surface of Earth, but can be treated as uniform over distances of a few hundred metres or more. The force due to Earth's electric field on a small charged particle such as an electron is many orders of magnitude greater than that due to the gravitational field. For large, neutral objects such as humans, the gravitational force is much larger.

On a smaller scale, a uniform electric field can be created by two charged parallel plates. This arrangement of charged plates is called a capacitor. Capacitors are a common circuit element. Figure 3.14 shows the electric field created by a pair of parallel plates.

▲**Figure 3.12** Lightning strikes a tree when the electric field between the cloud and the ground becomes too large.



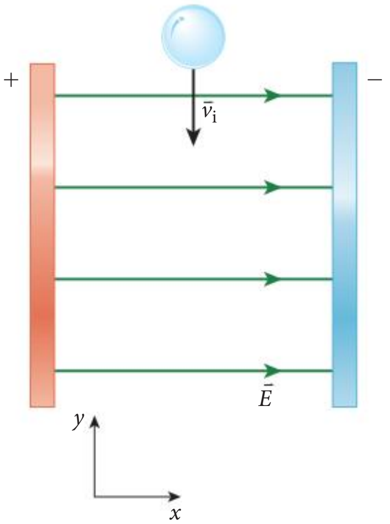
▲**Figure 3.13**
Earth's electric field in a) fair weather and b) stormy weather



▲**Figure 3.14**
The electric field between the plates of a parallel plate capacitor is uniform.

WORKED EXAMPLE 3.5

An approximately uniform electric field is created by charging two large parallel metal plates. A charged soap bubble drifts between the plates, with an initial velocity v_i parallel to the plates, as shown in Figure 3.15. The bubble has mass m and charge $+q$, the magnitude of the field between the plates is E. Assume that the gravitational force on the bubble is negligible in comparison to the electrostatic force.



◀**Figure 3.15**
A bubble enters the uniform electric field between a pair of charged plates.

1 Find the acceleration of the bubble. (4 marks)

2 Write an expression for the position of the bubble as a function of time once it enters the field between the plates. (4 marks)

Answers	Logic	
1 $a_x = \frac{F}{m} = \frac{E_x q}{m}$	The force acts only in the direction of the field. It is zero in a direction perpendicular to the field. This is analogous to projectile motion in a gravitational field, as in the previous example.	2 marks
	Relate acceleration to data given.	
$a_y = \frac{E_y q}{m} = 0$	Noting that the field is entirely in the *x* direction.	2 marks
2 $x = v_{x,i}t + \frac{1}{2}a_x t^2$ $= 0 + \frac{E_x q t^2}{2m}$ $= \frac{E_x q t^2}{2m}$	As the situation is analogous to projectile motion, apply the kinematics equations.	2 marks

$$y = v_{y,i}t + \frac{1}{2}a_y t^2 = v_{y,i}t + 0 = v_{y,i}t$$

We can see from the equations that the particle curves towards the direction of the field lines, but always has some component of its velocity perpendicular to the field – it does *not* follow the field lines. Hence the path followed is the same as that shown in Figure 3.11 in the previous example. 2 marks

Try this yourself

Repeat the questions above for a soap bubble with charge –2*q* and the same mass. (6 marks)

Fields apply forces; hence they are also able to apply a torque and cause rotation of an object. You studied torques Chapter 1.

The model that we have applied here is non-relativistic. This is reasonable for a bubble, but an object with very small mass such as an electron could quickly be accelerated to relativistic speeds. For any speed greater than about 0.1c, we should really be using the relativistic equations, given in Chapter 6.

A dipole, such as that shown in Figure 3.7, experiences no net force in a uniform electric field. This is because each charged part of the dipole experiences an equal but opposite force. The two forces, one on the positive part and one on the negative part, add to give a total force of zero. This does not mean that a dipole is not affected by the field. Each force produces a torque on the dipole, causing it to rotate and line up with the field. A pair of forces like this are called a 'couple', and the torque produced is called a 'couple moment'. As you saw in Chapter 1, a gravitational field can also exert a torque and cause rotation. However, the gravitational field only exerts a force in one direction, so it always applies a net force as well as a torque.

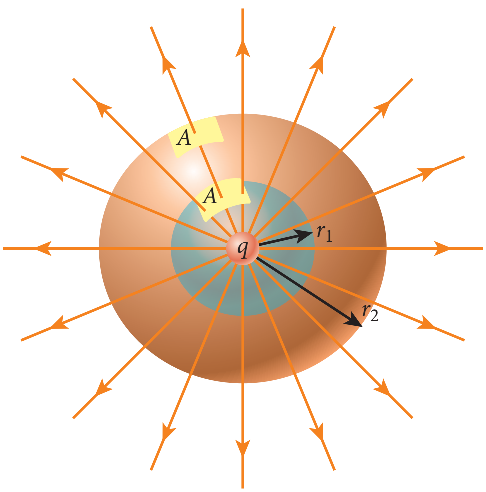


Figure 3.16▲
Field lines due to a point charge. The number of field lines passing through the sphere with radius r_1 is the same as the number passing through the sphere with radius r_2. However, the number passing through any small area, *A*, which is a measure of the density of field lines and hence field strength, is greater for the smaller sphere.

Electric field due to a point charge

Point charges occur commonly in nature. Fundamental particles such as protons and electrons can be modelled as point charges.

We saw in the previous section that up close, mass and charge distributions look flat and give uniform fields. In contrast, any finite mass or charge distribution seen from a great distance away can be modelled as a point mass or charge. Earth seen from a great distance away looks like a point mass. Newton's law of universal gravitation treats it as such. The dome of a van de Graaff generator looks like a point charge from a great distance away. Hence it is very useful to us to know what the field due to a point charge is.

We know from experiment that the bigger the charge, the bigger the field. So we expect field to be proportional to charge. We also expect the field to have spherical symmetry, as there is no preferred direction around a point. We can see this from our field line diagrams.

We also know from experiments that the field becomes weaker at greater distances from the charge. A charged balloon does not pull as hard at your hair if you move it further away. We need to know how the field strength varies with distance. We can use field line diagrams to work it out.

The field lines become less dense the further away they are from a point charge (Figure 3.16). If you draw concentric circles around a point charge, the number of field lines crossing each circle is constant. However the density decreases with the radius of the circle. In three dimensions, the field lines radiate out in all directions. If you imagine concentric spheres around a point charge, the *number*, n, of field lines crossing each sphere is the same. However the *density* of lines decreases with the square of the radius because the area of each sphere is proportional to r^2. Density is $\frac{\text{number}}{\text{area}}$, and area of a sphere $= 4\pi r^2$. So density $= \frac{n}{4\pi r^2}$.

Hence, the field strength, which is proportional to the field line density, decreases with distance squared from the source. This $\frac{1}{r^2}$ form is common to *all* point sources, whether they are sources of electric field, gravitational field, light or sound waves.

Recall from Chapter 11 of Nelson Physics Units 1 & 2 for the Australian Curriculum that light intensity decreases with $\frac{1}{r^2}$ for a point source of light.

The electric field due to a charged particle or a spherical charged object is:

$$E = \left(\frac{1}{4\pi\varepsilon_0}\right)\frac{q}{r^2}$$

where $\varepsilon_0 = 8.9 \times 10^{-12}\,\text{C}^2\,\text{N}^{-1}\,\text{m}^{-2}$ so the constant $\frac{1}{4\pi\varepsilon_0} = 9 \times 10^9\,\text{N}\,\text{m}^2\,\text{C}^{-2}$.

This expression meets the requirements described above.

The constant $\frac{1}{4\pi\varepsilon_0}$ is sometimes known as the **Coulomb constant**, after Charles-Augustin de Coulomb, and given the symbol k_e. The constant ε_0 is called the **permittivity of free space**. It tells us how the field varies in a vacuum, as well as giving us the correct units for field. This is an important physical constant. The speed of light depends on this constant, as we shall see in Chapter 5.

Figure 3.17 shows a plots of field strength as a function of distance for a positive and a negative charge.

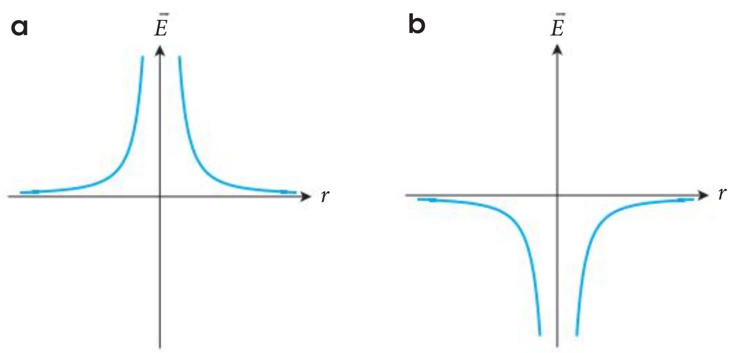


Figure 3.17 Plots of $E(r)$ for a) $+q$ and b) $-q$

You already know that electrons and protons have electric charge. We can now calculate the field around these particles.

WORKED EXAMPLE 3.6

Calculate the electric field due to a proton at a distance of 8.0×10^{-11} m. This is the average orbital radius of the electron in a hydrogen atom. This distance is four orders of magnitude larger than the size of the proton, which is about 10^{-15} m, so we can treat the proton as a point charge. (4 marks)

Answer	Logic	
$E = \left(\frac{1}{4\pi\varepsilon_0}\right)\frac{q}{r^2}$	Relate field to data given.	1 mark
$q = 1.6 \times 10^{-19}$ C	Find values needed but not given.	1 mark

$$E = (9.0 \times 10^{9}\ \text{N m}^2\ \text{C}^{-1}) \times \frac{(1.6 \times 10^{-19}\ \text{C})}{(8.0 \times 10^{-11}\ \text{m})^2}$$ Substitute values, including units. 1 mark

$E = 2.3 \times 10^{11}\ \text{N C}^{-1}$ Calculate the final value. 1 mark

Try this yourself

Calculate the electric field due to a proton at distances of 1×10^{-10} m (a typical distance between atoms), 1 cm and 1 m. (3 marks)

QUESTION SET 3.1

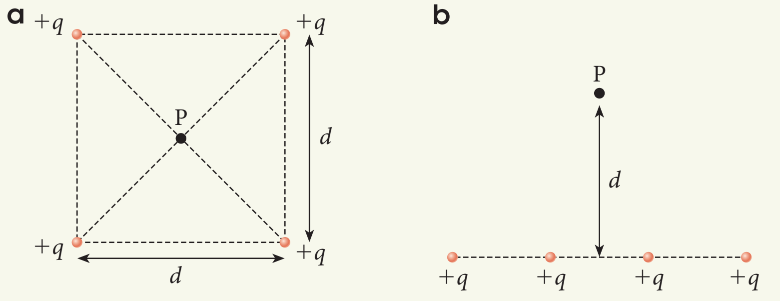
Remembering

1 Write down the relationship between the electrostatic force and electric field.

2 When is an electric field directed towards a charge, and when is it directed away from a charge?

Understanding

3 Consider the two arrangements of charges shown below in Figure 3.18. Compare the electric field at point P in the two cases. Is it the same in each case? If not, which is larger and why?



◀Figure 3.18

4 Draw a table comparing the properties of electric fields and gravitational fields. List similarities and differences.

Applying

5 Calculate the electric field at a distance of 10 cm away from a point charge of +5.0 mC.

6 Calculate the acceleration of an electron in a constant electric field of 95 N C^{-1}, directed upwards. In which direction does the electron accelerate?

7 A pollen grain of mass 0.001 g and charge 0.05 nC starts at rest in an electric field of 80 N C^{-1}. Assuming no other forces, how far does the pollen grain move in 1 minute?

Analysing

8 Draw a field diagram for the:

a electric field around two nearby positive charges.

b electric field around two nearby negative charges.

c gravitational field around the Moon and Earth.

Coulomb's law

You already know that charges exert forces on each other. Now that we know what the field due to a point charge is, we can quantify those forces that the charge exerts on a second charged particle.

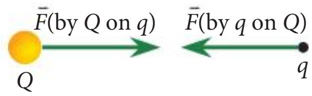
If our first point charge has charge Q then it creates an electric field:

$$E = \left(\frac{1}{4\pi\varepsilon_0}\right)\frac{Q}{r^2}$$

Using the definition of electric field, $\vec{E} = \frac{\vec{F}}{q}$, the force exerted on a second point charge with charge q is:

$$F = Eq = \left(\frac{1}{4\pi\varepsilon_0}\right)\frac{Qq}{r^2}$$

This is known as **Coulomb's law**. It describes the force that a point charge Q exerts on a second point charge q when they are separated by a distance r. The force will be directed towards or away from charge Q. Remember that forces are vectors. The form of Coulomb's law given here only gives the magnitude of the force. You should always draw a diagram to show the direction of the force. Remember: like charges repel and unlike charges attract.



▲**Figure 3.19** Two charges exert forces on each other.

If we now write an equation for the force that charge q exerts on charge Q, we get exactly the same expression. This should not come as a surprise. From Newton's third law, whatever force q exerts on Q, Q must exert an equal and opposite force on q. According to Newton's third law:

F(by q on Q) and F(by Q on q):

- are equal in magnitude
- are opposite in direction
- have the same fundamental nature

AND

- each force acts on a different object.

This is shown in Figure 3.19.

The gravitational and electric forces depend on the product of the property of the objects that creates the field, and decreases with the square of the distance between the objects.

You will study Bohr's planetary model in the chapter on quantum mechanics, as well as the more modern quantum model of the atom.

Coulomb's law states that the force that one charged particle exerts on a second charged particle is given by:

$$F = \left(\frac{1}{4\pi\varepsilon_0}\right)\frac{Qq}{r^2}$$

where Q and q are the charges on the particles, r is the distance between them and ε_o is the permittivity of free space.

The force is attractive if the particles have unlike charges, and repulsive if they have like charges.

WOW

Quantisation

You may wonder how negatively charged electrons can orbit the positively charged nucleus of an atom without being attracted into the nucleus and colliding with it. As you saw in the previous chapter, the development of Newton's law of universal gravitation and Kepler's laws gave us a model for planetary motion that helps us understand how Earth and the other planets orbit the Sun without crashing into it.

Niels Bohr noted that Coulomb's law has the same form as Newton's law of universal gravitation. He applied the planetary model developed by Newton and Kepler to atoms, substituting the Coulomb force for gravity. However, this still left other problems unsolved. To make the model work, he introduced quantisation. In Bohr's model the electrons can only sit in defined, quantised, orbits and so are prevented from falling into the nucleus.

PUT AN ELECTRON IN ORBIT AROUND A PROTON

See if you can find an initial position and velocity that will result in the electron orbiting the proton and not colliding with it.

WORKED EXAMPLE 3.7

Calculate the force exerted by a proton on an electron at a distance of 8×10^{-11} m. This is the orbital radius of an electron in hydrogen. (4 marks)

Answer	Logic	
$F = \left(\frac{1}{4\pi\varepsilon_0}\right)\frac{Qq}{r^2}$	Use Coulomb's law to relate force to data given.	1 mark
$q = -Q = 1.6 \times 10^{-19}$ C	Find values needed but not given.	1 mark
$F = (9.0 \times 10^9 \text{ N m}^2 \text{ C}^{-1}) \times \frac{(-1.6 \times 10^{-19} \text{ C})(1.6 \times 10^{-19} \text{ C})}{(8 \times 10^{-11} \text{ m})^2}$	Substitute values, including units.	1 mark
$F = -3.6 \times 10^{-8}$ N	Calculate final value.	1 mark

The negative sign indicates that the force is towards the proton. We could also have done this by simply taking the answer to the previous example and multiplying it by the charge of an electron:

$F = Eq = 2.3 \times 10^{11} \text{ N C}^{-1} \times -1.6 \times 10^{-19}$ C

$= -3.6 \times 10^{-8}$ N

Try these yourself

1 Calculate the force on an electron due to a helium nucleus at the same distance. (4 marks)

2 Calculate the force on the helium nucleus due to the electron. Compare it with the force on the electron due to the helium nucleus. (2 marks)

QUESTION SET 3.2

Remembering

1 Copy and complete the following table by writing 'attractive' or 'repulsive' in the spaces.

Force by → on ↓	Positive charge	Negative charge
Positive charge		
Negative charge		

Understanding

2 Two charges, q_1 and q_2, are such that $q_1 = 3q_2$. What is the ratio of the force $F_{q_1 \text{ on } q_2}$ to $F_{q_2 \text{ on } q_1}$?

3 Draw a table comparing the forces between two spherical or point-like masses and two point-like charges. List both similarities and differences.

Applying

4 Calculate the force on a charge of +2.0 mC a distance 10 cm away from a point charge of +5.0 mC.

5 a At what distance from a proton does an electron experience a force of 2.3×10^{-19} N?

b At what distance does the force have half this value?

c What is the force if the distance is doubled?

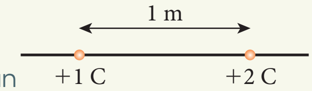
6 What charge is needed to exert a force of 1 N on a charge of 1 mC at a distance of 1 m?

Analysing

7 Two positively charged objects are a distance 1 m apart. The first object has a charge of +1 C. The second has a charge of +2 C, as shown in Figure 3.20.

a Along the line shown, is there a position at which a positively charged object can be placed such that it experiences no net force? Mark the approximate position.

b Along the line shown, is there a position on the line at which a positively charged object can be placed such that it experiences no net force? Mark the approximate position.



▲**Figure 3.20** Find the position on the line at which a negatively charged object experiences no net force.

8 A dipole, such as that shown in Figure 3.7, is in a uniform electric field such that initially the line joining the two charges is perpendicular to the field lines.

a Draw a diagram showing the dipole in this position and show the forces acting on the two charges.

b Explain why there is a torque but no net force acting on the dipole.

c Describe what happens to the dipole if it is free to move.

d Draw a diagram showing the forces acting on the dipole if it is aligned with the field. Is there a net torque in this case?

Energy transfers and transformations in electric fields

You already know a lot about energy. You know that there are two types of energy: potential energy and kinetic energy. Objects that are moving have kinetic energy and we can say that objects that are subject to a force have potential energy. You have already seen and used this many times for the gravitational field close to Earth. Although, strictly, it belongs to the field, we can say that the potential energy of an object of mass m in the gravitational field g at a height h above the surface of Earth is mgh.

You also know that energy is conserved. When you lift up an object such as a pencil and then drop it, gravitational potential energy is transformed into kinetic energy (Figure 3.21a). This transformation happens because of the gravitational force exerted by the gravitational field on the pencil.

Remember from Chapter 2 that the gravitational force is given by $F = mg$. From your studies of work and heat in *Nelson Physics Units 1 & 2 for the Australian Curriculum* you know that the work, W, done by a force, F, on an object as it is displaced a distance, s, in the direction of the force is $W = Fs$. Hence, when the pencil falls through a height h, the gravitational field does work on the pencil, $W = mgh$.

The same happens to a charged object in an electric field (Figure 3.21b). If the object starts at rest and is allowed to move, it will accelerate. We can say that it accelerates because the field's potential energy is converted to kinetic energy; we can say that it accelerates because the force (due to the field) does work on it. These two views are equivalent.

Just as the gravitational field does work on an object falling or being lifted in a gravitational field, the electric field does work on a charged object moving in an electric field. The electric field does work because it exerts a force through some distance. The work done is equal to the change in the potential energy of the system.

Although we often say that an object has potential energy, this is not strictly correct. It is the combination of objects that make up the system that has potential energy. A pencil held above the ground does not really have potential energy. It is the combination of Earth and pencil that has the potential energy. We could also say it is the gravitational field and the pencil, because in the field model of gravity it is the field that applies the force to the pencil. Potential energy exists whenever a force acts between objects. Similarly, a small isolated charge cannot be said to have potential energy. It only has potential energy because of its interaction with an electric field, which is due to some other charged object. Again, the potential energy is due to the application of a force to the charge, which occurs via the field.

You studied energy in Nelson Physics Units 1 & 2 for the Australian Curriculum, when you looked at heat and work in Unit 1 and again when you looked at energy conservation and forces in Unit 2. The idea of work being done by an applied force is used again here.

ELECTRIC FIELD

Position two or more charges in the electric field. Click 'go' to see how they accelerate. You can also view the electric field lines and lines of equipotential.

Remember that although we talk about the potential energy of an object, the potential energy really belongs to the system of objects – to their field.

Hence, fields are not only a way of exerting a force at a distance: *fields also store energy*. The gravitational field stores gravitational potential energy and the electric field stores **electric potential energy**.

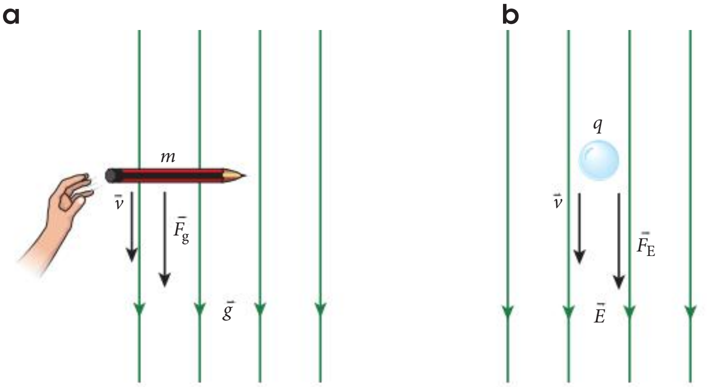


Figure 3.21 ▶ a) A dropped pencil has gravitational potential energy that is transformed into kinetic energy. b) A charged bubble in an electric field has electric potential energy that is converted into kinetic energy.

Electric potential energy and electric potential

Potential difference in circuits was studied in Nelson Physics Units 1 & 2 for the Australian Curriculum.

When you studied circuits in *Nelson Physics Units 1 & 2 for the Australian Curriculum*, you met the idea of **potential difference**. To understand potential difference we first need to understand **electric potential**.

Electric potential (or simply 'potential') has the same relationship to potential energy that field has to force. The field is defined as the force per unit charge: $E = \frac{F}{q}$. The potential is defined as the potential energy per unit charge: $V = \frac{U}{q}$.

The units for V are joules per coulomb, J C^{-1}, also known as **volts, V**.

When you studied circuits, the potential was defined relative to the ground. We assign the ground (or earth) connection a potential of 0 V. If there is no earth connection, then we usually define the zero of potential as being infinitely far away from any charged objects.

We can change where we decide to place the 'zero' of potential, so we usually talk about potential differences rather than potentials. A potential difference is the difference in electric potential between two points:

$$\Delta V = V_{\text{final}} - V_{\text{initial}} = \frac{\Delta U}{q}$$

V is the symbol for potential or potential difference. It has the unit volt, V. Try not to confuse the quantity with its unit!

The unit for potential difference is the volt, V, the same unit as for potential. Sometimes just the symbol V is used for potential difference, but really the Δ should be included to indicate that it is a *difference*, not an absolute value. Sometimes potential difference is called **voltage**, from the unit volt, just as power is sometimes referred to as 'wattage'. The term 'potential difference' is more precise, but 'voltage' is more common.

The change in potential energy when a charged particle is displaced in a field is equal to the work done on the charged object. This is a consequence of conservation of energy. Any work done on the system by moving the charge in the field appears as an increase in the potential energy of the system, assuming that the charge begins and ends at rest or at the same speed. The change in potential is the work done per unit charge. Hence we can also write the potential difference as:

$$\Delta V = \frac{W}{q}$$

Think carefully about your system when you apply conservation of energy. Make sure you know whether your system is open, closed or isolated.

When an object moves in a field, work is done by or on the field and the potential energy of the object in the field changes. The change in potential can be written as either:

$$\Delta V = \frac{\Delta U}{q} \text{ or } \Delta V = \frac{W}{q}$$

WORKED EXAMPLE 3.8

An alpha particle of charge 3.2×10^{-19} C moves a distance of 1.0 m along Earth's fair weather field lines. The field is $100\,\text{V m}^{-1}$ so the particle passes through a potential difference of 100 V. What is the change in potential energy of this particle? (4 marks)

Answer	Logic	
$\Delta V = \dfrac{\Delta U}{q}$	Relate energy to potential.	1 mark
$\Delta U = q\Delta V$	Rearrange for change in energy.	1 mark
$\Delta U = (3.2 \times 10^{-19}\,\text{C})(100\,\text{V})$	Substitute numbers including units.	1 mark
$\Delta U = 3.2 \times 10^{-17}\,\text{J}$	Calculate final answer.	1 mark

Note that we have not specified the direction in which the alpha particle moves, so we do not know whether the change in potential energy is positive or negative.

Try this yourself

What is the change in potential energy of an electron moving through this potential difference? (4 marks)

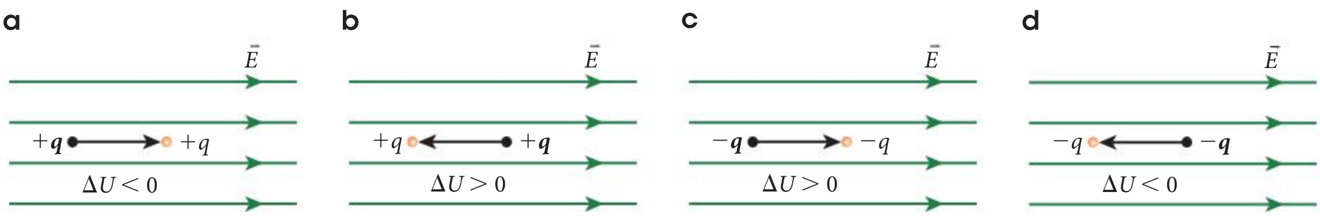
A positively charged object released from rest in an electric field will be accelerated in the direction of the field. The force exerted on the object by the field acts to increase its kinetic energy. From conservation of energy, this kinetic energy must come from somewhere. It comes from a decrease in potential energy. Hence the positive charge has moved from a point of higher electric potential to one of lower electric potential, so ΔV must be negative.

If a negatively charged object moves from higher to lower potential so that ΔV is negative, then the change in potential energy, ΔU, is positive. This can only happen if the charge has some initial kinetic energy, or if some external force is doing work on the system. This is shown in Figure 3.22. This is the case for an electron in a circuit passing through a battery.

Table 3.2 summarises the changes in potential and energy when a charge moves in a field. The field will do work on the charge when a positive charge moves with the field or a negative charge moves against the field. Work must be done on the system to make a positive charge move against the field or a negative charge move with the field.

CHARGES, FIELDS AND POTENTIALS

Set up some charges then move the gauge around to measure the potential at different points. Try it with a single charge, then with a dipole.



▲**Figure 3.22** Energy changes when a charge moves in an electric field. a) A positive charge moving in the direction of the field, b) a positive charge moving against the direction of the field, c) a negative charge moving in the direction of the field, d) a negative charge moving against the direction of the field.

We can use the idea of derivatives to relate the field to the potential mathematically.

The field is $E = -\frac{dv}{dx}$.

You can see this from the units: V m^{-1} and V respectively. Force and energy are related in the same way: $F = -\frac{du}{dx}$.

Be clear about potential, potential difference and potential energy. Potential is the potential energy per unit charge at a point, relative to some zero point. Potential difference, sometimes called voltage, is the difference in potential between two points.

Table 3.2 Changes in potential and energy for a charge moving in a field

Charge	Movement *with* or *against* the field lines	Change in potential	Change in potential energy	Work done *by* or *on* the field
+	With	Negative (decrease)	Negative (decrease)	By
+	Against	Positive (increase)	Positive (increase)	On
–	With	Negative (decrease)	Positive (increase)	On
–	Against	Positive (increase)	Negative (decrease)	By

In a circuit, when an electron passes through a resistor, it goes from a point of lower potential to one of higher potential. (Remember that electrons travel in the opposite direction to conventional current.) As it does so, the potential energy of the electron decreases. We also know that the current through a single loop circuit is the same at all points. Electrons are not created or destroyed in a circuit. Therefore, if the current is the same, the electrons will have the same drift speed, and the kinetic energy of the electrons has not changed. We know from conservation of energy that the energy must be somewhere, and in fact it appears as heating of the resistor.

Electrons and other subatomic particles moving through potential differences are so common that a special unit is used to describe the change in energy when this happens. When an electron moves through a potential difference of 1 V it has a change in energy of $\Delta U = q\Delta V = e\Delta V = (1.6 \times 10^{-19}\,\text{C})(1\,\text{V}) = 1.6 \times 10^{-19}\,\text{J} = 1$ electron volt. The **electron volt** or eV is a good size for describing the energy of subatomic particles. The SI unit joule is very large by comparison. You will see the unit electron volt a lot whenever you study quantum physics, particle physics or radiation. It is also used in chemistry to describe the energy of reactions.

The zero of potential energy

You may have noticed that we have mainly been talking about *changes* in potential and potential energy because this is what we measure. These changes can be positive or negative. Potential and potential energy can themselves also be positive or negative, depending on how we define the zero of potential energy. We use the same convention in electrostatics as we use in gravity. The zero of potential energy is when the components of the system are infinitely separated, so all the charges that make up the system are very far apart.

Consider the case of two positive charges very far apart. This arrangement has zero potential energy. If we want to bring them closer together into some final arrangement, we have to do work on them because they will repel each other. We put work into the system, so the final potential energy is positive. Now consider a positive charge and a negative charge very far apart, so they have zero potential energy. They will act to attract each other without external work needing to be done. If we release them from far apart they will move closer together, gaining kinetic energy and hence losing potential energy. A system of a positive and negative charge at any separation less than infinite has negative potential energy. This is also the case for a pair of masses, as gravity is always attractive.

We define the zero of electric potential energy as all charges in the system being infinitely far apart. Any other arrangement may have positive or negative potential energy.

WORKED EXAMPLE 3.9

An electron in an X-ray machine is accelerated through a potential difference of 100 kV before colliding with a target and emitting X-rays. What is the energy of the electron just before it hits the target? Give your answer in electron volts and joules. (5 marks)

Answer	Logic	
$E_k = \Delta U$	By conservation of energy the kinetic energy is the change in potential energy.	1 mark
$E_k = \Delta U = q\Delta V$	Relate the energy to the parameters given.	1 mark
$E_k = e(100000\,\text{V})$	Substitute numbers, with units.	1 mark
$E_k = 100\,\text{keV}$	Calculate final value.	1 mark
$E_k = 100\,\text{keV} \times 1.6 \times 10^{-19}\,\text{C} = 1.6 \times 10^{-14}\,\text{J}$	Convert units.	1 mark

You can see that the unit eV is much more suitable for describing energies on this scale than the joule, even when large potential differences are involved.

Try this yourself

Find the speed at which the electron will be moving after passing through this potential difference. Note that to get a better estimate of the speed we should really use a relativistic model, as described in Chapter 6. (3 marks)

EXPERIMENT 3.1

MAPPING EQUIPOTENTIALS

Equipotentials are lines or surfaces along which the potential is constant. Hence, the potential difference between any two points on an equipotential is zero. Lines of equipotential are always perpendicular to field lines. This means that we can use lines of equipotential to construct field lines.

Aim

To map lines of equipotential and use these to draw field lines for a dipole

Materials

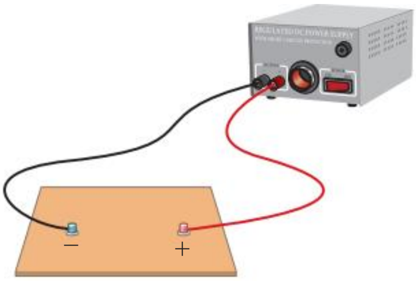
- 12 V DC power supply with leads
- conductive paper
- voltmeter

What are the risks in doing this experiment?	How can you manage these risks to stay safe?
Power supplies can be dangerous if not used correctly.	Only connect the power supply as instructed. Keep liquids away from the power supply.

In your write-up, add any more risks you can think of, as well as ways to manage them.

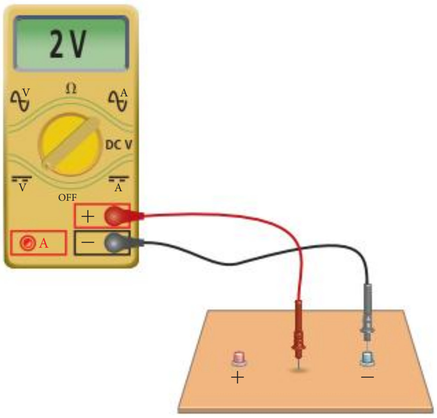
Procedure

1 Attach the positive terminal of the power supply to a point near one end of your conductive paper and the negative terminal near the other end as shown in Figure 3.23. You now have a dipole with a positive and negative electrode on your paper.



◀**Figure 3.23** Experimental set-up

2 With one probe from your voltmeter touching the negative electrode, move the other probe around on the paper until you get a reading of 2 V (Figure 3.24). Mark this point.



◀**Figure 3.24** Measuring the potential

3 Carefully move the probe around and mark other points of 2 V potential on the paper. Join the dots: this is your first equipotential line. Label this line.

4 Repeat steps 2 and 3 to map out equipotential lines of 4 V, 6 V, 8 V and 10 V potential. Label these lines on the paper.

Results

You should now have a piece of paper showing a set of lines of equipotential. Record the positions of the electrodes on your paper by tracing around them.

Analysis of results

Use the equipotential lines to draw electric field lines for your arrangement of electrodes.

Plot graphs of potential as a function of distance from one of your electrodes. Do this for the line joining the two electrodes and one other line.

Discussion

Are the equipotential lines equally spaced?

Do the field lines look the way you would expect from the field line drawings in Figure 3.7?

Describe the relationship between potential and distance from the electrodes. Can you write an equation to describe this relationship?

Taking it further

How could you extend this experiment?

QUESTION SET 3.3

Remembering

1 What is the relationship between electric potential and electric potential energy?

2 What are units for electric potential, potential difference and potential energy?

Understanding

3 Show that 1 V is equal to $1\,kg\,m^2\,s^{-3}\,A^{-1}$ in SI base units.

4 An alpha particle is ejected from a radioactive atom, leaving behind a negative ion. As the alpha particle moves away from the ion, is work done by or on the alpha particle by the field due to the ion?

Applying

5 A charge of 2.0 C flows through a resistor. The potential difference across the resistor is 1.5 V. Given that the electrons have the same kinetic energy at each end of the resistor, how much energy must be dissipated by the resistor?

6 A current of 1.0 A flows through a 12 V battery for 1.0 s. What is the change in potential energy of the battery?

Analysing

7 What is the speed of a particle that has been accelerated from rest through a potential difference of 1000 V when the particle is:

a an electron?

b a proton?

c an alpha particle?

Reflecting

8 How has your understanding of what goes on in electric circuits been changed as a result of this Unit? Draw a concept map linking the concepts that you learnt in *Nelson Physics Units 1 & 2 for the Australian Curriculum* Unit 1 with the new ideas introduced in this unit.

9 Draw diagrams summarising your understanding of potential energy in gravitational and electric fields. Include diagrams showing how potential energy changes when charged objects and objects with mass are moved from the zero of potential energy to some final position.

CHAPTER SUMMARY

- Electric charges create electric fields.
- Electric fields exert forces on electric charges.
- The electric field is the force per unit charge acting on a small positive test charge. It has units $N\,C^{-1}$ and is a vector.
- Electric field lines show the direction of the electrostatic force acting on a positive test charge.
- Electric field lines go in to negative charges and come out from positive charges.
- An electric dipole is a closely spaced pairs of positive and negative charges.
- Dipoles experience a torque in a field.
- The kinematics equations are used to model the motion of a charged particle in a uniform field.
- The electric field due to a point charge varies with the size of the charge and decreases with the distance squared from the charge:

$$E = \left(\frac{1}{4\pi\varepsilon_0}\right)\frac{Q}{r^2}$$

- The force that one point charge, Q, exerts on a second point charge, q, is given by Coulomb's law:

$$F = \left(\frac{1}{4\pi\varepsilon_0}\right)\frac{Qq}{r^2}$$

- Electric fields store electric potential energy.
- The potential energy stored by the field changes when work is done on or by the field.
- Electric potential is the potential energy per unit charge in an electric field. It has units J C^{-1} or V.
- Potential difference is the difference in potential between two points in an electric field. It also has units V and is sometimes called voltage.

CHAPTER GLOSSARY

coulomb the unit of charge, named after Charles Augustin de Coulomb

Coulomb constant the constant of proportionality for electric fields, $k_e = \frac{1}{4\pi\varepsilon_0} = 9 \times 10^9\,\text{N m}^2\text{C}^{-2}$

Coulomb's law the relationship between force, charge and distance for two point charges

dipole a closely spaced pair of positive and negative charges (electric dipole) or north and south poles (magnetic dipole)

electron charge charge of an electron, -1.6×10^{-19} C

electric field the field due to electric charges, which applies a force to electric charges

electric potential potential energy per unit charge in an electric field

electric potential energy the potential energy stored in an electric field; the potential energy of a charge in an electric field

electromagnetic force the combination of electric and magnetic forces acting on a charge, due to an electric and a magnetic field

electron volt, eV a unit of energy equal to 1.6×10^{-19} J

electrostatic field model the model that assigns an electric field to stationary charges; it is this field that exerts forces on other charges

electrostatic force force due to and acting on stationary charges

permittivity of free space, ε_0 the physical constant that determines how large an electric field is produced by a charge in vacuum. It has the value $8.9 \times 10^{-12}\,\text{C}^2\text{N}^{-1}\text{m}^{-2}$

polar having regions of positive and negative charge

potential difference the difference in potential between two points in an electric field

uniform field a field with a constant magnitude and direction over some region in space

van der Waals forces electrostatic forces due to charged regions of molecules, resulting in weak chemical bonds including hydrogen bonds

voltage more correctly called the potential difference, it is measured in volts

volt, V unit of potential difference, $1\,\text{V} = 1\,\text{J C}^{-1}$

CHAPTER REVIEW QUESTIONS

Remembering

1 How is electric field related to force?

2 How is electric potential related to electric potential energy?

3 List three ways in which electric fields and gravitational fields are similar.

Understanding

4 Two charged particles enter a uniform electric field, as shown in Figure 3.25. Both particles have the same magnitude charge.

a Which of the two charged plates creating the field shown is the positively charged plate?

b What can you say about the charges and relative masses of the two particles?

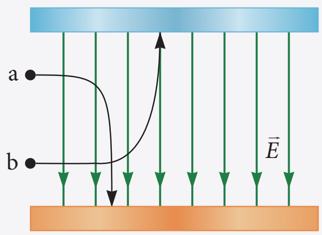
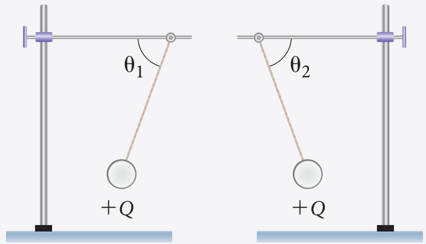


Figure 3.25

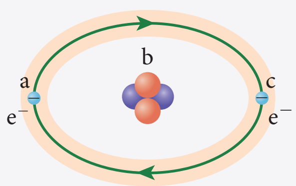
5 Two small balls with equal positive charges are hung from retort stands as shown in Figure 3.26. Draw a diagram showing how the balls will hang if the charge on the second ball is reduced but that on the first ball remains the same. Clearly show how the angles change.



◀Figure 3.26

6 Two small balls with equal positive charges are hung from retort stands as shown in Figure 3.26 above. Draw a diagram showing how the balls will hang if the charge on the second ball is removed but that on the first ball remains the same. Clearly show how the angles change.

7 Consider three charged particles, a helium nucleus and its two orbiting electrons, arranged as shown in Figure 3.27. Rank the magnitudes of the forces $F_{\text{a on b}}$, $F_{\text{a on c}}$, $F_{\text{b on a}}$, $F_{\text{b on c}}$, $F_{\text{c on a}}$ and $F_{\text{c on b}}$ from smallest to largest.



◀Figure 3.27

8 The dome of a van de Graaff generator is charged up such that there is an electric field in the region surrounding it. A small test charge with charge *q* is brought into this region and placed at a point P. If the test charge is replaced with a small test charge of 2*q*, what happens to:

a the electric field at point P?

b the force on the test charge?

c the electric potential at point P?

d the potential energy of the dome or test charge system?

9 The field at a distance of 1 m from the dome of a van de Graaff generator is 1000 V m^{-1}. Model the dome as a point charge at this distance.

a By how much would the charge on the dome need to increase for the field at this point to double?

b At what distance from the dome is the field now 1000 V m^{-1}?

Applying

10 A dust particle has a mass of 1×10^{-6} g. What charge must this dust particle have for the electrostatic force on it due to Earth's fair weather field with magnitude 100 V m^{-1} to balances the gravitational force? Give the sign and magnitude of the charge.

11 A polystyrene bead with mass 1.0 g and charge +1.0 nC is in an electric field of 50 V m^{-1}. Calculate the force on and acceleration of the bead due to the field.

12 An electron moving at 25 km s^{-1} enters a uniform electric field of magnitude 550 V m^{-1} in the direction of its velocity.

a What is the force on the electron?

b After what distance does the electron come to a stop?

13 Two protons in a molecule are 3.80×10^{-10} m apart.

a Find the magnitude of the electric field due to one of the protons at this distance.

b Find the magnitude of the electrostatic force exerted by one proton on the other.

14 An electron moves in an electric field from a point at which the potential is 1.5 V to a point at which the potential is 3.5 V.

a By how much does the potential energy of the electron–field system change?

b Is this change positive or negative?

c Where does the energy come from or go to?

15 Two protons in a nucleus are a distance 2.0×10^{-15} m apart.

  a What is the magnitude of the electrostatic force that they exert on each other?

  b What is the magnitude of the gravitational force that they exert on each other?

  c Repeat parts a and b for a pair of neutrons the same distance apart.

Analysing

16 A proton is released from rest in a uniform electric field.

  a Describe the motion of the proton.

  b Describe the changes in the:

    i proton's kinetic energy.

    ii potential energy of the proton–electric field system.

  c Does the field do work on the proton, or does the proton do work on the field?

17 An electron moving at $20\,km\,s^{-1}$ enters a uniform electric field of magnitude $500\,V\,m^{-1}$ in the direction of its velocity.

  a What is the kinetic energy of the electron?

  b Use an energy approach to calculate the distance it takes for the electron to come to a stop.

  c What is the potential difference between the initial point at which the electron entered the field and the point at which it comes to a stop?

Reflecting

18 Read about Earth's electric field and the global circuit. Compare and contrast this circuit with circuits that you learnt about in Unit 1. Prepare a 10-minute talk on Earth's electric field and the global circuit that could be presented to students currently studying the Unit 1 module on circuits.

19 The field model allows us to explain how forces act at a distance. Research one alternative model and summarise its main features. What evidence supports this other model? Contrast this model with the field model. Evaluate the two and determine which is the better model.

20 Draw a spider diagram to connect all the concepts and equations in this chapter.

CHAPTER 4 MAGNETIC FIELDS

Shutterstock.com/Daniel Gale

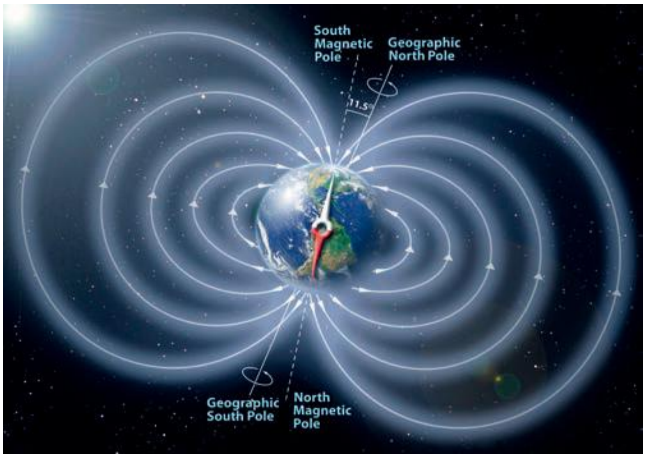
By the end of this chapter you will have covered the following material.

Science Understanding

- Current-carrying wires are surrounded by magnetic fields; these fields are utilised in solenoids and electromagnets (ACSPH106)
- The strength of the magnetic field produced by a current is called the magnetic flux density (ACSPH107)
- Magnets, magnetic materials, moving charges and current-carrying wires experience a force in a magnetic field; this force is utilised in DC electric motors (ACSPH108)

Introduction

We have seen that gravitational fields are created by objects with mass, and that electric fields are created by objects with charge. The third field model that we will now explore is the **magnetic field**. Magnetic fields are important to us in many ways. Earth has a magnetic field that protects us from cosmic rays from the Sun (see Figure 4.1). Magnetic materials are used in many devices, such as transformers and USB drives. Magnetic fields interact with electric fields to make motors and generators work.



Peter Reid (peter.reid@ed.ac.uk)

Figure 4.1 ▲ Earth's magnetosphere

Earth's magnetic field has been used for navigation by animals such as pigeons, sharks, bees and bacteria for millions of years. Humans have used it for thousands of years: there is evidence that the compass was used in China in the 13th century BCE. Animals such as pigeons, sharks, bees and bacteria have used it for millions of years. The Greeks knew about magnetism as early as 800 BCE. They discovered that the stone **magnetite** (Fe_3O_4) attracts iron. One legend says that the material magnetite was named after the shepherd Magnes. The iron nails of his shoes reportedly stuck to pieces of magnetite as he pastured his flocks. The words 'magnet' and 'magnetism' derive from 'magnetite'.

Most magnetic materials in use today still contain iron. These materials make transformers and electromagnets more efficient. We use magnetic materials for most of our data storage including hard drives, USB drives and credit cards.

William Gilbert, an English physician in the 16th century, hypothesised that Earth itself is magnetic. He deduced from this that the centre of Earth is made of iron. He also performed many experiments with magnets and discovered that cutting a magnet in half results in two smaller magnets, *not* in separated north and south poles.

In the 19th century, experiments by Michael Faraday and Hans Christian Oersted, as well as many others, showed that magnetic fields arise from currents. Hence, moving charges are a source of magnetic fields.

Magnetic materials such as magnetite remained a mystery; it is only in the last hundred years that quantum mechanics has begun to explain these materials. It is the properties and behaviour of fundamental particles such as electrons that give rise to the magnetic behaviour of iron.

This means that to understand magnetism we need to extend our previous model, the electrostatic field model. We must include moving charges. When we do this, our new model is called the **electromagnetic field model**. We will look at the magnetic field part of this model now. In the next chapter we will look at how the two fields interact to give us electromagnetism.

The magnetic field

Remember that a vector quantity is written as $\vec{A}$, and it has a magnitude and direction. The magnitude of vector $\vec{A}$ is written as A.

We can measure the magnetic field by its effect on moving charges, for example a current in a wire. Experiments using lengths of current-carrying wire in magnetic fields show that the **magnetic force** on the wire depends on three things:

- the size of the current carried
- the strength of the magnetic field
- the angle between the direction of the current and the field.

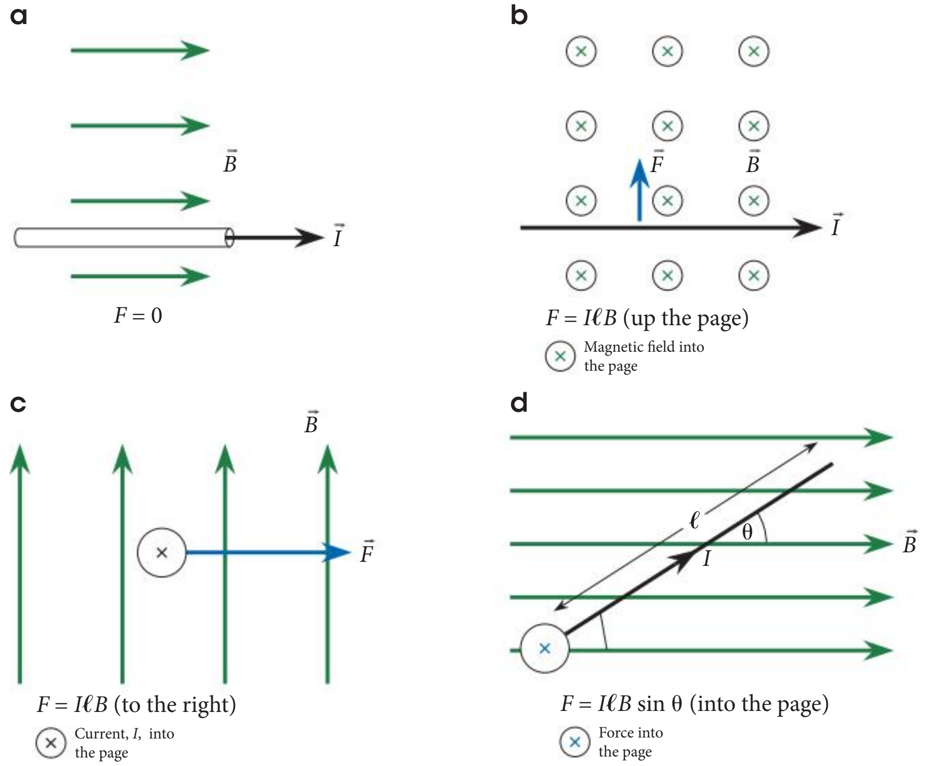
These observations can be summarised mathematically as:

$$F = I\ell B\sin\theta$$

where F is the magnitude of the force vector $\vec{F}$, B is the magnitude of the magnetic field, $\vec{B}$, I is the magnitude of the current, $\vec{I}$, ℓ is the length along which the current flows and θ is the angle between the vectors I and B.

This is not a vector equation, but it does use the magnitudes of two vectors, $\vec{B}$ and $\vec{I}$, to give us the magnitude of a third vector, the force, $\vec{F}$. $\vec{I}$ is the vector representing the current and θ is the angle between the vectors $\vec{B}$ and $\vec{I}$. The angle θ tells us how the directions of the vectors $\vec{B}$ and $\vec{I}$ affect the direction and magnitude of vector $\vec{F}$.

This expression tells us that the force is zero if $\vec{B}$ and $\vec{I}$ are in the same direction, and a maximum when they are perpendicular to each other. Figure 4.2 shows some examples.



◀ **Figure 4.2**
The *magnetic force* applied to a current-carrying wire depends on the angle between the current element within the field and the magnetic field.

We have said that F is the *magnitude* of the force $\vec{F}$. A complete description of the force includes the *direction* of the force, because force is a vector. The magnetic field, $\vec{B}$, is also a vector, as is the current $\vec{I}$ because it points in a particular direction. So, just as with other forces, the force due to the magnetic field is a vector and we need to think about directions. Magnetic fields apply a force perpendicular to the direction of both the magnetic field and the current.

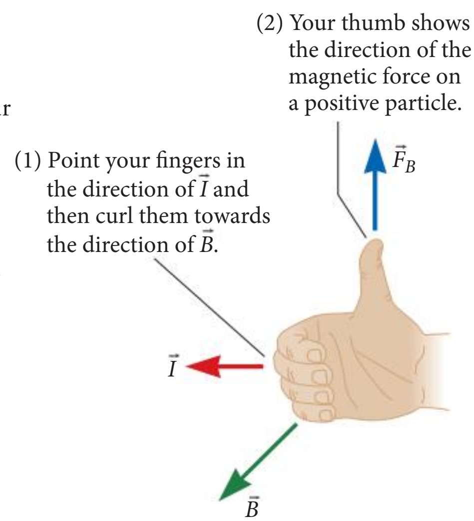
The right-hand rule

To find the direction of the force we use the right-hand rule as shown in Figure 4.3. Hold the fingers of your right hand together and point your thumb straight out, at a right angle to your hand. Move your hand so your fingers point in the direction of the current. Curl them towards the field. Your thumb is now pointing in the direction of the force.

You may come across other ways of applying the right-hand rule. You can use whichever version you like, because they all give the same answer.

For example, you can point your index finger towards $\vec{I}$, then point your middle finger towards $\vec{B}$. Your thumb will then point towards $\vec{F}$, as shown in Figure 4.4(a). This is exactly the same as the rule shown in Figure 4.3 except that you leave your index finger behind when you curl your fingers.

Figure 4.3 ▼
The right-hand rule for currents in magnetic fields



Some people point their thumb in the direction of $\vec{I}$, their fingers in the direction of $\vec{B}$, and then their palm is facing towards $\vec{F}$. This is sometimes called a right hand slap rule because it looks as though you are about to slap in the direction of $\vec{F}$. This is shown in Figure 4.4(b).

These are all just mnemonics for solving a **vector cross product**. The important thing is that you find the way that you can remember, and then stick to it. Just remember to always use your *right hand*!

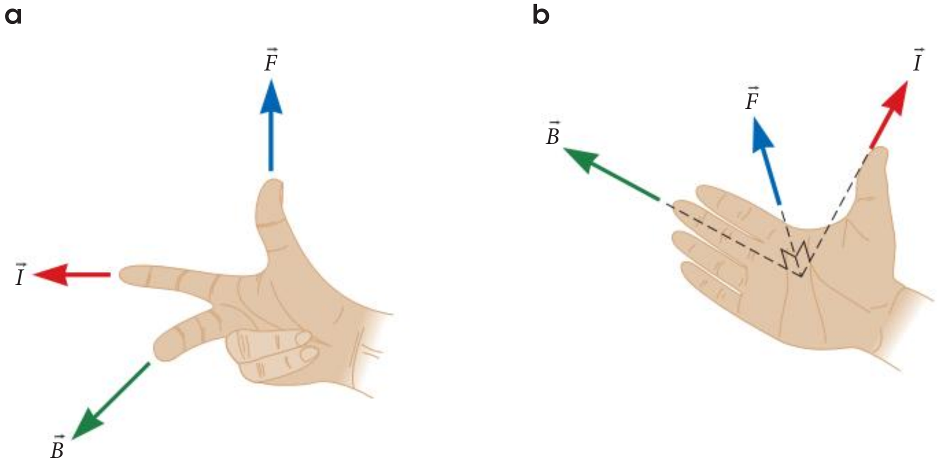


Figure 4.4 ▶ There are two variations on the right-hand rule: a) the pointing version and b) the right-hand slap.

WOW

Vector cross product

The equation for the force acting on a current element due to a magnetic field is more correctly written as:

$$\vec{F} = \ell(\vec{I} \times \vec{B})$$

This is a vector equation using a vector cross product ($\vec{I} \times \vec{B}$). The cross product is necessary because the force depends on the angle between $\vec{I}$ and $\vec{B}$. The cross product of two vectors is another vector. The vector $\vec{F}$ is perpendicular to both $\vec{I}$ and $\vec{B}$. It is perpendicular to the plane defined by $\vec{I}$ and $\vec{B}$. It is important to note that $\ell(\vec{I} \times \vec{B}) \neq \ell(\vec{B} \times \vec{I})$: in fact $\ell(\vec{I} \times \vec{B}) = -\ell(\vec{B} \times \vec{I})$, indicating that they are in opposite directions.

If you are studying specialist maths you will already have used both the vector dot product and the vector cross product. You will see the vector cross product $\vec{C} = \vec{A} \times \vec{B}$ often in physics. $\vec{A}$, $\vec{B}$ and $\vec{C}$ are vectors and the magnitude of $\vec{C}$ is $C = AB\sin\theta$. The direction of $\vec{C}$ is given by the right-hand rule, shown in Figure 4.3.

The definition of magnetic field

VECTOR CROSS PRODUCT

Work through this simulation to see how the vector cross product $\vec{A} \times \vec{B} = \vec{C}$ varies as you change the magnitudes of $\vec{A}$ and $\vec{B}$ and the angle between them.

From the equation $F = I\ell B\sin\theta$ we get our definition of magnetic field. The field is the force per unit current, per unit length, which we write mathematically as:

$$B = \frac{F}{I\ell\sin\theta}$$

The magnetic field, $\vec{B}$, has units of $\mathrm{N\,s\,C^{-1}\,m^{-1}}$ or in basic units $\mathrm{kg\,s^{-1}\,C^{-1}}$. This has the name **tesla, T**, after Nikolai Tesla. $\vec{B}$ is also called the **magnetic flux density**.

This definition of magnetic field is similar to our definitions of electric and gravitational fields. All three fields are defined as force per unit 'something', where 'something' is the property that the field acts on. In the case of gravity, the something is mass, *m*; for electric fields it is charge, *q*; for magnetic fields it is the current element, $\vec{I}\ell$. This 'something' – *m*, *q* or $\vec{I}\ell$ – is also the property that creates the field.

The unit tesla, T, is a large unit. Most magnetic fields that you come across are measured in micro- or millitesla (see Table 4.1). Only in the last few decades have we been able to create magnetic fields larger than a few tesla.

Table 4.1 Some typical magnetic field strengths

Source of field	Magnitude, *T* (approximate)
Earth (surface)	3×10^{-5} to 6×10^{-5}
Typical fridge magnet (surface)	5×10^{-3}
Modern rare earth magnet (surface)	1
Medical MRI system (inside)	3
Strong research facility field	100
Neutron star	$>10^{8}$

WORKED EXAMPLE 4.1

An overhead power line carrying a current of 10 000 A is in a magnetic field of 30 μT. The field is perpendicular to the wire. What force per unit length does the wire experience? (4 marks)

Answer	Logic	
$F = I\ell B\sin\theta$	Relate the force to the field and current.	1 mark
$F/\ell = IB\sin\theta$	Rearrange for force per unit length.	1 mark
$F/\ell = (10000\,\text{A})(30 \times 10^{-6}\,\text{T})\sin(90°)$	Substitute values with units.	1 mark
$F/\ell = 0.30\,\text{AT} = 0.30\,\text{N m}^{-1}$	Calculate the final answer.	1 mark

This is quite a large force considering how long a span of high voltage power line is. However, remember that the current is alternating 50 times per second. The force also alternates direction rather than constantly pulling the wire in a single direction.

Try this yourself

What total force would be exerted on a span of the wire if it was 200 m long? How does this force compare with the gravitational force acting on the wire if the wire has a linear density of $750\,\text{kg km}^{-1}$? (4 marks)

Charges moving in magnetic fields

A current is a collection of charges moving in the same direction. Recall from Unit 1 that current is charge per unit time, $I = \frac{q}{t}$. A current experiences a force in a magnetic field; therefore, we expect that a single moving charge will also experience a force. Experiments show that this is the case. The equation for the force is the same, but we need to replace $I\ell$ with qv. Now $I\ell = \frac{q\ell}{t} = qv$. The magnitude of the force is given by

$$F = (qv)B\sin\theta$$

where q is the charge on the particle, measured in C, v is the magnitude of its velocity in m s^{-1}, and θ is the angle between the vectors $q\vec{v}$ and $\vec{B}$. The direction of the force is again perpendicular to both the magnetic field and the velocity.

We need to use the right-hand rule again (Figure 4.3). Point the fingers of your right hand in the direction of the velocity, then curl them towards the field. Your thumb gives the direction of the force. Note that q and $\vec{v}$ are together in brackets. Hence if the charge is negative, the direction of $q\vec{v}$ is opposite the direction of $\vec{v}$. In this case your fingers start by pointing opposite to $\vec{v}$ and then curling towards $\vec{B}$.

The force acting on a charged particle moving with velocity, $\vec{v}$, in a magnetic field, $\vec{B}$, is given mathematically by the vector cross product $\vec{F} = q\vec{v} \times \vec{B}$. This is another example of the use of vector cross products in physics.

A particle with charge q moving at velocity v in a magnetic field B experiences a force given by:

$$F = qvB \sin\theta$$

where θ is the angle between the field and the velocity of the particle. The direction of the force is perpendicular to both the velocity and the field and is given by the right-hand rule.

Once we have the force on the moving charge (Figure 4.5a), we use Newton's laws to calculate the acceleration. We can then use kinematics to determine its trajectory (Figure 4.5b).

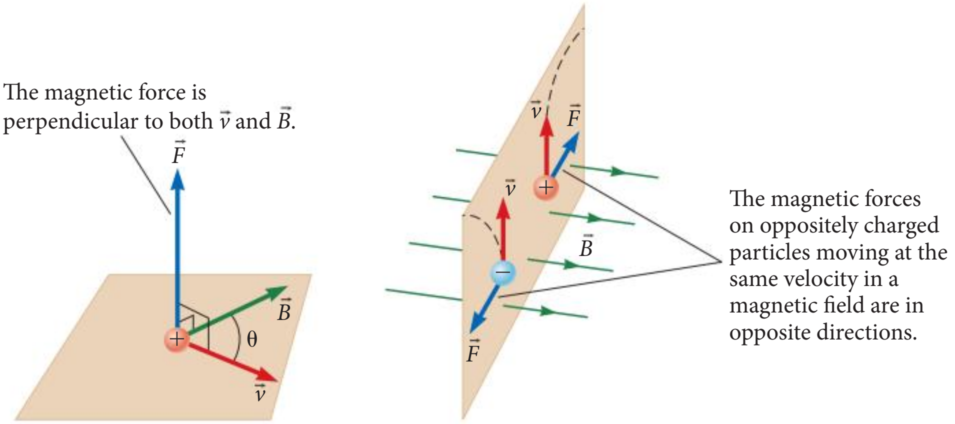


Figure 4.5 ▶ a) The force on a positively charged particle in a magnetic field is perpendicular to both the field and the particle's velocity. b) The forces on positively and negatively charged particles in a magnetic field. The dashed lines show the paths of the particles.

WORKED EXAMPLE 4.2

An alpha particle enters Earth's magnetic field at a velocity of 55000 m s^{-1}. The local field strength (magnetic flux density) is 40 μT. What is the range of possible accelerations of the alpha particle? (9 marks)

Answer	Logic	
$a = \frac{F}{m}$	Relate acceleration to force using Newton's second law.	1 mark
$F = qvB\sin\theta$	Relate the force to the given parameters.	1 mark
$a = \frac{qvB\sin\theta}{m}$	Relate acceleration to the given parameters: acceleration, a, can have any value in the range $\frac{-qvB}{m}$ to $\frac{+qvB}{m}$.	3 marks
$q = 2e = 2 \times 1.6 \times 10^{-19}$ C $= 3.2 \times 10^{-19}$ C, $m = 6.6 \times 10^{-27}$ kg	Find the values needed.	1 mark
$a = \frac{qvB}{m} = \frac{(3.2 \times 10^{-19}\text{ C})(55\,000\text{ m s}^{-1})(40 \times 10^{-6}\text{ T})}{6.6 \times 10^{-27}\text{ kg}}$	Substitute values with the correct units.	1 mark
$a = 1.1 \times 10^{8}$ m s^{-2}.	Calculate the final value.	2 marks

Acceleration can have any value between -1.1×10^{8} m s^{-1} and $+1.1 \times 10^{8}$ m s^{-2}. The value depends on the angle between the velocity and the magnetic field.

Try these yourself

What acceleration would an electron entering this field at the same speed have if it was moving in a direction: (7 marks)

a perpendicular to the field?

b parallel to the field?

c at an angle of 45° to the field?

Paths of particles in magnetic fields

As you saw in Worked example 4.2, fast-moving charged particles experience large forces even in small magnetic fields. These forces result in large accelerations. You might expect, therefore, that the field is doing a lot of work on the particle. This is not the case. Work done by a force is equal to the product of the force and the distance over which it acts. The distance is the component *in the direction of the force* only.

$$W = Fx\cos\theta$$

where θ is the angle between the force and the displacement, x.

Magnetic fields do no work on charged particles. The force is perpendicular to the velocity, and hence to the displacement, at any moment in time. The field changes the direction of the velocity, and hence the path of the particles, but not their speed.

The path of a charged particle in a uniform magnetic field depends on the angle between the initial velocity and the field. The path may be a straight line, a circle or a helix.

The particle has a straight-line trajectory only when its velocity is parallel to the field. In this case the force on the particle is zero.

If the particle is moving in a plane perpendicular to the field then the force is always perpendicular to the displacement, and the particle moves in a circle.

If the particle has a velocity with a component in the direction of the field, this component of the velocity is not altered by the field. However, the perpendicular component is altered by the acceleration due to the field. In this case the particle follows a helical path, with the axis of the helix in the direction of the field. Figure 4.6 shows some possible paths of charged particles in magnetic fields.

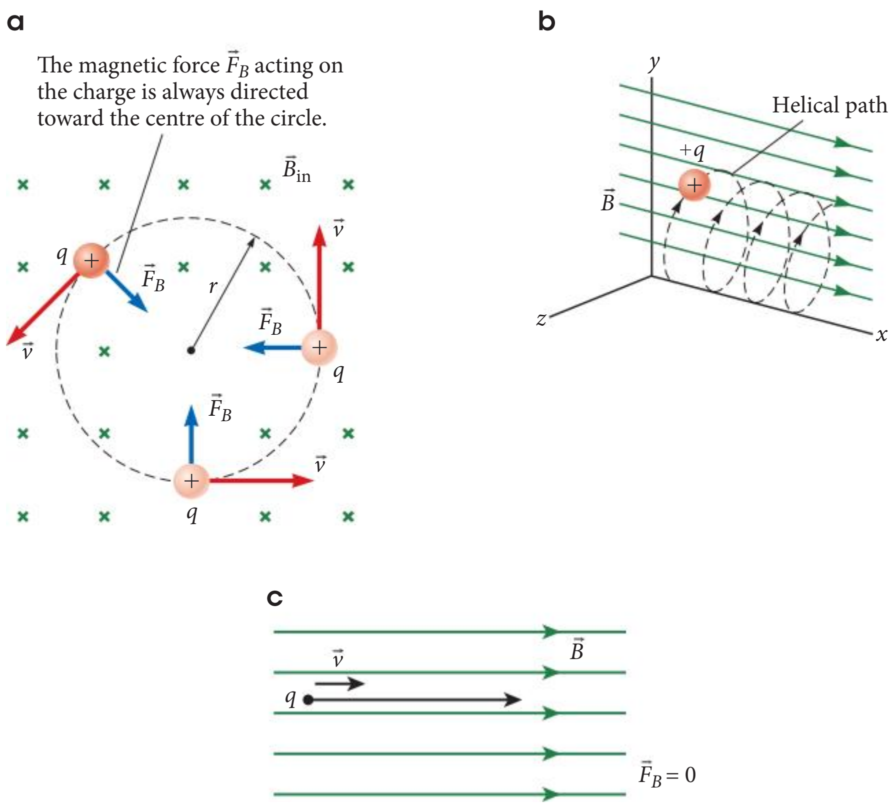
This behaviour of charged particles in magnetic fields is very useful. **Synchrotrons** use magnetic fields to contain fast-moving charged particles. The particles are contained in giant rings, tens to hundreds of metres across. The magnetic fields keep the particles circulating in the ring. As the particles circle, they emit high-energy light. The high-energy light is used for medical, materials science and fundamental physics research.

You saw this relationship between force and work in Nelson Physics Units 1 & 2 for the Australian Curriculum, Chapter 9.

A satellite in a circular orbit experiences a gravitational force perpendicular to its velocity. No work is done by the gravitational field. This is analogous to a charged particle circling in a magnetic field. No work is done, so there is no change in energy.

CHARGED PARTICLE IN ELECTRIC AND MAGNETIC FIELDS

This simulation shows you how a charged particle moves in an electric or a magnetic field. Note the shapes of the paths.



◀ Figure 4.6 In a uniform magnetic field, the path of a charged particle is a) circular when the particle's velocity is perpendicular to the field, b) helical when the particle's velocity has components perpendicular and parallel to the field and c) an un-deflected straight line when the velocity is parallel to the field.

The orbiting particles in synchrotrons move so fast that they need to be modelled using relativistic equations. The classical field equations predict speeds much higher than are actually possible. Relativity is discussed in Chapter 6.

In a **synchrotron**, the magnetic field acts perpendicular to the orbiting particle's velocity. Hence the force on the orbiting particle is

$$F = qvB$$

In chapter 1 we saw that when an object is undergoing uniform circular motion, with radius of orbit *r*, its centripetal acceleration is given by

$$a = \frac{v^2}{r}$$

and hence by Newton's second law the net force acting on it is

$$F = ma = \frac{mv^2}{r}$$

It is the magnetic field which supplies this centripetal force. Equating these two expressions gives

$$qvB = \frac{mv^2}{r}$$

The Australian Synchrotron is described in more detail in the Scientific literacy box in Chapter 5.

In a synchrotron, the orbital radius, *r*, is constrained by the size of the ring. The mass and charge are characteristic of the type of particle. Hence the orbital velocity is controlled by varying the size of the magnetic field. The velocity is given by

$$v = \frac{qBr}{m}$$

Given that the particle travels a distance $2\pi r$ in each period, *T*, the period is given by

$$T = \frac{2\pi m}{qB}$$

The particles used in synchrotrons are typically protons and electrons. They reach speeds more than half that of the speed of light and hence need to be treated as relativistic particles. This means that the actual velocity attained is not as high as calculated using the expression above. The classical electromagnetic field model does not give an accurate prediction of the particle speed. To more accurately calculate the speed attained we need to use a relativistic model. Relativity is discussed in Chapter 6.

Magnetic fields are also used in **mass spectrometers** to determine which atoms or molecules are in a sample of material. Mass spectrometers are used in airport security, forensic analysis and even in surgery. The 'iKnife' is an electrosurgical knife that burns away tissue as it cuts. The smoke from the tissue is passed to a mass spectrometer for analysis, and cancerous tissue can be instantly identified.

iKNIFE

Read more about the iKnife and its uses.

WOW

Mass spectrometers and the fight against terrorism

You may have seen mass spectrometers at airport security checkpoints. These are usually used for random spot checks after you pass through the metal detector. (The metal detector also uses a magnetic field.) These small mass spectrometers are used to detect nitrates and other chemicals commonly contained in explosives.

A small sample is taken by wiping a swab over a person's hands or clothes. The sample is then ionised inside the mass spectrometer, and the resulting ions passed first through an electric field and then through a magnetic field. The electric field accelerates the ions in a straight line. The moving ions then enter a perpendicular magnetic field. This curves their paths into an arc of a circle. The radius of curvature depends on the mass and charge of the ions. Hence the curvature can be used to determine the chemicals that are present.

If nitrates or other suspicious substances are detected, then you and your luggage will be searched.

Alamy/David Burton

▲ **Figure 4.7** A sample is placed inside a mass spectrometer at airport security.

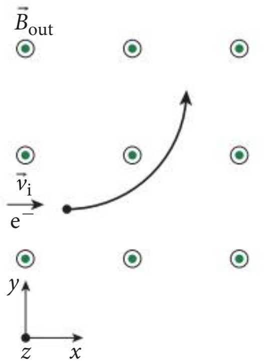
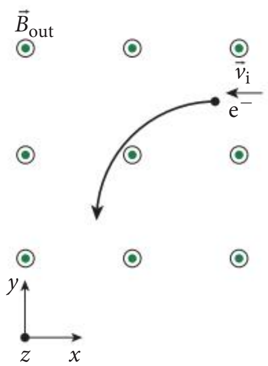
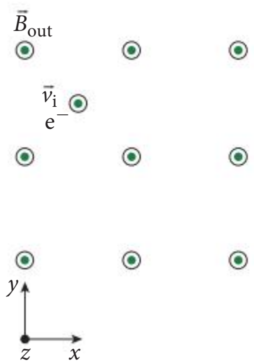
WORKED EXAMPLE 4.3

A magnetic field points in the positive *z* direction. Draw the path of an electron in the field with an initial velocity:

a in the positive *x* direction. (2 marks)

b in the negative *x* direction. (2 marks)

c in the positive *z* direction. (1 mark)

Answers	Logic	
a  ▲ **Figure 4.8** An electron entering the field in the positive *x* direction follows a curved path.	Use the right-hand rule, and remember that we reverse the direction of $q\vec{v}$ because it is a negatively charged particle. The field and velocity are perpendicular so the path is a section of a circle. Looking from above, the path is anticlockwise.	2 marks
b  ▲ **Figure 4.9** An electron entering the field in the negative *x* direction follows a curved path.	Use the right-hand rule, and remember that we reverse the direction of $q\vec{v}$ because it is a negatively charged particle. The field and velocity are perpendicular so the path is a section of a circle. Looking from above, the path is again anticlockwise.	2 marks
c  ▲ **Figure 4.10** An electron enters the field in the positive *z* direction. The path is a straight line.	The field and velocity are parallel, so the acceleration is zero and the path is a straight line.	1 mark

Try this yourself

How would the paths be different for a proton? (5 marks)

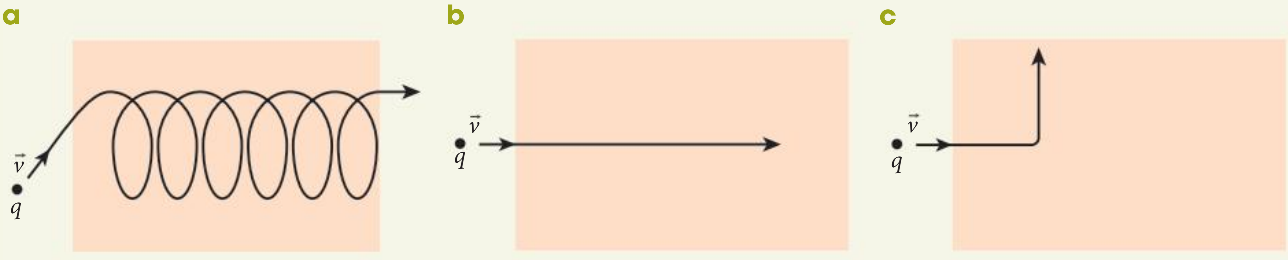
QUESTION SET 4.1

Remembering

1 Define 'magnetic flux density'. What are its units?

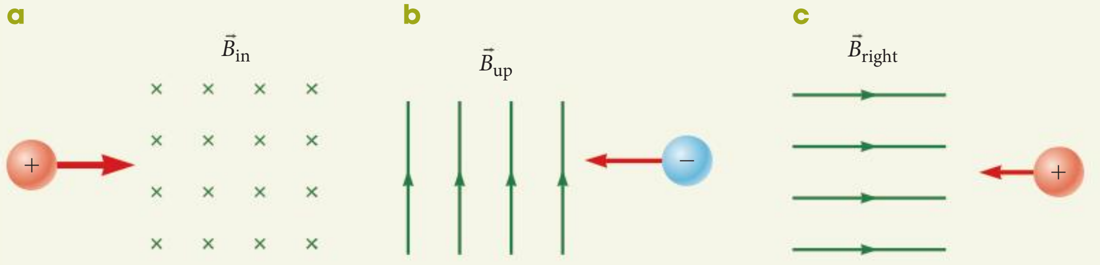
Understanding

2 Which of the possible paths shown in Figure 4.11 is not possible for a charged particle entering a region of uniform magnetic field?



▲ Figure 4.11
The shaded area is a region of uniform magnetic field. Which path is impossible?

3 Determine the initial direction of the deflection of the charged particles shown in Figure 4.12 as they enter the magnetic fields shown.



▲ Figure 4.12

Applying

4 A 1 m length of wire carrying 9 A is in a perpendicular magnetic field of 0.01 T. What is the force on the wire due to the magnetic field?

5 A 2.0 m length of wire carrying 5.0 A is at an angle of 30° to a magnetic field of 50 μT.

a What is the force on the wire due to the magnetic field?

b What is the size of the field that would give the same magnitude force if the wire was at an angle of 45° to the field?

6 What is the size of magnetic field necessary to balance the gravitational force on an electron moving horizontally north at 100 m s^{-1}? In which direction does this field need to be?

7 An electron is circulating inside the ring of a synchrotron with an orbital radius of 125 m. The electron has a velocity of 1.5×10^8 m s^{-1}. Ignoring relativistic effects, calculate:

a the period of the electron's orbit.

b the frequency with which it orbits.

c the magnetic field required to keep the electron in orbit.

Analysing

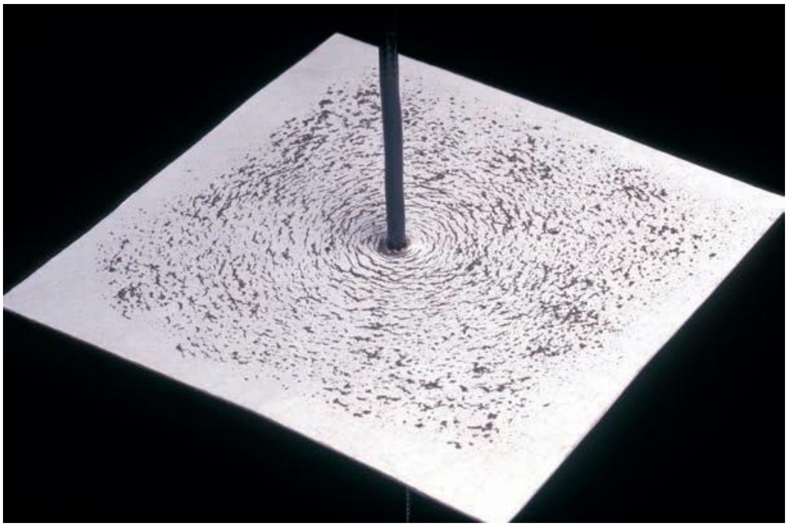
8 A proton and an electron enter a uniform magnetic field. The particles are travelling at the same speed perpendicular to the field. Draw a diagram showing their paths and explain the differences.

9 Mass spectrometers need to ionise molecules or atoms to be able to determine what they are. Explain why.

Reflecting

10 How has your understanding of magnetic fields and forces contributed to your ability to visualise three-dimensional effects in physics? What strategies or models have you used to help you visualise these effects? Compare your answer with those of others in your class.

Sources of magnetic fields



◀ **Figure 4.13** Iron filings show the magnetic field lines around a current-carrying wire.

Alamy/© sciencephotos

The relationship between magnetism and electricity was discovered in 1819 by Hans Christian Oersted. During a lecture demonstration he found that an electric current in a wire deflected a nearby compass needle. This showed that a magnetic field is created by the flow of charge, or current.

When iron filings or lots of compass needles are placed around a current-carrying wire, as in Figure 4.13, they form loops. These loops show the field lines due to the current.

Iron filings do this because they become magnetised, and point along the direction of the field. Compass needles are made from already magnetised iron or steel. A compass needle has a north and a south pole. The north pole is attracted to south poles, and repelled by other north poles. The south pole is attracted to north poles, and repelled by other south poles. Hence compass needles are useful for visualising magnetic fields. The circular nature of these magnetic field lines was first observed and published by Faraday.

Jean-Baptiste Biot and Felix Savart performed many experiments with magnets and current-carrying wires. They found that for a point P some distance from a long, straight current-carrying wire:

- the magnetic field is perpendicular to both the direction of the current and to a line between the wire and P.
- the magnitude of the field is inversely proportional to the distance from the wire to P, as shown in Figure 4.14(b).
- the magnitude of the field is proportional to the current, as shown in Figure 4.14(c).

These observations can be summarised mathematically as:

$$B = \frac{\mu_0 I}{2\pi r}$$

where B is the magnitude of the magnetic field at a distance r from a wire carrying current I. The constant μ_0 is called the permeability of free space.

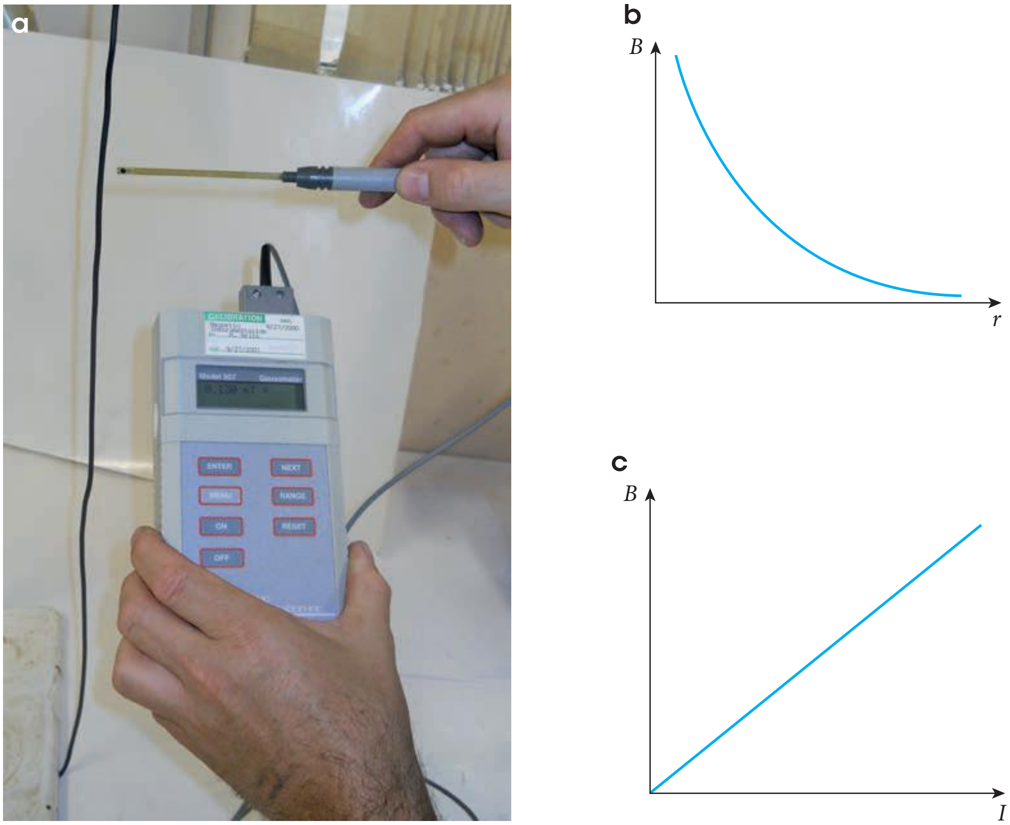
$$\mu_0 = 4\pi \times 10^{-7}\,\mathrm{T\,m\,A^{-1}}$$

In Nelson Physics Units 1 & 2 for the Australian Curriculum you studied light and saw that its speed in a medium depends on the electric and magnetic characteristics of the medium. These characteristics are the permittivity and permeability. In vacuum these are the constants ε_0 and μ_0.

The constant μ_0 has the same function for magnetic fields as the permittivity of free space, ε_0, has for electric fields. It tells us the strength of a field created by a given current in vacuum, and ensures the units of field are correct. As we shall see in the next chapter, the speed of light depends on these two constants: μ_0 and ε_0.

Remember that the field is a vector. We use the right-hand rule again (see Figure 4.3) to find the direction of the field.

Figure 4.14 ▶ a) Measuring field strength due to a current-carrying wire. b) Field strength, *B*, as a function of distance, *r*, from a wire. c) Field strength, *B*, as a function of current, *I*, carried by a wire.



Point sources create fields that vary with $\frac{1}{r^2}$. A line source, such as a current-carrying wire, creates a field that varies with $\frac{1}{r}$. Large flat sources create fields that, up close, do not vary with distance at all.

WORKED EXAMPLE 4.4

Calculate the magnetic field strength at a distance of 10 cm from a long wire carrying a current of 10 A. (3 marks)

Answer	Logic	
$B = \frac{\mu_0 I}{2\pi r}$	Relate the field to the current and distance.	1 mark
$B = \frac{(4\pi \times 10^{-7}\ \text{T m A}^{-1})(10\ \text{A})}{2\pi \times 0.1\ \text{m}}$	Substitute numbers with units.	1 mark
$B = 2 \times 10^{-5}\ \text{T}$	Calculate the final value.	1 mark

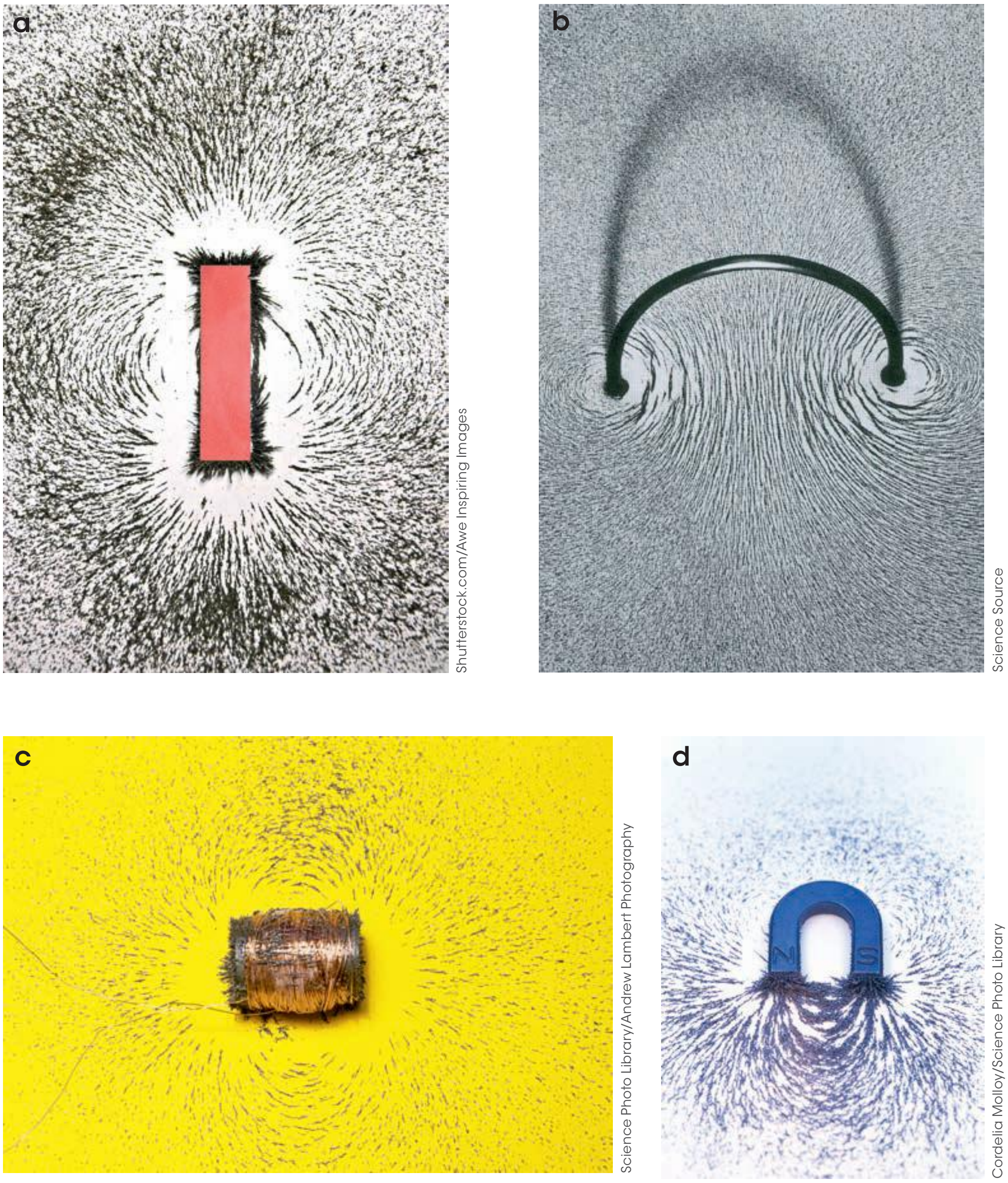
Try these yourself

1 Find the field at distances of 20 cm, 40 cm, 60 cm, 80 cm and 1 m from the wire. (5 marks)

2 Plot a graph of field as a function of distance from the wire. (2 marks)

Magnetic field lines

Magnetic field lines show the direction of force acting on a magnetic north pole, such as the north pole of a compass needle. Magnetic field lines are easy to visualise using iron filings, which act as tiny compass needles, as shown in Figure 4.13. Figure 4.15 shows the field lines for permanent magnets and arrangements of current-carrying wires. These diagrams allow you to see a slice through the magnetic field in the plane containing the iron filings. Remember that fields are actually three dimensional. You can use iron filings suspended in oil to see magnetic field lines in three dimensions.



▲ **Figure 4.15**
Iron filings can be used to show the magnetic field lines for a) a permanent bar magnet, b) a current-carrying loop of wire, c) a current-carrying coil or solenoid (note the dense, parallel field lines – iron filings – coming out of the left end of the solenoid) and d) a horseshoe magnet.

ACTIVITY 4.1

VISUALISING FIELD LINES IN THREE DIMENSIONS

Aim

To observe the magnetic field lines of a bar magnet in three dimensions

You will need

A bottle of baby oil, some iron filings, a test tube that fits snugly into the neck of the bottle and a bar magnet that fits inside the test tube.

What to do

1 Peel the label off the bottle. Tip a little of the oil out of the bottle so that the test tube fits in. Ask your teacher where to put this excess oil; do not pour it down the sink. Add some iron filings (about 15 mL is enough) to the oil in the bottle. Insert the test tube into the neck of the bottle. It should fit snugly so oil does not seep out around it. *Clean up any spills immediately!*

2 Now *carefully* slide the bar magnet into the tube. *Do not drop it in*, as the tube may break.

What did you discover?

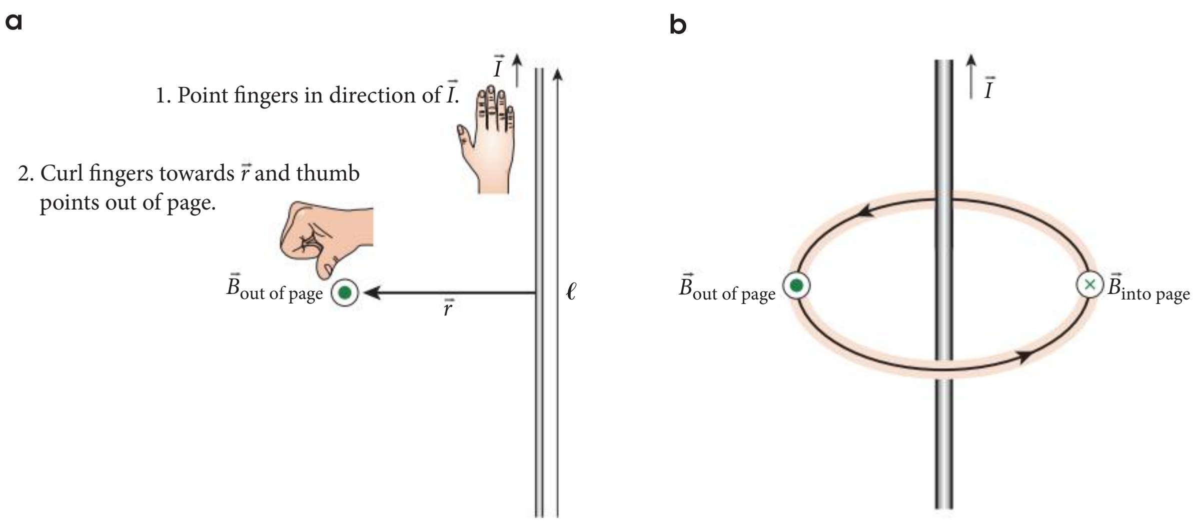
1 Draw a diagram or take a photograph showing how the iron filings have arranged themselves in the oil.

2 Try it with any other magnets that fit into the test tube, for example disc or ball magnets.

When you have finished, replace the lid on the baby oil bottle to avoid spills. Do not use the oil with filings in it on skin.

We saw in Figures 4.13 and 4.15 that current-carrying wires produce magnetic fields that form loops. We also know that field is a vector quantity, so the loops must have a direction. To find the direction we can use the right-hand rule again. This time the cross product is between the direction of current and the line joining the current to the point of interest, as shown in Figure 4.16(a).

Point your fingers in the direction of the current then curl them along the line joining the current to the point of interest. Your thumb now points in the direction of the field. You can do this at various points around the wire and join up the arrows to form loops.

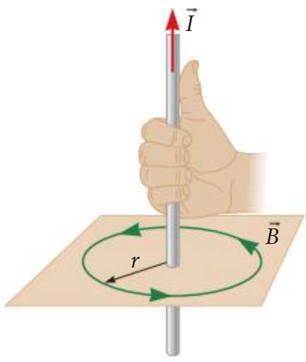


▲ **Figure 4.16**
Finding the direction of the field near a current-carrying wire using the right-hand rule. a) Use the right-hand rule to find the direction of the field at a point near the wire. b) Find the direction at other points and join the arrows to form loops.

You can see that this field looks very different from the electric field due to a single positive or negative charge. Rather than radiating out as the field lines from an electric charge do, the magnetic field due to a line of current forms loops. Magnetic field lines generally form closed loops.

A quick way to work out the direction of the magnetic field lines is to point your right thumb in the direction of the current. Your fingers then naturally curve in the direction of the field lines (see Figure 4.17). This rule is not different from the right-hand rule. It is just a quick way of applying it at all positions around the first of the vectors in the cross product.

Just as with electric field lines, there is an infinite number of possible magnetic field lines that you can draw. Choose a sensible number, and space them so that the density of field lines is proportional to the field strength. This means that for a current-carrying wire the distance between field lines should get bigger as you get further from the wire.



▲ **Figure 4.17**
Quick 'rule of thumb' for finding the direction of magnetic field lines: point your right thumb in the direction of the current and your fingers curl in the direction of the field lines.

> **Magnetic field lines show the direction of force acting on a north pole.**
> **The density of the field lines is an indication of the field strength.**
> **Currents produce field lines that form loops.**
> **The direction of the field is given by the right-hand rule.**

WORKED EXAMPLE 4.5

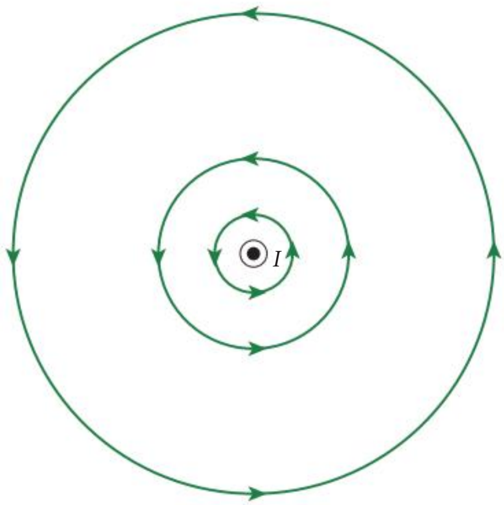
A long wire is carrying current directly upwards.

a Draw the magnetic field lines due to the current as seen from above. (2 marks)

b Draw arrows on your diagram showing the direction of the magnetic force on an electron to the north, south, east and west of the wire. The electron is moving upwards. (4 marks)

Answers

a



▲ **Figure 4.18**
Magnetic field lines as seen from above for a current coming out of the page

Logic

Use the right-hand rule to determine the direction of the loops. 2 marks

b

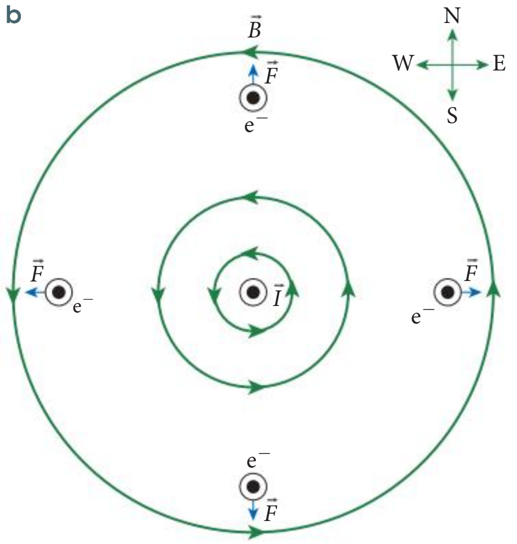


Figure 4.19 ▲
Force on an electron moving upwards out of the page

Use $F = qvB\sin\theta$ and the right-hand rule again to find the direction of the forces, remembering that q is negative in this case. 4 marks

Try this yourself

Repeat the question above with a proton instead of an electron. (5 marks)

Solenoids and electromagnets

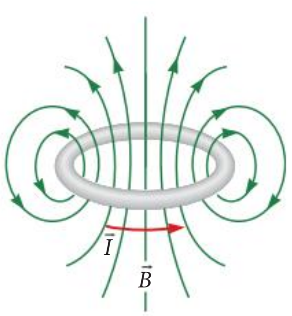
As we have seen, a current-carrying wire creates a magnetic field. If we want to create a large field, we use lots of wires. One way to do this is to coil a single wire into a **solenoid**, or coil.

Each loop of the wire creates a magnetic field, as shown in Figure 4.20. Inside the coil these fields add up to give a large and approximately uniform magnetic field. The more loops or turns of wire, the greater the field. In a tightly wound solenoid, the internal field lines are straight and parallel. Outside the coil the fields mostly cancel each other out. Hence there is a large, uniform internal field and a very small external field.

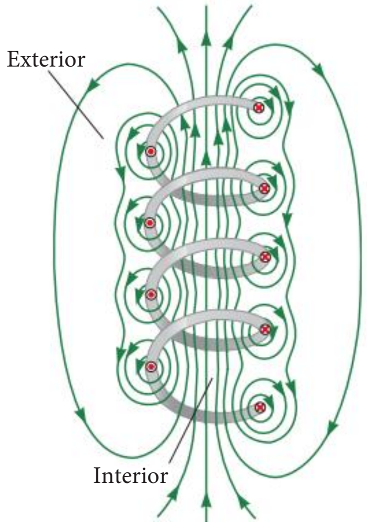
The result is an extremely useful device. Solenoids are used in transformers, electromagnets, magnetic switches, and many other applications. Inductors, which are common circuit elements, are small solenoids.

Figure 4.20 ▶
Magnetic field due to a) a single turn of wire and b) a solenoid

a



b



EXPERIMENT 4.1

A CURRENT BALANCE

A current balance is a current element that is placed in a magnetic field where it experiences a force. The current balance acts like a see-saw. The rotational effect of the magnetic force at one end is equal to the rotational effect of the gravitational force on the mass at the other end of the balance when the balance is in equilibrium.

Aims

To measure the magnetic force on a current balance

To find the magnetic field strength in a solenoid

Materials

Part A

- air core solenoid
- materials to make a current balance:
  - thin, stiff, lightweight cardboard or plastic
  - stiff conducting wire
  - copper or zinc sheet
  - pin
  - fine sandpaper
  - scissors
  - sticky tape
- short pieces of wire of known mass (the weights)
- 2 power supplies
- 2 ammeters (0–5 A)
- 2 variable resistors or rheostats
- 2 switches
- crocodile clips and leads

Part B

- current balance circuit comprising power supply, rheostat, ammeter, switch
- solenoid circuit comprising the other power supply, rheostat, ammeter, switch

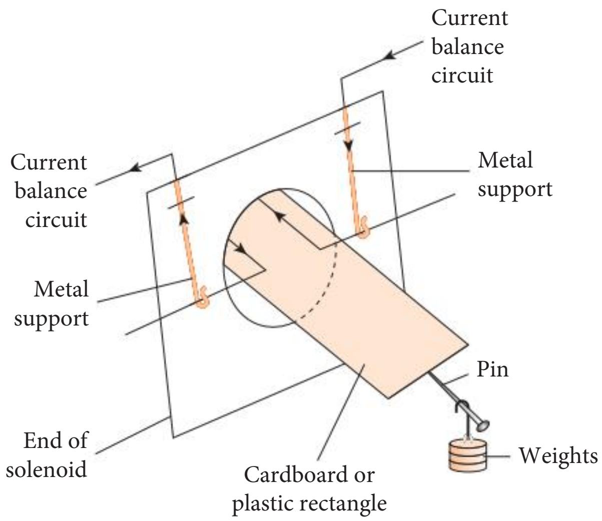
What are the risks in doing this experiment?	How can you manage these risks to stay safe?
Electricity can shock and cause damage to equipment.	Use low voltages and currents only.
Scissors can cut skin and have sharp tips.	Be very careful when using scissors. Do not run with scissors.

In your write-up, add any more risks you can think of, as well as ways to manage them.

Procedure

Part A

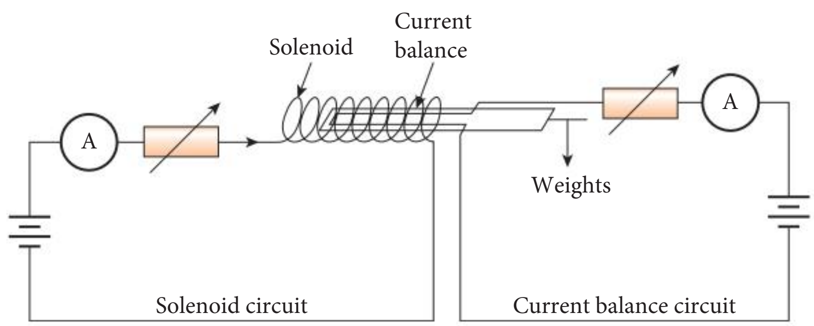
1. Cut a rectangle from the cardboard or plastic so that half will fit into the solenoid and half is outside.
2. Attach a small pin to the middle of one of the short sides of the rectangle so that it overhangs the end that will sit outside the solenoid (see Figure 4.21).
3. Make a rectangular half loop of conducting wire, to sit near the edges of the cardboard or plastic rectangle that goes into the solenoid. Make sure it is attached to the rectangle.
4. Cut two supports out of the metal sheet, bend them and attach them to the end of the solenoid as shown. Use the crocodile clips to connect them to the current balance circuit. Note that they should not make an electrical contact with the solenoid.
5. Bend the ends of the rectangular half loop so that they sit on the metal supports.
6. Use the sandpaper to clean the metal and ensure a good electrical connection.
7. Measure the length of the current element that is perpendicular to the magnetic field of the solenoid.
8. Balance the current balance by hanging pieces of wire or small weights over the pin.



◀ **Figure 4.21**
Metal supports attached to the end of the solenoid allow the current balance to rotate freely when current flows in the current balance circuit.

Part B

9 Connect the balance and solenoid circuits as shown in Figure 4.22.



◀ **Figure 4.22**
Circuits for current balance and solenoid

10 Close the switch in the solenoid circuit and adjust the current to 2.5 A–3.5 A.

11 Close both switches and observe what happens to the current balance – it needs to act as a seesaw, with the inner end being pushed down by the magnetic force. Make changes to the solenoid circuit if necessary.

12 Adjust the number of weights and their positions until the current balance is balanced.

Results

Record the following in a table. Don't forget to include units on all your measurements.

- Current in solenoid
- Current in current element
- Distance from pivot to current element
- Distance from pivot to balancing weights

Analysis of results

Calculate the following and add them to your table. Remember to include units.

- Masses of weights used
- Gravitational force on balancing masses
- Torque by weight force on current balance (torque is the product of force applied and the distance from the point of application to the pivot point)
- Magnetic force on current element
- Magnetic field in solenoid

The torque by the magnetic force is equal to the torque by the weight force when the current balance is balanced.

Discussion

1 Comment on the quality of your data and how this affects your results.

2 How could you improve the quality of your data?

ACTIVITY 4.2

MAKE YOUR OWN ELECTROMAGNET

Aim

To make an electromagnet

You will need

Several metres of fine, insulated wire, a small battery, a large steel nail, sticky tape and some paperclips.

Use a small battery, such as an AA size, or a low voltage power supply. *Do not* use a large battery such as a 12 V car battery; you could melt the wire and hurt yourself.

What to do

1. Check that your nail is not magnetised by trying to pick up a paperclip with it.
2. Wrap the wire tightly around the nail, always winding in the same direction. Leave enough wire free at each end to connect to your battery.
3. Sticky tape one end of the wire to each battery terminal.
4. Now try again to pick up a paperclip with your (electromagnet) tail.

What did you discover?

1. How many paperclips can you pick up with your electromagnet?
2. Does the electromagnet still work after you disconnect the battery?

Magnetic moments

The magnetic field of a bar magnet, shown in Figure 4.15a, is called a magnetic dipole field; it is caused by two **magnetic poles**. The magnetic field shown in Figure 4.20 is also called a magnetic dipole field because it has the same shape. We shall see in the section below, when we look at magnetic materials, that magnetic poles always come in pairs, or dipoles, consisting of a north pole and a south pole.

For any loop of current we can define a **magnetic moment,** $\vec{\mu}$. The magnitude of μ is:

$$\mu = IA$$

where I is the current and A is the area of the loop. μ has units of A m^2.

The magnetic moment is a measure of the strength of the magnetic field due to a current loop or a magnetic dipole. The larger the current, the greater the magnetic moment and the larger the field produced.

$\vec{\mu}$ is a vector, and its direction is given by a right-hand rule. Curl your fingers in the direction of the current, and your thumb points in the direction of $\vec{\mu}$. This is just the same as the rule of thumb for magnetic field lines shown in Figure 4.17. In this case, however, your fingers follow $\vec{I}$ (not $\vec{B}$) and your thumb points in the direction of $\vec{\mu}$ (not $\vec{I}$). The direction of $\vec{\mu}$ is that of the magnetic field, $\vec{B}$, at the centre of the loop. It points in the direction a compass needle would point if placed at this position. We can always define a magnetic moment for any magnet, including a bar magnet, from the torque that it experiences in a magnetic field. This is discussed in Chapter 5.

A pair of closely spaced positive and negative charges is an electric dipole (see Chapter 3). A pair of closely spaced north and south magnetic poles is a magnetic dipole. Both produce a characteristic dipole field and both experience a torque in a field.

The idea of a magnetic moment, which has a direction, is important in particle physics (see Chapters 9 and 10).

Magnetic materials

Currents create magnetic fields because of the moving charges that make up the current. But how does a fridge magnet or any other piece of magnetic material create a magnetic field? Where is the moving charge in this case?

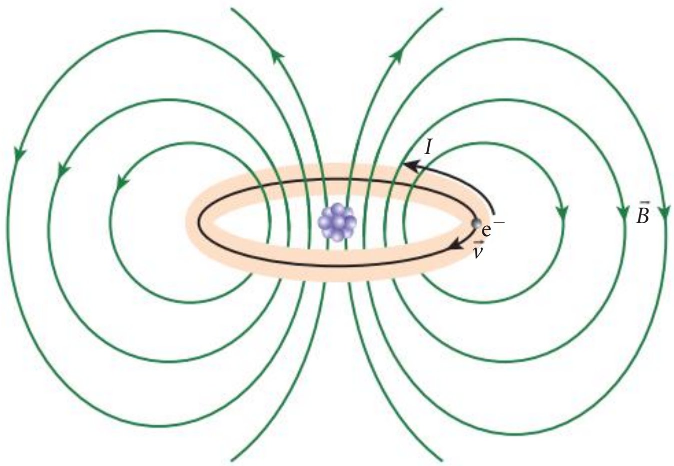
The answer is the electrons. Every electron in an atom creates its own tiny magnetic field as it orbits the nucleus, because it is acting as a tiny circular current. This circular current creates a dipole magnetic field like that shown in Figure 4.23. Hence the orbiting electron has a magnetic moment.

You know from chemistry that electrons orbit the nucleus in shells. This shell structure is important for understanding magnetic materials. Magnetism arises from unpaired electrons in the shells.

Some cosmological theories predict the existence of magnetic monopoles (an isolated north or south magnetic pole). However, they are not predicted by currently accepted models in particle physics, such as the Standard Model (see Chapter 10), nor have they ever been observed in nature.

Electrons also have a property known as **spin**. The electrons are probably not really spinning. But they do have their own tiny magnetic field even apart from the field due to their orbital motion. The spin of a particle is a measure of its magnetic moment, and hence is a measure of a particle's intrinsic magnetic field. We shall see in Chapter 8 that spin is quantised, both in magnitude and direction.

So electrons in materials have two magnetic moments, and hence two magnetic fields – one due to their orbital motion, and the other due to their spin.



▲ **Figure 4.23**
The electron orbiting the nucleus is a current loop, hence it has a magnetic field.

In most materials the magnetic fields from the electrons cancel out because they are pointing in all different directions. For a material to be magnetic it must have unpaired electrons in its shells *and* the directions of the magnetic moments (and hence fields) of many atoms in the material must line up in the same direction. There are very few materials that meet these requirements. The only common natural one is iron. Hence these types of materials are called **ferromagnetic** because they behave like iron.

Magnets are made of ferromagnetic materials that have been magnetised. If you stroke a steel pin with a strong magnet it will line up the magnetic moments in the iron. The pin then becomes magnetised and can be used as a compass needle. You can also use a solenoid, as in Activity 4.2, to magnetise a pin (or a nail).

The magnetic field lines due to a bar magnet are shown in Figure 4.24. Note that just as for the current-carrying wire, the field lines form loops. Magnetic field lines begin at a north pole and end at a south pole of a magnet. Any individual magnet is always a dipole with a north and a south pole. As Gilbert observed about 500 years ago, if you break the magnet you get two smaller magnets, each with a north and a south pole. This now makes sense to us because we understand that the poles are due to the movement of electrons. If you kept breaking your magnet you would eventually get to a single electron, still with spin, and creating its own magnetic field with the field forming loops. You cannot produce a magnetic monopole by breaking a bar magnet into small pieces; you will always get more dipoles.

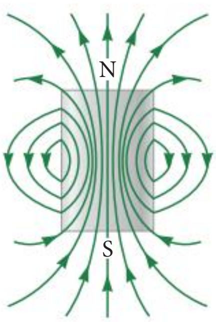


Figure 4.24 ▲
Magnetic field of a bar magnet

WOW

Quarks

We have seen that magnetic fields are created by moving charges, so it makes sense that electrons and protons have magnetic fields. But uncharged neutrons also have an intrinsic magnetic field! We cannot explain this if the neutron is a fundamental particle with no internal structure. The experimental observation of the neutron's magnetic field led to an important development in particle physics. The quark theory of matter says that particles such as protons and neutrons are actually made up of even smaller particles called 'quarks'. These quarks carry charges of $\pm\frac{1}{3}e$, or $\pm\frac{2}{3}e$. Protons and neutrons are each made of 3 quarks. The quark charges add up to zero for a neutron, but the neutron still has the charges inside creating a magnetic field.

Quarks are part of the Standard Model of particle physics, which you will study in Chapter 10.

Case study

Dr Jeff Tallon – high-temperature superconductors

Dr Jeff Tallon is Distinguished Scientist at Callaghan Innovation (a New Zealand Government-owned research institute) and was formerly Professor of Physics at Victoria University of Wellington. He has received the inaugural Prime Minister's Prize for Science (NZ) and he and his group are internationally known for their work in superconductivity. One of his group's most important discoveries was the high-temperature superconductor, $Bi_2Sr_2Ca_2Cu_3O_{10}$ or BSSCO. BSSCO, which is manufactured by their spin-off company HTS-110 Ltd, is now the most commercially important high-temperature superconductor in the world. It took 14 years of international patent battles for Jeff and his group to win the patent rights for BSSCO, the material whose structure he first jotted down on a greasy paper lunch bag.

Superconductors are materials whose resistance drops to zero below a critical temperature, T_c. The first superconductors discovered had critical temperatures of only a few kelvin. Then in 1986 the first high-temperature superconductor was discovered, $La_{2-x}Ba_xCuO_4$ with a T_c of 32 K. Since then many others have been discovered, with scientists racing to create materials with ever higher critical temperatures. BSSCO has a T_c of 105 K.

Superconductivity is a quantum mechanical effect that occurs when electrons of opposite spin pair up and form a wave that extends across the entire superconductor. Exactly why this happens in materials like BSCCO is not yet fully understood.

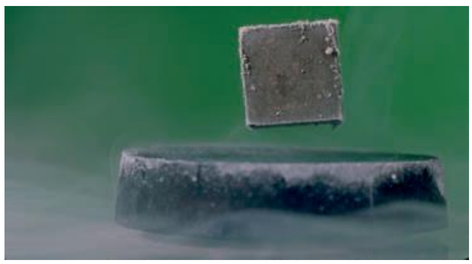
One of the other properties of these high-temperature superconductors is that they are perfect diamagnets. A diamagnet does the opposite to a ferromagnet in a magnetic field. Instead of the spins lining up with the magnetic field to give an even-larger magnetic field, they line up the opposite way. In a superconductor at temperatures below T_c, the induced field inside the material completely cancels out the applied field. This results in the Meissner effect, as shown in Figure 4.26. A magnet will levitate above the cooled superconductor. This effect is used in maglev trains in Japan. Superconducting magnets are also used to create the large magnetic fields used in MRI machines (see the Scientific Literacy box on page 117).

One of the difficulties of making superconductors into wires is that they are ceramics, and so are brittle. Unlike metals, superconductors cannot be stretched and bent. Dr Tallon and his team have devised a clever way of making wires using powdered BSSCO in silver tubes, which are then rolled and bundled to produce wires. These wires are used in various applications, such as generators (described in the next chapter).

The research into understanding how these high-temperature superconductor materials work is ongoing and relies on contributions from many scientists, and collaborations with groups around the world. Once we know how they work, scientists hope to design superconductors with higher critical temperatures and better mechanical properties. The applications will impact on health, transport, energy, communications, manufacturing and science.

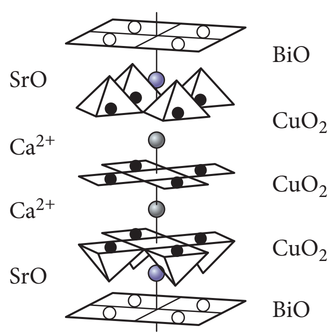
Courtesy of Dr Jeff Tallon

▲ **Figure 4.25**
Dr Jeff Tallon



Alamy/Phototake Inc.

▲ **Figure 4.26**
The Meissner effect. A magnet hovers over a disc of superconductor cooled with liquid nitrogen.



▲ **Figure 4.27**
The structure of BSSCO. Note the layers of different types of atoms.

Questions

1. Name two properties of high-temperature superconductors.
2. Define 'critical temperature of a superconductor'.
3. What is the difference between diamagnets and ferromagnets?
4. Why is it so difficult to make wires out of high-temperature superconductors?
5. Imagine that a room-temperature superconductor is discovered. Write a page (about 300 words) describing how *one* important technology would change. Consider the economic, social and environmental implications.
6. Why do you think the patent battle for BSSCO was so important that Dr Tallon and his colleagues kept fighting for 14 years? Can you think of reasons other than economic ones?

AURORAS IN REAL TIME

You can check forecasts for the aurora australis and aurora borealis here, and follow the links to viewing tips and more information.

MAGNET AND COMPASS

See the magnetic field lines due to a bar magnet and Earth. Move a compass around to see how the direction of the needle changes.

Earth's *B* field

Earth's magnetic field forms a shield that keeps out high-energy charged particles from space (mostly from our Sun). These particles would otherwise be extremely damaging to organisms on the surface of Earth. Fortunately for us, Earth's magnetic field deflects most of those particles past Earth. The few that leak in follow helical paths towards the poles. When these high-energy charged particles reach Earth's atmosphere they collide with air molecules, which are then ionised. The energy released when the atoms and molecules re-form are the auroras – the **aurora borealis** or northern lights and the **aurora australis** or southern lights (Figure 4.28).

Corbis/© Stefan Christmann

▲ **Figure 4.28** The aurora australis

Some other planets and the Sun also have a magnetic field. The origin of the Sun's magnetic field is convective currents, which you read about in Chapter 2 of *Nelson Physics Units 1 & 2 for the Australian Curriculum*. The origin of planetary magnetic fields is as yet not completely understood.

The configuration of Earth's magnetic field, shown in Figure 4.29, is very much like that due to a gigantic bar magnet deep in Earth's interior. However, the high temperatures in Earth's core, around 6000 K, prevent iron from retaining any permanent magnetisation. It is more likely that Earth's magnetic field is due to convection currents in Earth's core. Charged ions or electrons circulating in the liquid interior could produce a magnetic field of the form observed.

The magnetic poles are not quite at the geographic poles. The magnetic pole closest to the north geographic pole is near Hudson's Bay in Canada. The north pole of a compass needle points to this pole, so it is actually a *south* magnetic pole. The magnetic pole closest to the south geographic pole is in Antarctica, about 2000 km from the south geographic pole. It is a *north* magnetic pole. Compare the field lines for Earth in Figure 4.29 with those for the bar magnet in Figure 4.24 and you can see that what is usually labelled as the north magnetic pole on maps is really a magnetic south pole.

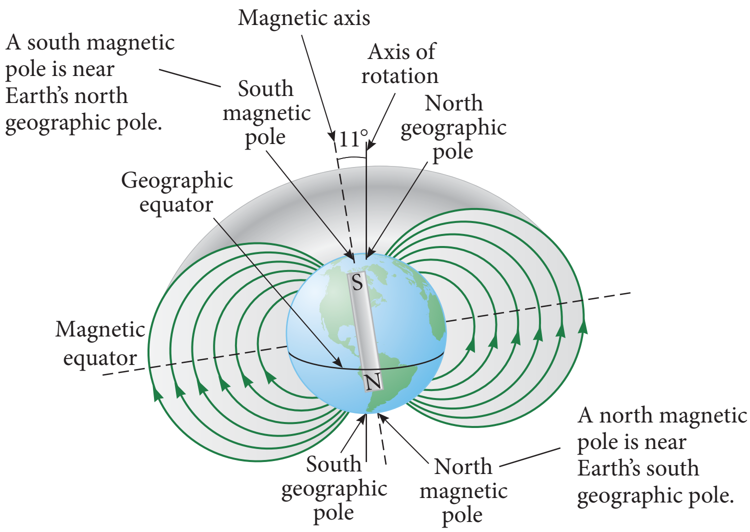


Figure 4.29
Earth's magnetic field

WOW

Earth's wandering magnetic poles

Earth's magnetic poles 'wander' on a daily and annual basis. Hence, the direction and magnitude of Earth's magnetic field changes in time. The direction of Earth's magnetic field has even reversed several times during the last million years. There is evidence for this in the magnetisation of iron in basalt rocks from volcanoes.

There is also variation on much shorter time scales. There are small daily variations due to the Sun's magnetic field. These variations are bigger in summer than in winter. Solar flares can result in variations that last hours to days. These can be observed in increased auroral activity.

There are larger variations on the scale of centuries. Recordings made in London since 1580 show that the direction of Earth's magnetic field there has varied by more than 30°. This variation seems to have a cycle of about 500 years. The reason for these variations is not yet understood.

Scientific literacy: magnetic resonance imaging (MRI)

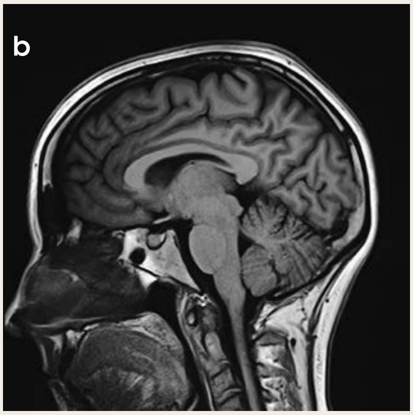
Magnetic resonance imaging or MRI (originally called nuclear magnetic resonance (NMR)) uses electromagnetic waves that match the resonance frequencies of the spins of the nuclei.

Nuclear magnetic resonance was discovered in the 1930s by Isidor Rabi, who won the Nobel Prize for Physics in 1944. Felix Bloch and Edward Purcell developed the technique in the 1940s for looking at chemical compounds and were awarded a Nobel Prize in 1952.

In the 1970s, physician Raymond Damadian, chemist Paul Lauterbur and physicist Peter Mansfied were all working independently on NMR.

Damadian discovered that the hydrogen NMR signal produced by cancerous tissue is different from that of healthy tissue. He applied for a patent for a machine to create images using NMR. His design had most of the basic components of today's MRI machines. Mansfield refined the technique and showed that images could be produced. Lauterbur produced the first actual NMR image.

Shutterstock.com/EPSTOCK



Courtesy of Kelly Robinson

Figure 4.30
a) A hospital MRI machine
b) An MRI of a brain

In 2003 Lauterbur and Mansfield shared the Nobel Prize for Medicine or Physiology for the development of MRI. Damadian was furious! He took out full page advertisements in newspapers including the *New York Times*, claiming that he should have been given a share of the prize.

This is how MRI works: The human body is almost two-thirds hydrogen atoms, and the nucleus of each is a single proton. A proton acts like a tiny bar magnet. In MRI a very large magnetic field, up to 3 T, is applied to the patient. This aligns the hydrogen spins with the field.

Then a pulse of electromagnetic waves is used to 'flip' some of the spins. The frequency of the electromagnetic waves has to be just right. The protons will only absorb energy from electromagnetic waves of the right frequency. This is why it is called *resonance*. The protons have quantised values of energy. They can move between energy levels only by absorbing or emitting light of the right frequency. You will see why in Chapter 8. This is a lot like atomic energy spectra, which arise because the electrons have discrete energies. In this case it is the protons that have discrete energies.

The spins that are no longer aligned with the field are in a higher energy state. They flip back, releasing energy as electromagnetic waves as they do so. This energy is detected and analysed, and the information is used to create images.

MRI gives better contrast between soft tissues than other techniques such as X-rays. This is very useful for imaging the brain, the heart and tumours. The resolution of MRI images is also very good, and can be used to investigate blood vessels. Much work has been done on the brain using MRI scans.

MRI is considered a very safe technique. Unlike X-rays and CT scans, it does not expose patients to ionising radiation. However, the large magnetic fields cannot be used on patients with metal implants such as pacemakers or most cochlear implants. The magnetic fields can also create eddy currents (see Chapter 5) and MRI machines are very noisy. Because the tube into which the patient is inserted is quite narrow, many people find it unpleasantly claustrophobic.

Sometimes patients are injected with a contrast agent. This helps make some tissues stand out more clearly. Contrast agents work by changing the magnetic field in the area into which they are injected. These agents usually contain gadolinium, which can be a problem for people with poor kidney function.

Questions

1 Draw a flowchart showing how the MRI technique works. You may need to conduct further research.

2 Describe two advantages of MRI over other medical imaging techniques.

3 Describe two risks associated with MRI. What can be done to minimise those risks?

4 Imagine you have been employed to produce an information leaflet for fMRI patients. Research fMRI and write a short summary (no more than 1 page or 300 words) explaining how the technique works.

5 MRI has better resolution than X-rays and does not use ionising radiation. Yet X-rays are still a much more common diagnostic tool. Compare the two techniques and justify why this is the case.

6 Imagine Dr Damadian has employed you to run a social media campaign to persuade the Nobel Prize committee to award him part of the 2003 prize. What techniques and arguments would you use to justify his case?

MAGNETIC RESONANCE IMAGING

You can find more information on the physics of MRI here.

Comparing the three field models: *g*, *E* and *B*

The gravitational field, g, was described in Chapter 2. The electric field, E, was described in Chapter 3.

We use field models to explain force at a distance. You have now seen the field models associated with three different forces: gravitational, electrostatic and magnetic. The models are similar in some ways and different in others.

Gravitational, electric and magnetic fields are all defined as field = force/property, where 'property' is the thing that creates the field and on which the field acts (Table 4.2).

Table 4.2 The three field models

Field	Created by	Acts on
Gravitational	Mass	Mass
Electric	Charge	Charge
Magnetic	Moving charge	Moving charge

Gravitational, electric and magnetic fields have several important differences.

- The gravitational and electric force vectors are in the direction of the field, but the magnetic force vector is perpendicular to the magnetic field.
- The electric force acts on a charged particle, and the gravitational force acts on a mass, regardless of whether the particle is moving, but the magnetic force acts on a charged particle only when the particle is in motion.
- The gravitational and electric forces can do work in displacing a charged particle, but the magnetic force associated with a steady magnetic field does no work when a particle is displaced by the field, because the force is perpendicular to the displacement.

These field models have been extremely successful in explaining and predicting phenomena. The electric and magnetic field models are the basis for electromagnetism, which we shall study in the next chapter. However, the field model has been unable to explain all the phenomena associated with the behaviour of charged particles such as electrons and protons. An alternative model, the exchange particle model, has been developed by quantum physicists to explain the interactions of fundamental particles. We will look at this newer model in Chapters 9 and 10. This newer model does not replace the field model any more than relativity replaces Newtonian mechanics – rather it complements it. Scientists work with many models and representations, and choose the one that works best for a given situation. The field model has been invaluable in helping us explain electric and magnetic phenomena, and is the basis of much of the technology that we take for granted. We shall look at some of this technology in Chapter 5.

QUESTION SET 4.2

Remembering

1 What is the source of magnetic fields?

2 How does magnetic field strength vary with distance from a long, straight current-carrying wire?

3 What are the units of the constant μ_0? Write these units in fundamental units.

Understanding

4 Why are most materials not magnetic? In what way are ferromagnets different?

5 Draw the magnetic field lines due to a wire carrying current directly downwards, as seen from above.

6 A current-carrying wire lies in a north–south line. A small magnetic compass needle is placed beneath the wire.

- a What happens to the compass needle when the current in the wire flows:
  - i from north to south?
  - ii from south to north?
- b If the compass was placed above the wire, what difference would it make to the direction in which the compass needle pointed?

Applying

7 What is the magnetic field a distance 1.0 cm from a wire carrying a current of 1.0 A?

8 At what distance from a wire carrying a current of 15 A is the field 1.0 μT?

9 What current is necessary in a long, straight wire to produce a magnetic field of 50 μT, approximately that due to Earth, at a distance of 1 cm?

Analysing

10 A wire carries a current of 10 A directly upwards.

- a Draw the magnetic field lines due to the current in the wire.
- b What is the direction of the force on an electron moving directly downwards, 5 cm from the wire? Draw the force on your diagram.
- c If the electron is moving at 1 km s^{-1}, what is the magnitude of the force acting on it?

CHAPTER SUMMARY

- Magnetic fields are created by moving charges and currents.
- Magnetic fields exert forces on moving charges and currents.
- The magnetic force is perpendicular to both the magnetic field and the direction of movement.
  $F = qvB\sin\theta$ for a single charge
  $F = I\ell B\sin\theta$ for a current
- The right-hand rule (Figure 4.3) is used to find the direction of the force.
- The magnetic field due to a current-carrying wire varies with the size of the current and decreases with the distance from the wire:
  $B = \frac{\mu_0 I}{2\pi r}$.
- Solenoids are used to create large, uniform magnetic fields.
- Magnetic field lines show the direction of force acting on a north pole.
- The density of the field lines is an indication of the field strength.
- The direction of the field is given by the right-hand rule.
- Magnetic field lines due to currents form loops.
- Magnetic materials always have a north pole and a south pole.
- Magnetic field lines start on north poles and end on south poles.
- Earth's magnetic field is similar to that of a bar magnet with a north magnetic pole near the south geographic pole.

CHAPTER GLOSSARY

aurora australis the southern lights; light produced by the collision of high-energy charged particles with air molecules close to the South Pole

aurora borealis the northern lights; light produced by the collision of high-energy charged particles with air molecules close to the North Pole

electromagnetic field model a combination of the electric field model and the magnetic field model, including the interaction between the two fields

ferromagnetic having magnetic properties like iron; able to be magnetised and retain the magnetisation so that the material is magnetic

magnetic field the field created by moving charges, including charges in magnetic materials, which exerts a force on moving charges and magnetic materials

magnetic flux density the magnitude of the magnetic field, measured in T

magnetic force the force that a magnetic field exerts on a moving charge or current

magnetic moment, μ also called the dipole moment, it is a vector of magnitude proportional to the magnetic field produced by a current loop, $\mu = IA$. The magnetic moment points in the direction of the field at the centre of the current loop

magnetic pole magnetic north or south pole, a point from which field lines come out or go in

magnetite an iron oxide, Fe_3O_4, that is a natural magnetic material

mass spectrometer a device that uses a magnetic field to characterise materials by the atoms and molecules they contain

permeability of free space, μ_0 the physical constant that determines the strength of the magnetic field produced by a current in vacuum. It has the value $4\pi \times 10^{-7}\,\text{T m A}^{-1}$

solenoid a coil of current-carrying wire that creates a large uniform field within the coil

spin in quantum theory, a property of particles, including electrons, that results in them having their own magnetic moment and hence magnetic field

synchrotron a machine that uses electric and magnetic fields to accelerate charged particles to large velocities while containing them in rings, to produce high-energy light

tesla the unit of magnetic field, $1\,\text{T} = 1\,\text{kg s}^{-1}\,\text{C}^{-1}$; it is named after Nikolai Tesla

vector cross product the vector cross product, $\vec{C} = \vec{A} \times \vec{B}$, gives a vector perpendicular to both $\vec{A}$ and $\vec{B}$ with magnitude $C = AB\sin\theta$, where θ is the angle between $\vec{A}$ and $\vec{B}$. The right-hand rule gives the direction of $\vec{C}$

CHAPTER REVIEW QUESTIONS

Remembering

1 How is magnetic field related to magnetic force?

2 List three ways in which magnetic fields and gravitational fields are different.

3 List three ways in which electric fields and magnetic fields are different.

Understanding

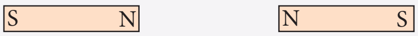
4 At what angle to a magnetic field can an electron and a neutron travel to have the same path?

5 Why does Earth's magnetic field not protect us from high-energy light (photons)? How does it protect us from the other forms of cosmic radiation?

6 Draw the magnetic field of a bar magnet, showing the north and south poles of the magnet. Add the vector $\vec{\mu}$, the magnetic moment of the bar magnet, to your diagram.

7 Figure 4.31 shows two bar magnets placed with their north ends together.

a Draw the magnetic field lines associated with this arrangement.



◀ Figure 4.31

b Are there any points in Figure 4.31 where the magnetic field is zero? If so, where?

8 Figure 4.32 shows a solenoid that carries a current in the direction as shown in the diagram.

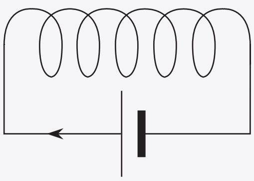
a Copy the diagram and carefully draw the magnetic field of the solenoid.

b What is the effect on the magnetic field within the solenoid of:

i increasing the current in the solenoid?

ii reversing the direction of the current in the solenoid?

iii increasing the number of turns of wire in the solenoid without changing the length of the solenoid?



◀ Figure 4.32

Applying

9 What is the minimum magnitude of a magnetic field necessary to apply a force of 1×10^{-12} N to an electron moving at a speed of 500 km s^{-1}?

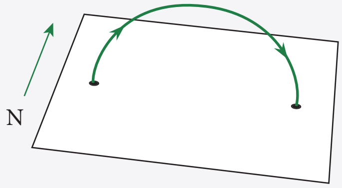
10 A 5 m long current-carrying wire is at an angle of 30° to a magnetic field. It carries a current of 30 A and experiences a force of 0.02 N. How large is the magnetic field?

11 a How large a current is necessary to produce a field of 0.15 T a distance of 1.0 cm from a wire?

b What is the field at a distance of 2.0 cm from this wire?

12 A wire 2.1 m long carrying a current of 0.85 A has a force of 5.0×10^{-2} N exerted on it by a uniform magnetic field at right angles to the wire. What is the magnitude of the magnetic field?

13 A small magnetic compass needle is placed at the centre of a single loop of wire that carries an electric current of 2 A, as shown in Figure 4.33. The loop has a radius of 2.0 cm. The plane of the coil is vertical and east–west. The magnitude of the magnetic field of the loop is much greater than Earth's magnetic field at this location.

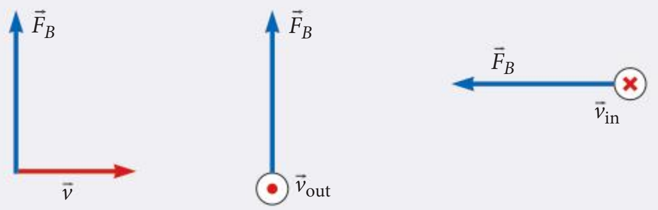


◀ Figure 4.33

a In which direction will the north pole of the compass needle point?
b The direction of the current in the wire is reversed. What is the direction in which the compass needle points now?
c Calculate the magnetic moment, μ, of the current loop. In which direction does it point?

Analysing

14 In Figure 4.34 a proton is moving in a magnetic field. The velocity of the proton and the direction of the magnetic force acting on it are shown. Find the direction of the magnetic field in each case.

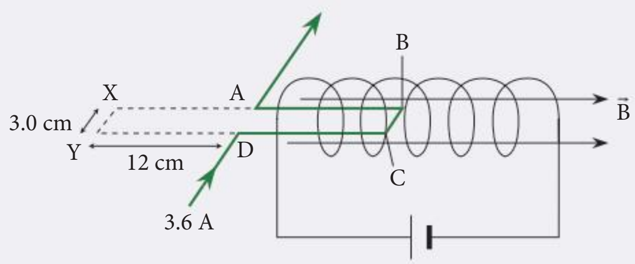


◀ Figure 4.34

15 A wire is carrying a large current directly upwards.
a Draw the magnetic field lines, as seen from above, due to this current.
b Consider a second vertical current-carrying wire close to the first wire. If the current in the second wire is also upwards, what is the direction of the force on this second wire? Draw it on your diagram.
c What is the direction of the force on the first wire due to the current in the second wire? Draw it on your diagram.
d How would your answers to parts (b) and (c) change if the current in the second wire was downwards?

16 How can the motion of a moving charged particle be used to distinguish between an electric field and a magnetic field? Give a specific example.

17 Figure 4.35 shows a current balance in which a loop of wire carrying a current of 3.6 A is balanced in the uniform field of a solenoid of field strength 0.20 T. The end of the loop BC has a length of 3.0 cm, while the length of side AB is 12.0 cm.



◀ Figure 4.35

a Find the magnitude and direction of the force on side:
i BC of the loop.
ii CD of the loop.
b What length of string of mass 6.5 g m^{-1} must be placed on the end XY of the loop to restore equilibrium?

Reflecting

18 Discuss the importance of Earth's magnetic field. How would life on Earth be different if Earth did not have a magnetic field? What technologies will be disrupted if the field changes direction again in the near future?

19 Draw a spider diagram to connect all the concepts and equations related to magnetic fields.

20 Compare and contrast electric and magnetic fields.

uttterstock.com/Benchaporn Maiwat

CHAPTER 5
ELECTROMAGNETISM

By the end of this chapter you will have covered the following material.

Science Understanding

- Magnets, magnetic materials, moving charges and current-carrying wires experience a force in a magnetic field; this force is utilised in DC electric motors (ACSPH108)
- A changing magnetic flux induces a potential difference; this process of electromagnetic induction is used in step-up and step-down transformers, DC and AC generators, and AC induction motors (ACSPH110)
- Conservation of energy, expressed as Lenz's Law of electromagnetic induction, is used to determine the direction of induced current (ACSPH111)
- Electromagnetic waves are transverse waves made up of mutually perpendicular, oscillating electric and magnetic fields (ACSPH112)
- Oscillating charges produce electromagnetic waves of the same frequency as the oscillation; electromagnetic waves cause charges to oscillate at the frequency of the wave (ACSPH113)

Introduction

Electromagnetism drives modern technologies such as large-scale electric power generators, transformers for computers and electric motors for public transport and transportation of metal ores and primary produce.

Alamy/age fotostock

Getty Images/Bloomberg

▲ **Figure 5.1** Electromagnetic induction at work. a) The Loy Yang power station in Victoria uses coal to power a generator to make electricity. b) A diesel electric freight train pulls trucks loaded with ore.

Many experiments in the 19th century showed that electricity and magnetism were intimately related. For example, Ampere performed experiments that showed that parallel currents attract and anti-parallel currents repel each other. The force on the wires is due to the magnetic fields created by the currents. Michael Faraday and Joseph Henry independently demonstrated **electromagnetic induction**. Electromagnetic induction is the production of an electric field by a changing magnetic field. Faraday was a brilliant experimentalist, and made many important contributions to physics and chemistry. However, his knowledge of mathematics did not extend beyond basic algebra.

James Clerk Maxwell, in contrast, showed talent in mathematics from an early age, and made many important contributions in theoretical physics. He took the observations made by Faraday and other physicists and constructed a mathematical model consisting of four differential equations. His model explained all the observed electromagnetic phenomena *and* predicted the existence and behaviour of **electromagnetic waves**. The central idea in Maxwell's model is that there is a symmetry between electric and magnetic fields.

JAMES CLERK MAXWELL: BIOGRAPHY

Read this short biography of this Scottish scientist.

HERTZ'S SPARK GAP TRANSMITTER

Hertz provided evidence for Maxwell's theory of electromagnetic waves.

MICHAEL FARADAY'S BIOGRAPHY

Read the biography of Michael Faraday.

- **An electric field that changes in time creates a magnetic field**

and

- **a magnetic field that changes in time creates an electric field.**

The development of the theory of electromagnetism demonstrates the need for both experiment and theory. Without experiments, existing theories could not be tested and falsified, and new theories would not be developed. The development of theory, in turn, leads to new technologies that allow further and more precise experiments. Maxwell's electromagnetic theory unified a set of ideas about electricity and magnetism and led us to an understanding of how the two fields are related. They are, in fact, two aspects of a single thing. The symmetry between electricity and magnetism was also what prompted Einstein to develop his theory of special relativity, which is described in the next chapter.

This new field model for electromagnetism led to the development of important technologies, including **motors** and **generators**. Generators, mainly powered by coal, are used to produce almost all of our electricity in Australia. Motors are used to run much of our transport including electric trains and trams in the cities and the enormous diesel electric locomotives that move commodities such as iron ore, wool and wheat around the country. We shall look at some of these technologies in this chapter.

WOW

Michael Faraday (1791–1867)

Faraday rose from humble beginnings to work with and lead some of the best scientists of his time. He was a brilliant experimentalist, and made many important contributions to physics and chemistry. Faraday investigated electromagnetic induction and introduced the concept of 'line of force', which we now call 'fields'. His knowledge of mathematics did not extend beyond basic algebra.

Science Photo Library/Chemical Heritage Foundation

Figure 5.2 ▶ Michael Faraday

Currents from fields: electromagnetic induction

Electromagnetic induction is the production of an electric field by a time-varying magnetic field. An electric field in a region means that there is a potential difference between points in that region. If there is a potential difference and free charge carriers, then a current will be generated. This is the basis of electromagnetic induction, as used in electricity generation. To understand how the process works, we first need to look at the idea of **magnetic flux**.

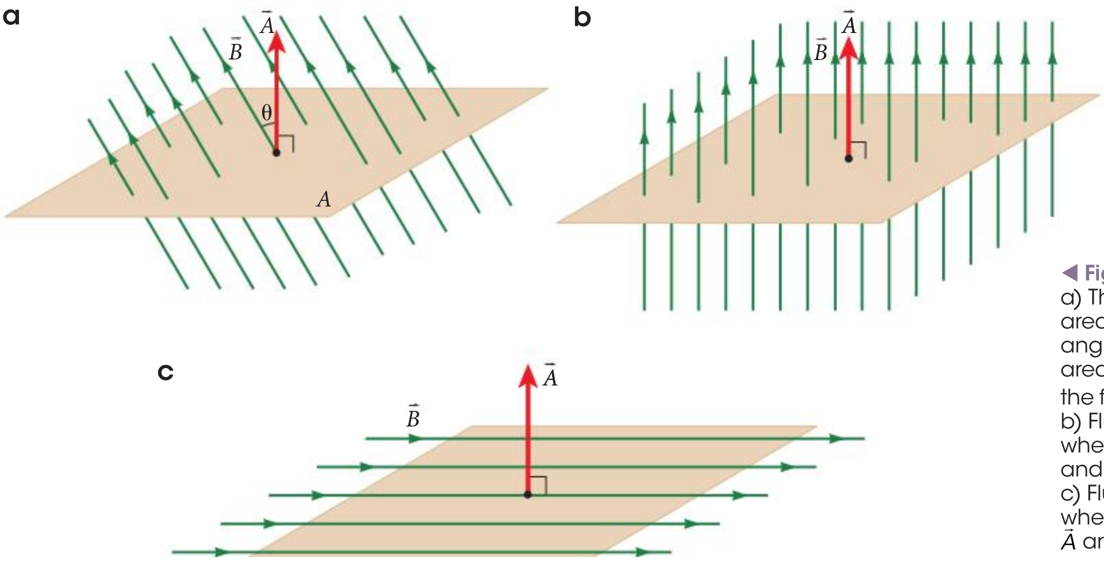
Recall what you learnt about circuits in Nelson Physics Units 1 & 2 for the Australian Curriculum, Chapter 6 .

Magnetic flux

In the previous chapter we noted that magnetic field strength is also called 'magnetic flux density'. This comes from the definition of magnetic flux. Consider a uniform magnetic field passing through an area, A, as shown in Figure 5.3. We choose the direction of the vector $\vec{A}$ to be perpendicular to the area, as this gives us a unique way of representing the area. The magnitude of the vector $\vec{A}$ is the area A, in units of m^2. The number of $\vec{B}$ field lines crossing the area depends on the directions of $\vec{B}$ and $\vec{A}$. Magnetic flux, Φ, is the magnetic field crossing through some area, multiplied by the size of the area:

$$\Phi = BA\cos\theta$$

The magnetic flux, Φ, has units of $T\,m^2$ or the weber, Wb, after Wilhelm Weber.



◀ **Figure 5.3**
a) The flux through the area A depends on the angle, θ, between the area vector $\vec{A}$ and the field, $\vec{B}$. $\Phi = \vec{B}\vec{A}\cos\theta$.
b) Flux is a maximum when θ is zero and $\vec{B}$ and $\vec{A}$ are parallel.
c) Flux is a minimum when $\theta = 90°$ and $\vec{B}$ and $\vec{A}$ are perpendicular.

The flux, Φ, has maximum amplitude when the field is in the same direction as, or opposite to, the vector $\vec{A}$; when $\vec{B}$ is perpendicular to the surface of the area. Φ is zero when $\vec{B}$ is parallel to the surface of the area.

WOW

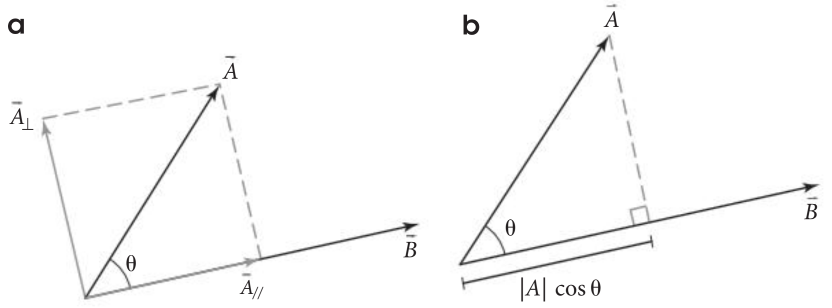
Vector dot product

The flux is more correctly written as the **vector dot product**, $\Phi = \vec{B} \cdot \vec{A}$.

The vector dot product, $C = \vec{A} \cdot \vec{B}$, is used often in physics. $\vec{A}$ and $\vec{B}$ are vectors, but the dot product $C = \vec{A} \cdot \vec{B}$ is a scalar, with magnitude $C = AB\cos\theta$. The dot product allows us to combine two vectors to produce a scalar.

You have met the dot product before (although it may not have been called that) when you studied forces and energy. The work done by a force is given by $W = \vec{F} \cdot \vec{s}$. Both force and displacement are vectors, but the energy transferred (the work done) is a scalar.

Geometrically, the magnitude C is the projection of vector $\vec{A}$ on to vector $\vec{B}$, multiplied by the magnitude of $\vec{B}$. When $\vec{A}$ and $\vec{B}$ are parallel, C has its maximum value. When $\vec{A}$ and $\vec{B}$ are perpendicular, then C is zero.



▲ **Figure 5.4** The dot product of two vectors, $C = \vec{A} \cdot \vec{B}$. a) Vector decomposition of A into components parallel and perpendicular to B. b) Projection of $\vec{A}$ on to $\vec{B}$.

VECTOR DOT PRODUCT

See how $C = \vec{A} \cdot \vec{B}$ varies.

WORKED EXAMPLE 5.1

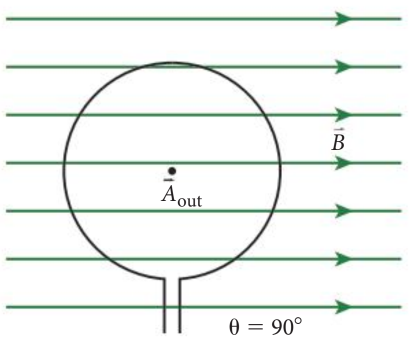
A loop of cross-sectional area $0.050\,m^2$ is in a uniform magnetic field of magnitude 0.24 T.

a Draw diagrams showing the loop and field and identifying the angle, θ, between the area vector $\vec{A}$ and the field $\vec{B}$, when the flux is a:

- **i** minimum. (2 marks)
- **ii** maximum. (2 marks)

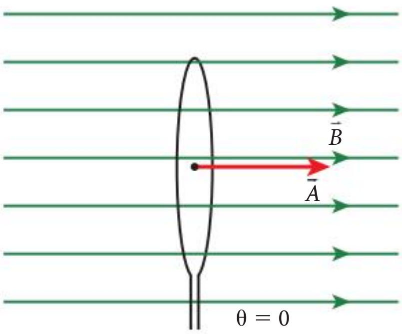
b Find the maximum and minimum values of the flux through the loop. (5 marks)

Answers	Logic	
a i	Identify that flux has a minimum value of $\Phi = BA\cos\theta = 0$ when $\theta = 90°$ and that this is when the loop is parallel to the field.	2 marks



▲ **Figure 5.5** Loop is parallel to field so $\theta = 90°$ and flux is zero.

ii



▲**Figure 5.6** Loop is perpendicular to field so $\theta = 0°$ and flux is a maximum.

Identify that flux has a maximum value of $\Phi = BA$ when $\theta = 0°$ and that this is when the loop is perpendicular to the field. 2 marks

b $\Phi = BA\cos\theta$	Relate flux to field and area.	1 mark
$\Phi_{max} = BA\cos(0) = BA$	Identify the maximum value.	1 mark
$\Phi_{max} = 0.24\,\text{T} \times 0.050\,\text{m}^2$	Substitute numbers including units.	1 mark
$\Phi_{max} = 1.2 \times 10^{-2}\,\text{Wb}$	Calculate the final value.	1 mark
$\Phi_{min} = BA\cos(90°) = 0$	Identify the minimum value.	1 mark

Try this yourself

Find the angle for which the flux is half its maximum value. (4 marks)

Induced emf

When the magnetic flux changes with time, an electric field is induced. Imagine a loop in the field. The induced electric field produces an **induced emf** around the loop.

The magnitude of the induced emf is given by Faraday's law:

$$\varepsilon = \frac{-(\Phi_f - \Phi_i)}{\Delta t} = \frac{-\Delta\Phi}{\Delta t}$$

where Φ_f is the final flux and Φ_i is the initial flux; ε has units of $\text{T m}^2\,\text{s}^{-1}$, which is the same as volt, V. The negative sign indicates that the induced emf *opposes* the change in flux.

Looking back at the expression for flux, we can see that:

$$\varepsilon = \frac{-\Delta(BA\cos\theta)}{\Delta t}$$

Inspecting this equation we can see that there are three ways to induce an emf:

- Change the magnetic field.
- Change the area, A.
- Change the angle, θ, between the area and the field.

In practice, it is usually either the magnetic field or the angle θ that is varied. For example, a coil connected to an **alternating current (AC)** produces a time-varying magnetic field. These are used in **transformers** and motors.

When a loop or coil of wire with area A is placed in a field, the flux through the loop can be varied by spinning the loop. This changes the angle, and induces an emf in the loop. The same effect can be achieved by spinning a magnet near the loop. In both cases the flux varies in time. This is used in generators.

We can take any parameter kept constant out of the brackets. For example, if area and angle are kept constant while B is varied, we write:

$$\varepsilon = \frac{-A\cos\theta\,\Delta(B)}{\Delta t} = \frac{-A\cos\theta\,(B_{\mathrm{f}} - B_{\mathrm{i}})}{\Delta t}$$

If the area and field are held constant but the angle is changed we write:

$$\varepsilon = \frac{-BA\Delta(\cos\theta)}{\Delta t} = \frac{-BA(\cos\theta_{\mathrm{f}} - \cos\theta_{\mathrm{i}})}{\Delta t}$$

The loops in a coil act like batteries connected in series (see Nelson Physics Units 1 & 2 for the Australian Curriculum Chapter 5).

To generate a larger emf, a coil containing multiple loops of wire is used. Each loop will have an emf induced between its ends, so connecting n loops in series is like connecting n batteries in series. Simply add the emf in all loops. Thus:

$$\varepsilon = \frac{-n\Delta\Phi}{\Delta t} = \frac{-n\Delta(BA\cos\theta)}{\Delta t}$$

Once an emf is induced, a current will flow if there are free charge carriers and a path for them to flow along. This is usually achieved by putting a metal coil in the field. This **induced current** is related to the emf by Ohm's law. As you saw in *Chapter 5 of Nelson Physics Units 1 & 2 for the Australian Curriculum*, $i = \frac{\text{emf}}{R}$, where R is the resistance of the path. We are using the symbol i for current, rather than I, as the current may vary with time.

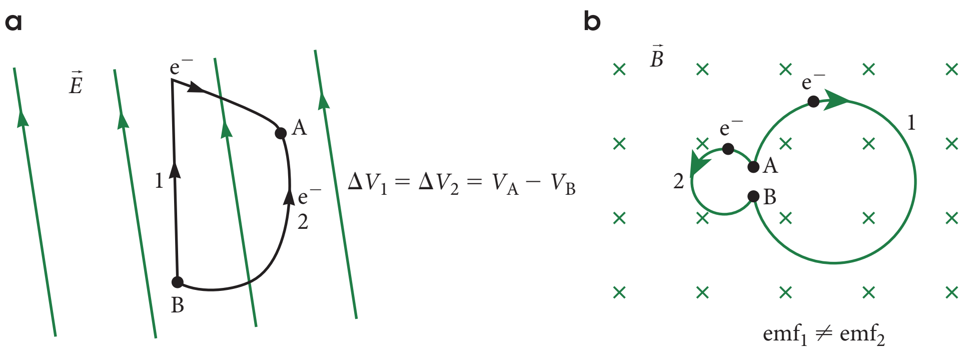
We use the term 'emf' here rather than potential difference for a reason. They are often treated as if they are the same thing, but they are not. The difference is explained in the Wow box.

WOW

emf

Electromotive force (emf) and potential difference often do the same job, which is to enable a current to flow. They have the same unit, volt (V), but they are different. Potential difference is the unique difference in potential energy per unit charge *between any two points* in an electric field.

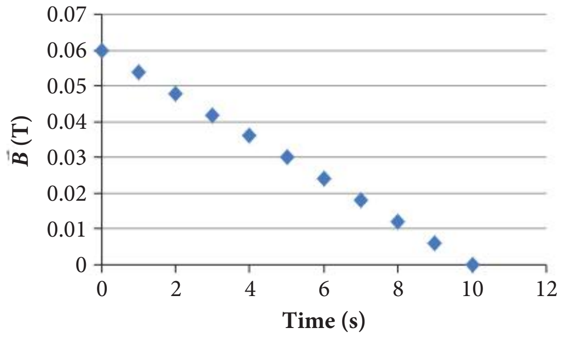
The emf is the energy per unit charge available to a charged particle. The induced emf between any two points in a changing magnetic field *does not* have a unique value, but depends on the path between the two points. This is because it depends on the flux enclosed. Different closed paths between two points may contain different fluxes.



▲ **Figure 5.7** a) Potential difference: an electron moves from point A to point B in an electric field. The change in potential energy of the electron is the same regardless of path taken. b) Emf: an electron passes through two loops in a changing magnetic field. The emf measured across the two loops is different because they contain different magnetic fluxes, even though they have the same beginning and end points.

WORKED EXAMPLE 5.2

A wire loop of cross-sectional area 0.050 m^2 is in a magnetic field. The loop is perpendicular to the field. The field changes with time as shown.



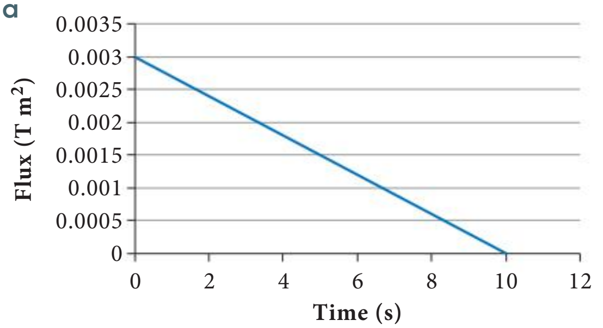
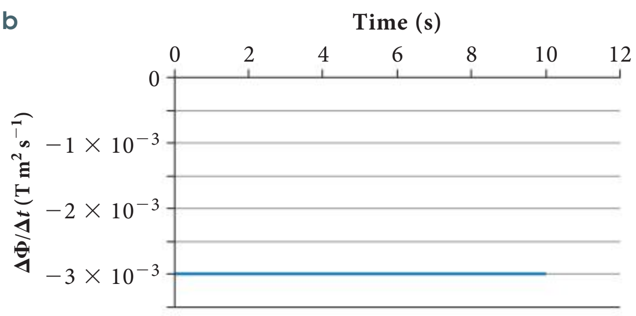
▲ **Figure 5.8** Magnetic field as a function of time

a Sketch a graph of flux through the loop as a function of time. (2 marks)

b Sketch a graph of $\frac{\Delta\Phi}{\Delta t}$ as a function of time. (2 marks)

c Find the emf induced between the ends of the wire. (1 mark)

d Find the current induced in the loop when the loop has a resistance of 0.15 Ω. (3 marks)

Answers	Logic	
a  ▲ **Figure 5.9** Flux as a function of time	Use $\Phi = BA\cos\theta$, noting that $\theta = 0$ and $A = 0.05\,m^2$ is given in the question.	2 marks
b  ▲ **Figure 5.10** $\frac{\Delta\Phi}{\Delta t}$ as a function of time	$\frac{\Delta\Phi}{\Delta t}$ is the gradient of the $\Phi(t)$ graph, which is constant. We find this gradient by taking the rise over run for a section of the $\Phi(t)$ graph.	2 marks
c +0.3 mV	The emf is the negative of the gradient of our $\Phi(t)$ graph, which we can see above is −0.3 mV; hence the emf is +0.3 mV.	1 mark

d $i = \frac{\varepsilon}{R}$	Use Ohm's law to relate emf to current.	1 mark
$i = \frac{-3.0 \times 10^{-4}\,\text{V}}{0.15\,\Omega}$	Substitute numbers including units.	1 mark
$i = -0.002\,\text{A}$ or $-2\,\text{mA}$	Calculate the final value.	1 mark

Try these yourself

1 What current would be produced if a five-loop coil was used instead? (5 marks)

2 How quickly would the field have to drop to zero to produce an induced current of 0.1 A? (6 marks)

ACTIVITY 5.1

ELECTROMAGNETIC INDUCTION

Aim

To investigate the current produced in a coil by a changing magnetic field

You will need

- bar magnet
- coil
- sensitive ammeter or centre-zero galvanometer

What to do

1. Connect the coil to the ammeter.
2. Slowly move the north pole of the magnet into one end of the coil and pull it out again. Note what happens to the ammeter or galvanometer. Now do it faster.
3. Predict what you will see if you put the south pole of the magnet into the coil. When you have written down your prediction, do the experiment.

What did you discover?

1. Did your observations match your predictions? Draw a sketch of the current as a function of time, noting what was happening with the magnet and coil along the time axis.
2. What can you say about the direction of the current?

LENZ'S LAW

View the animations of Lenz's law, magnetic braking and metal detectors.

Lenz's law

The negative sign in the equation for induced emf tells us about the direction of the induced current.

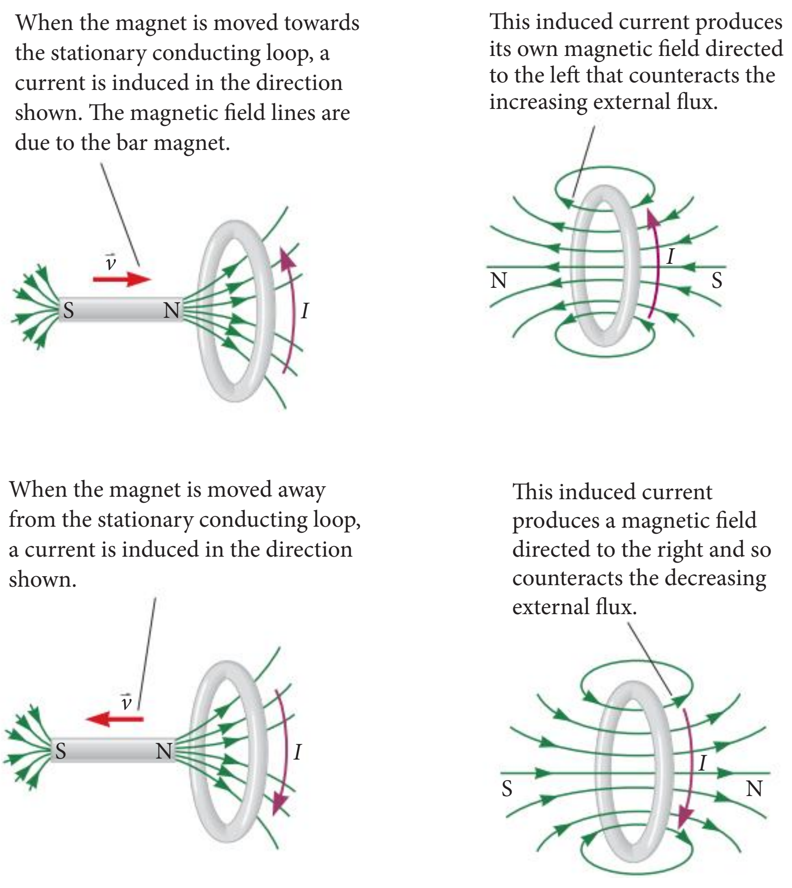
Consider a loop in a magnetic field that is getting stronger with time, as shown in Figure 5.11. There are two possible directions in which the induced current can flow. How do we know which way it will go? The negative sign tells us that the current must flow such that the flux through the loop decreases. Why is this? The short answer is: conservation of energy.

The potential energy of the changing magnetic field is transformed into electric potential energy. The result is an electric field and, as we saw in Chapter 3, an electric field can do work by applying a force. Work is done on any free electrons by the induced electric field. The electrons then flow, giving the induced current. The induced current takes its energy from the changing magnetic flux (via the electric field), and so reduces the rate at which the flux changes.

Consider what would happen if the current acted to produce a further increase in the magnetic flux through the loop. The flux would increase more, giving a bigger induced current, giving a bigger flux and so on. We would have a ‘perpetual motion machine’ that made more and more current without any source of energy! This would violate conservation of energy and cannot happen. So the current *must* flow in the other direction and act to decrease the magnetic flux through the loop. Lenz’s law is essentially a statement of *conservation of energy*.

You studied conservation of energy in Nelson Physics Units 1 & 2 for the Australian Curriculum, Chapter 1.

Lenz’s law:
An induced emf acts to produce an induced current. The induced current is in the direction that causes a magnetic flux change that opposes the change in flux which induced the emf.



◀ Figure 5.11
A moving bar magnet induces a current in a conducting loop. The direction of the current is determined by Lenz’s law.

Eddy currents

Induced currents are not only seen in wire loops, but in any material in which there are free charge carriers. If a magnet is moved around over a piece of metal, the changing magnetic field will induce **eddy currents** in the metal. The electrons move in circles in the region where the field is changing. They form loops and spirals of current, like eddies in a cup of tea when you stir it. These eddy currents create magnetic fields that oppose the changing flux from the moving magnet. They act to slow down or brake the magnet. This is called **magnetic braking**.

ACTIVITY 5.2

MAGNETIC BRAKING

Aim

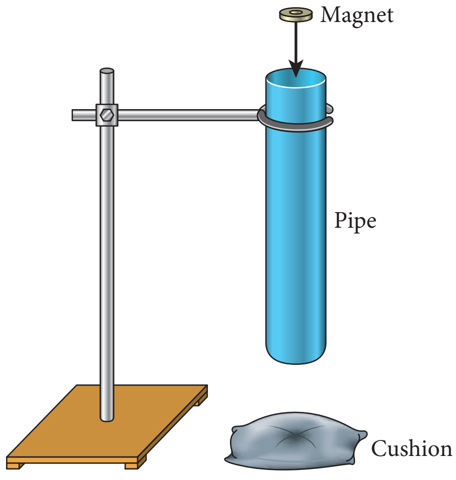
To investigate magnetic braking

You will need

- plastic pipe
- copper pipe of the same length
- retort stand and clamps
- strong magnet
- stopwatch
- something to cushion the magnet's fall

What to do

1 Attach the plastic pipe to the retort stand so it stands vertically. Drop the magnet down the pipe and time how it long it takes to fall through the pipe.
2 Try dropping it north pole first and then south pole first. Does it make any difference?



◀ **Figure 5.12** Experimental set-up

3 Replace the plastic pipe with the copper pipe.
4 Can you predict in advance what will happen? Will the magnet fall faster or slower than in the plastic pipe? Will it fall faster with the north pole down than with the south pole down? Write down your predictions and your reasoning.
5 Now make some more measurements using the copper pipe.

What did you discover?

1 Did your observations match your predictions? How can you explain what you observed?

An interesting extension to this activity is to repeat the measurements using a copper pipe with a slit along its length. Why do you think this might give a different result to a complete copper pipe?

QUESTION SET 5.1

Remembering

1 What are the units of magnetic flux?

2 Define 'electromagnetic induction'.

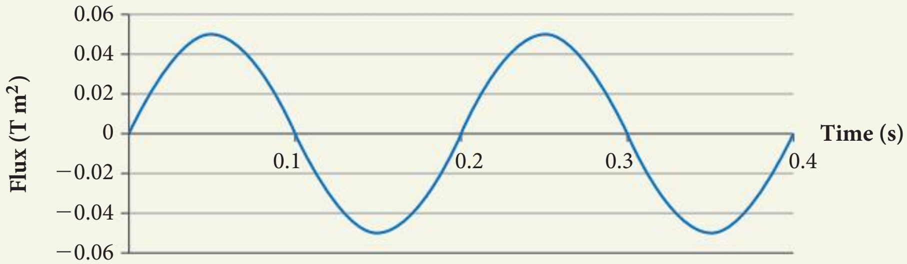
Understanding

3 A loop is placed in a uniform magnetic field and moved in a straight line. In this case, no current is induced in the loop. Why?

4 Figure 5.13 shows the flux through a loop as a function of time.

a Sketch the rate of change of flux through the loop $\frac{\Delta\Phi}{\Delta t}$. Use the same time scale.

b Sketch the induced emf across the loop as a function of time. Use the same time scale.



◀Figure 5.13

5 Draw a flowchart that summarises Lenz's law for a magnet being pushed into a solenoid (coil).

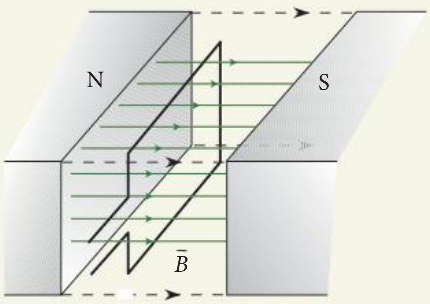
Applying

6 A solenoid produces a magnetic field of 0.25 T in its interior. The field is approximately uniform. What is the radius of the coil, given that the flux through any loop of the solenoid is 5.0 mWb?

7 A loop of cross-sectional area 0.050 m^2 with resistance 1.0 Ω is in a variable magnetic field. At what average rate does the field need to change to induce a current of 0.050 A in the loop? Use graphs to explain your answer.

Analysing

8 A loop of cross-sectional area 0.015 m^2 is in a magnetic field of 0.030 T. Initially the loop is perpendicular to the field lines (Figure 5.14). The loop is rotated about an axis parallel to its long sides at a uniform angular velocity of 5 revolutions per second.



◀Figure 5.14

a What is the flux through the loop at $t = 0$ s?

b Draw a graph of the flux through the loop as a function of time. Mark important features on your graph including the maximum flux and the period.

c Write an equation to describe the flux as a function of time.

d On the same axes draw a graph of the induced emf across the loop as a function of time.

Reflecting

9 Follow the weblinks on page 124 about Maxwell and Faraday. Read more about them. Compare and contrast their contributions to electromagnetism. Whose contribution do you think was more significant? Justify your answer.

Applications of electromagnetism

Many physicists find the symmetry between electric and magnetic fields very beautiful. It is these underlying symmetries that help inspire physicists to find patterns in nature, leading to a better understanding of how the universe works.

This symmetry is not only beautiful, it is incredibly useful! In this section we will look at some of the applications of electromagnetism. Most of Australia's electricity power production relies on generators. Most appliances have a plug pack or built-in transformer. Blenders, hairdryers and sound speakers are electric motors.

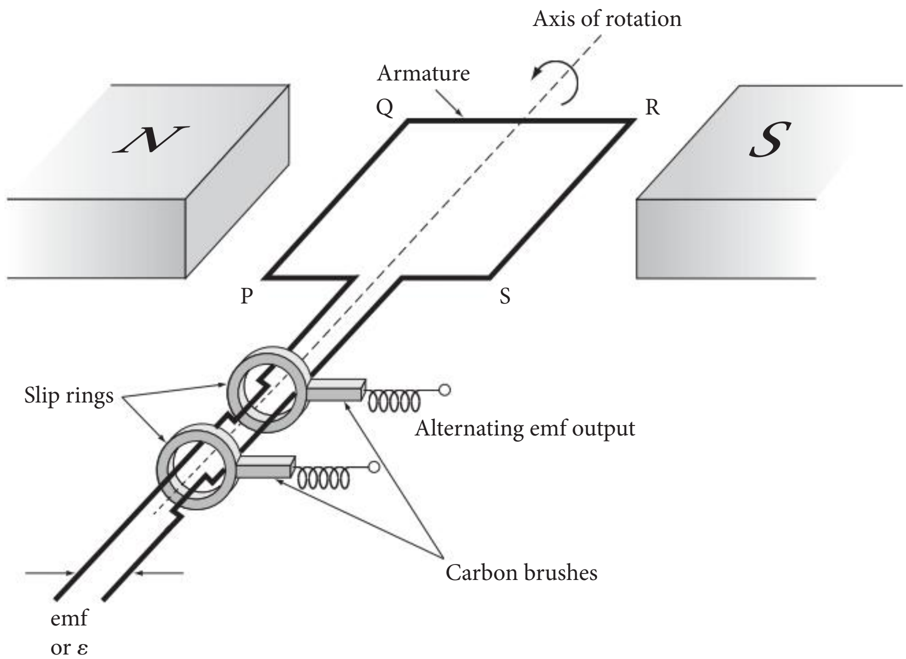
Generators

A generator uses the relative movement of coils of wire and magnets to induce an emf across the coils to generate a current. The energy required to produce the movement may come from any source. In Australia it is mostly supplied by burning coal. A small fraction comes from the gravitational potential energy of water (hydroelectric power stations) and the kinetic energy of air molecules (wind turbines). Many other nations use nuclear energy. All of these energy sources require generators to produce electricity.

AC generators

Figure 5.15 shows a very simple alternating current (AC) generator. An alternating current is one that varies between positive and negative values. Usually AC varies sinusoidally. The coil is attached to an **armature** that rotates in the magnetic field between the poles of the two magnets. As it rotates the flux through the loops of the coil varies, causing an emf across the ends of the coil. Each end of the coil is attached to a conducting slip ring that slides against a brush. The brushes are then connected to the external circuit that uses the emf generated.

Figure 5.15 ▶ The rotation is circular, so the circular functions, *sine* and *cosine*, represent the loop's motion.

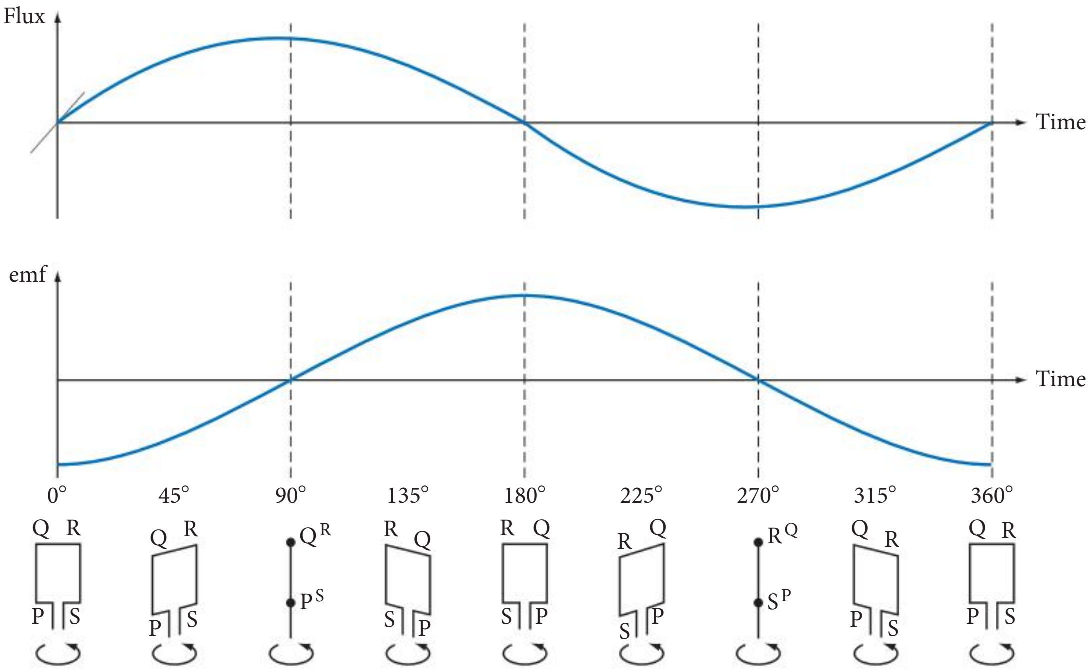


In Figure 5.16 the flux vs time graph is a sine curve because the original flux is zero.

The flux varies with the angle, θ, which varies in time such that $\theta(t) = \left(\frac{2\pi}{T}\right) t = 2\pi ft$, where T is the period of rotation. The frequency is $f = \frac{1}{T}$. Hence, the flux as a function of time is given by:

$$\Phi = nBA\sin(2\pi ft)$$

where a coil with n turns of area A rotates in a magnetic field of magnitude B.



▲ **Figure 5.16** A loop rotates in a circle. The flux through the loop describes a sine curve (top graph). The rate of change of flux – the gradient – also describes a sinusoidal curve, but the emf is the negative of the rate of change of flux (lower graph). The rotation of the coil is shown for each one-eighth of a turn.

The emf is the negative of the gradient of the flux as a function of time.

So the emf is given by:

$$\text{emf} = 2\pi fnBA\cos(2\pi ft)$$

This is shown in Figure 5.16. Note that when the flux is changing most rapidly, the emf has its maximum values. For a sine curve, the gradient is greatest at $t = 0$, $t = \frac{T}{2}$ and again at the end of each cycle. When the flux is at a peak, at $t = \frac{T}{4}$ and $t = \frac{3T}{4}$, the gradient is momentarily zero, so the emf is zero. The flux and emf have the same frequency.

The maximum emf occurs when $\cos(2\pi ft) = \pm 1$; then:

$$\varepsilon_{max} = 2\pi fnBA$$

The maximum emf can be changed by changing f, n, B or A. If f is changed, the period changes as well as the emf.

Usually the armature has a large coil of wire, as the emf is proportional to the number of loops or turns in the coil. However, the bigger the coil, the heavier it is, so sometimes it is the magnets that are rotated instead of the coil.

WORKED EXAMPLE 5.3

A square coil of side length 0.10 m is made up of 400 turns. It is rotated at 25 Hz in a magnetic field of magnitude 0.10 T.

a Find the maximum emf induced. (3 marks)

b If, at $t = 0$, there is a maximum flux through the coil, sketch the emf as a function of time. (4 marks)

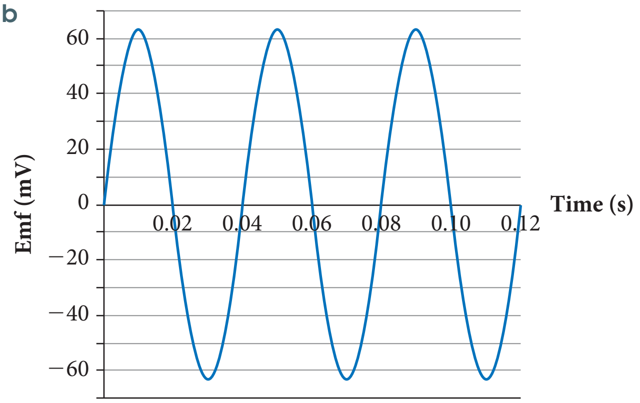
Answers

a $\varepsilon_{\text{max}} = 2\pi fnBA$

$\varepsilon_{\text{max}} = 2\pi \times 25\,\text{Hz} \times 400 \times 0.10\,\text{T} \times (0.10\,\text{m} \times 0.10\,\text{m})$

$\varepsilon_{\text{max}} = 63\,\text{V}$

b



▲ **Figure 5.17** Emf as a function of time

Logic

Logic	Marks
Relate emf to parameters given.	1 mark
Substitute the correct values with units.	1 mark
Calculate the final value.	1 mark
We know from part a that the emf varies between −63 V and +63 V. The period of its oscillations is the same as the period of rotation, which is $T = \frac{1}{f} = \frac{1}{25}\text{Hz} = 0.04\,\text{s}$. Note that as we do not know which way the coil is turning, a sketch showing emf starting from zero and decreasing first is also possible.	4 marks

Try this yourself

What effect would doubling the frequency have on the maximum emf? Sketch emf(t) for this case. (5 marks)

EXPERIMENT 5.1

A SIMPLE AC GENERATOR

A simple AC generator can be made by attaching a magnet to a spring and oscillating it in a coil, as shown in Figure 5.18.

GENERATORS

This simulation shows how a generator works.

Aim

To measure the current produced by the simple AC generator

Materials

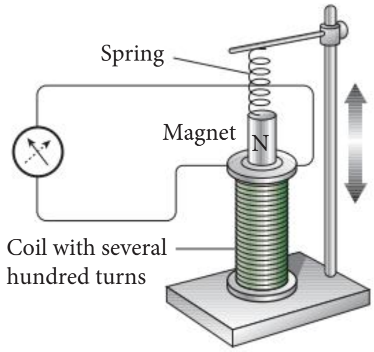
- retort stand and clamps
- spring
- magnet
- weights
- coil
- oscilloscope or data logger connected to computer

What are the risks in doing this experiment?	How can you manage these risks to stay safe?
The magnet could fly off the spring and hit someone.	Make sure the magnet is well attached and do not oscillate it too vigorously.

In your write-up, add any more risks you can think of, as well as ways to manage them.

Procedure

1 Attach the spring to the retort stand as shown in Figure 5.18.
2 Attach the magnet and one weight to the spring.
3 Place the coil below the magnet. Adjust the height of the spring and magnet so that at equilibrium the magnet is just inside the coil.
4 Connect the coil to an oscilloscope or data logger so that you can measure the emf produced.
5 Pull the magnet down to start it oscillating. You may need to move the coil out of the way to do this, then put it back in place.
6 Record the period of oscillation and the maximum emf produced. It is more precise to measure 10 complete oscillations, then divide by 10, to get the period of oscillation. Don't forget to include uncertainties in your results.
7 Repeat the measurements, adding weights to vary the frequency of oscillation. You will need to adjust the height of the spring each time.



▲**Figure 5.18** A simple generator can be made by connecting a magnet to a spring and oscillating it in a solenoid.

Results

Record the maximum emf as a function of period of oscillation in a table like the one below. Don't forget to include units and uncertainties in your results.

Calculate the frequency of oscillation and add this to your table as shown.

Time for 10 oscillations	Period	Frequency	Maximum emf

Analysis of results

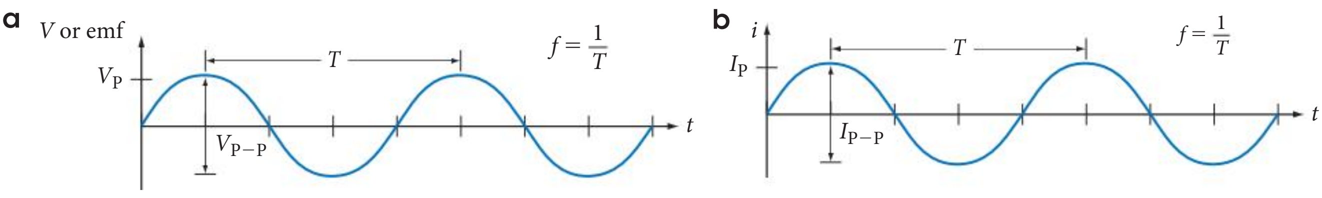
Draw a graph of maximum emf as a function of frequency of oscillation. Comment on the shape of your graph.

Discussion

1 Do your results agree with what you would expect from the equation for the emf of a generator?
2 How could you improve this experiment to make it more accurate? What could you do to extend it?

AC quantities

For a sinusoidal potential difference, the relevant quantities are peak potential difference, V_P, peak-to-peak potential difference, V_{P-P}, period, T, and frequency, f. These are shown graphically in Figure 5.19.



▲**Figure 5.19** Induced emf gives rise to an induced current. The peak voltage and peak current are proportional to each other and they have the same frequency. a) The induced emf or voltage has a peak value, V_P, that is half the peak-to-peak value, V_{P-P}. It describes one cycle in one period of time, T. b) The induced current, i, has a peak value, I_P, that is half the peak-to-peak value I_{P-P}. It describes one cycle in one period of time, T.

When dealing with **direct current (DC)**, potential difference and current are constant, or at least constantly in the same direction. AC values vary between a peak positive value and a peak negative value, oscillating back and forth in each cycle. The average of the AC potential difference over one cycle is zero, yet the AC potential difference obviously delivers energy during that time; that is, it delivers power to a circuit. Power is proportional to the square of the potential difference. If we square the potential difference and find the average, we can get a value for the average power. To convert this to a single potential difference that would deliver the same power as the original AC potential difference, we take the square root of this average. The single value of potential difference that we get when we square, average and take the square root is called the **root mean square** or rms value:

$$V_{rms} = \frac{V_{peak}}{\sqrt{2}} = 0.707V_{peak}$$

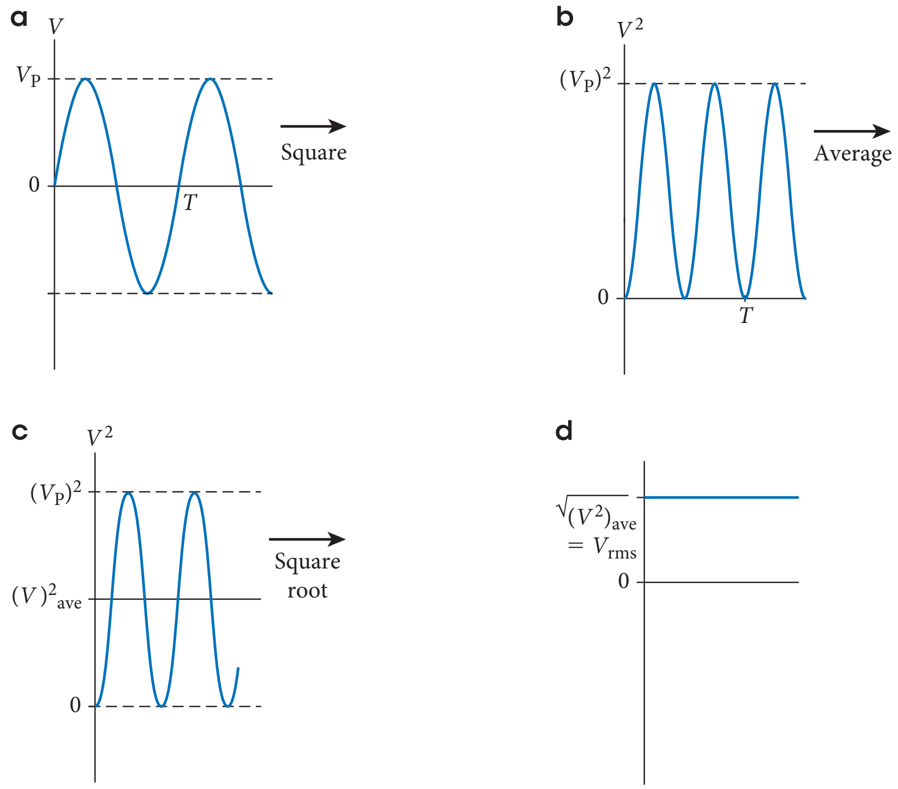


Figure 5.20 a) A sinusoidal $V(t)$. b) We square this to get $V^2(t)$. This has a peak value of V_P^2. c) The average value of $V^2(t)$, which has the value $\frac{1}{2}V_P^2$. d) Taking the square root of this value gives us the rms value: $V_{rms} = \frac{1}{\sqrt{2}}V_{peak}$.

Similarly: $$I_{rms} = \frac{1}{\sqrt{2}}I_{peak} = 0.707I_{peak}$$

The rms potential difference is an average AC potential difference that produces the same power in a resistive component as a constant DC potential difference of the same magnitude. AC systems are usually described using rms values.

For an AC generator, the maximum output emf is $\varepsilon_{max} = 2\pi fnBA$. The rms value of the output emf is therefore:

$$\varepsilon_{rms} = \frac{\varepsilon_{max}}{\sqrt{2}} = \frac{2\pi fnBA}{\sqrt{2}} = \sqrt{2}\pi fnBA$$

The rms current produced is

$$I_{rms} = \frac{I_{max}}{\sqrt{2}} = \frac{\sqrt{2}\pi fnBA}{R}$$

where R is the resistance of the load attached to the generator output.

AC AND RMS VALUES

This page has more details of AC signals and how to find rms values.

Recall from your studies of circuits in Nelson Physics Units 1 and 2 for the Australian Curriculum that current, I, and voltage, V, for a load of resistance R are related by Ohm's law: V = IR.

WORKED EXAMPLE 5.4

The rms potential difference produced by an AC generator is 240 V. To what peak value does this correspond? (4 marks)

Answer	Logic	
$V_{rms} = \frac{1}{\sqrt{2}} V_{peak}$	Relate V_{rms} to V_{peak}.	1 mark
$V_{peak} = \sqrt{2} V_{rms}$	Rearrange for V_{peak}.	1 mark
$V_{peak} = \sqrt{2}(240\ \text{V})$	Substitute the correct values with units.	1 mark
$V_{peak} = 340\ \text{V}$	Calculate the final value.	1 mark

Try this yourself

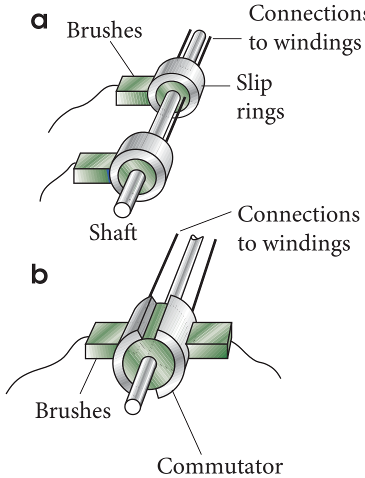
What is the peak potential difference of an AC generator built for use in the USA where $V_{rms} = 110\ \text{V}$? (3 marks)

DC generators

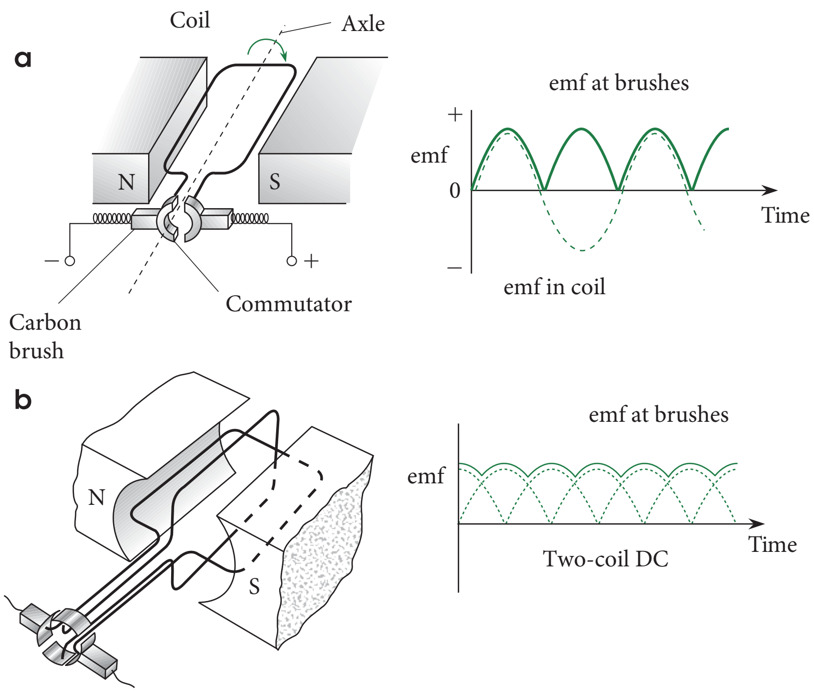
A generator that provides an output emf that is always positive is called a DC generator. In a DC generator, a **commutator** is used instead of the slip rings used in an AC generator, as shown in Figure 5.21.

In the commutator, each side of the coil is connected to a conducting copper strip. These are separated by insulators. As the commutator rotates past the carbon brushes, the effect is the same as if the connections to the slip rings were reversed each half-cycle. The output for a single coil is shown in Figure 5.22(a). The output is 'lumpy' or pulsed, rising to a maximum and dropping back to zero each half turn.

To smooth out the lumpy DC current, it is usual to use many coils that are offset relative to each other to provide a steady current. Each coil has separate pairs of connections to the



▲Figure 5.21 a) An AC generator uses slip rings. b) A DC generator uses a commutator.



◀Figure 5.22 In a DC generator, the carbon brushes make contact with a commutator, which allows the connections to the coils to switch every half cycle. a) A single coil gives a fluctuating output. b) This problem is overcome by using two or more coils.

commutator. The output for a two-coil DC generator is shown in Figure 5.22(b). These two coils are offset by 90°. One of the problems of using multiple commutators is that there can be sparking across the insulating gap. This can be dangerous in car engines where there are flammable gases nearby.

QUESTION SET 5.2

Remembering

1 Name three sources of energy used in electricity production.

2 What is the purpose of the commutators in a DC generator?

Understanding

3 Why is there an alternating potential difference produced in a coil when it is rotated in a uniform magnetic field?

Applying

4 An AC current source produces a current as a function of time given by $i(t) = 30A \sin(6\pi t)$.

- a What are the maximum, average and rms values of current supplied by this current source?
- b What is the frequency of this source?
- c What is the period of the source?

5 The armature of an AC generator is rotating at a constant speed of 35 revolutions per second in a horizontal field of flux density 1.0 T. The diameter of the cylindrical armature is 24 cm and its length is 40 cm.

- a What is the maximum emf induced in the armature if it has 30 turns?
- b What is the rms emf produced by this generator?

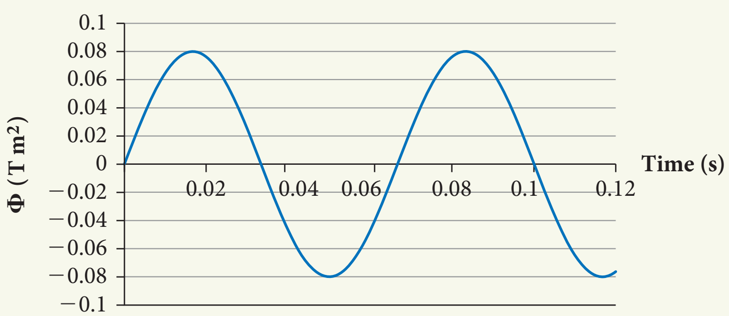
6 A flat rectangular coil 15 cm by 25 cm has 300 turns.

An alternating emf of peak value 340 V is produced when the coil rotates at 3000 revolutions per minute in a uniform magnetic field. What is the value of the magnetic field strength?

7 A rectangular coil of 30 turns and area 100 cm^2 rotates at 1200 revolutions per minute in a uniform magnetic field of flux density 0.50 T.

- a Find the frequency of the generated emf.
- b Find the maximum emf.
- c What is the rms emf?
- d Write the equation that gives the emf at any instant.

8 Figure 5.23 shows the magnetic flux as a function of time through each loop of a 30-turn coil in a generator.



◀Figure 5.23

- a Find the maximum and minimum emf and the period of oscillation of the emf.
- b Draw a graph showing the emf produced as a function of time.
- c Mark the rms emf on the graph.

9 The armature of a 50 Hz AC generator rotates in a magnetic field of strength 0.15 T. If the area of the coil is 2.5×10^{-2} m^2, how many turns must the coil contain if the maximum emf produced is 150 V?

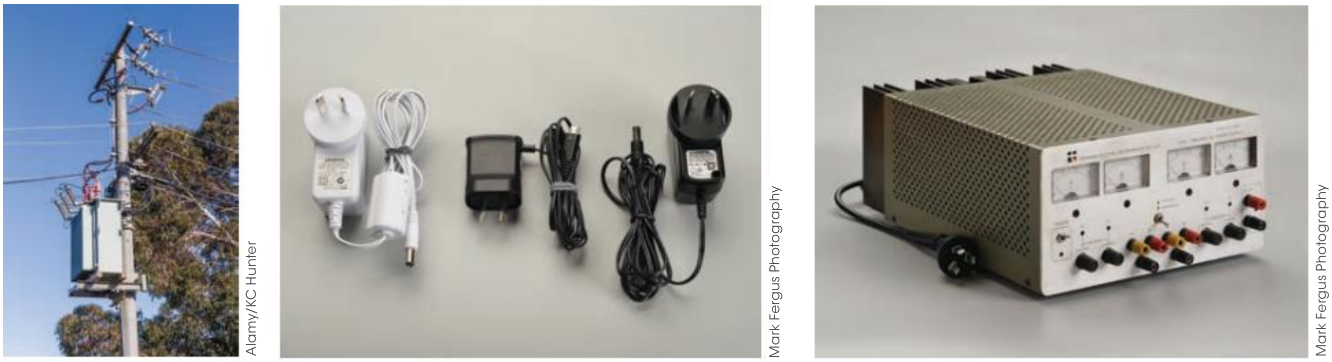
Reflecting

10 Australia's electricity production and transmission infrastructure is based on the use of generators producing AC power. How would things be different if instead we converted completely to power generated by photovoltaic cells? What infrastructure would have to change? Construct an argument for or against such a change.

Transformers

Most appliances use a transformer to convert the 240 V mains power to a lower potential difference. Some also convert the AC potential difference to DC. The transformer may be inside the device, or it may be a separate plug pack. There are also transformers at electricity substations. These drop the potential difference from the thousands of volts at which it is transmitted to the 240 V that is supplied to homes and businesses.

▼Figure 5.24
Some transformers



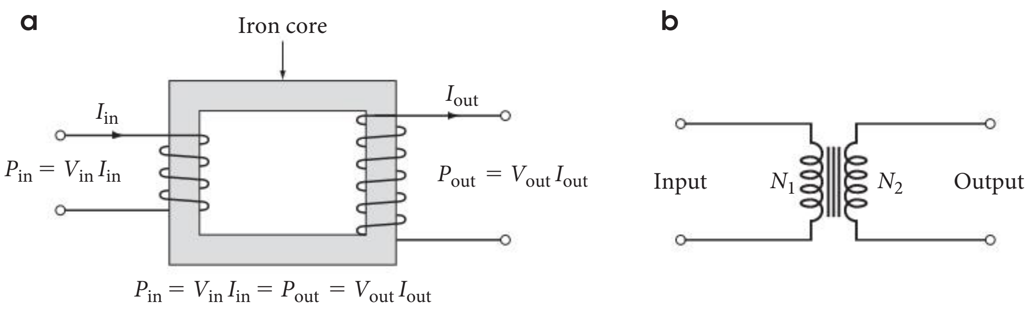
A transformer consists of two solenoids or coils of wire placed near each other so that an alternating current in the primary coil can induce a current in the secondary coil. The link between input and output is by electromagnetic induction; there is no electrical connection.

Solenoids are used because they produce a large and approximately uniform magnetic field inside the coil. The field in the primary varies sinusoidally with the alternating current in the coil. The coils need to be coupled so that the changing magnetic field in the primary coil causes a changing magnetic flux in the secondary coil. There are two ways of doing this. First, the coils can share the same space by placing one within the other. This is sometimes used in cordless appliances such as kettles. The second, and more usual, way is to link the coils using a ferromagnetic core.

Figure 5.25 shows how this works. The primary coil is wound around one side of an iron core. The current in the primary coil magnetises the whole core, not just the part within the primary coil. The time-varying current in the primary coil causes a time-varying magnetic field inside the secondary core. This creates a time-varying electric field, hence an emf and current in the secondary coil.

Solenoids produce a large uniform magnetic field, as described in Chapter 4.

A ferromagnetic material becomes magnetised in a magnetic field. All the tiny magnetic dipoles, which are individual atoms or molecules, align to give a large magnetic field. This was described in Chapter 4.



◀Figure 5.25
a) A schematic diagram of a transformer and b) the circuit symbol for a transformer.

The flux through any loop is the same for both coils. If the primary coil has N_p turns then:

$$V_p = -N_p\left(\frac{\Delta\Phi}{\Delta t}\right)$$

Similarly:

$$V_s = -N_s\left(\frac{\Delta\Phi}{\Delta t}\right)$$

V_s is proportional to the rate of change of the flux. Flux is proportional to V_p. Hence V_s is proportional to the rate of change of V_p, or $V_s \propto \frac{dV_p}{dt}$. A transformer is a differentiator!

As $\frac{\Delta\Phi}{\Delta t}$ is the same for both coils:

$$\frac{V_s}{V_p} = \frac{N_s}{N_p}$$

Assuming that the transformer is 100% efficient, power out = power in:

$$P_{out} = P_{in}$$

Recall from *Nelson Physics Units 1 & 2 for the Australian Curriculum* that $P = VI$, so:

$$I_s V_s = I_p V_p$$

$$\Rightarrow \frac{I_p}{I_s} = \frac{N_s}{N_p}$$

TRANSFORMERS

This page shows how a transformer works. It also has a transformer calculator.

Transformer equations:

$$\frac{V_s}{V_p} = \frac{I_p}{I_s} = \frac{N_s}{N_p}$$

In reality, transformers are not 100% efficient, and a small amount of energy is lost as heat through resistance and eddy currents.

A **step-up transformer** ($N_s > N_p$) has a higher emf and lower current on the secondary side. A **step-down transformer** ($N_s < N_p$) has a lower emf and higher current on the secondary side.

Step-down transformers are sometimes connected to a **rectifier** circuit, such as that shown in Figure 5.26. A rectifier converts an alternating current to a direct current. This is done by passing the current through two pairs of **diodes**. The current output across the load varies in time but is always in the same direction. The variation is called a 'ripple', which is reduced by using a smoothing capacitor and resistor.

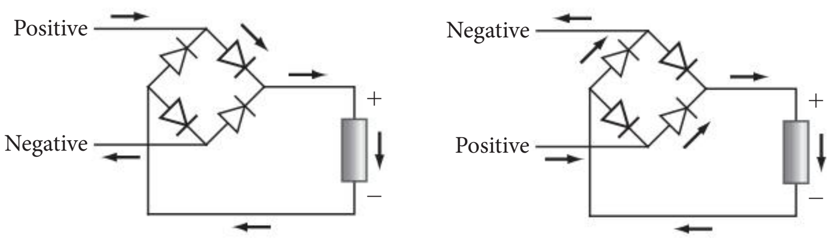


Figure 5.26 ▶ Converting AC to DC – the potential difference across, and the current through, the resistor is the same for both input conditions.

WORKED EXAMPLE 5.5

A 120 W, 24 V AC supply is connected to the input terminals of a transformer. The primary coil is wound with 240 turns. The output emf is 72 V. Assume there is no power loss in the transformer.

a Find the number of turns on the secondary coil. (4 marks)

b Is this a step-up or step-down transformer? (1 mark)

c What is the output current? (5 marks)

Answers	Logic	
a $\frac{V_s}{V_p} = \frac{N_s}{N_p}$	Relate the number of turns to the parameters given.	1 mark
$N_s = N_p \frac{V_s}{V_p}$	Rearrange for N_s.	1 mark
$N_s = 240 \left(\frac{72\text{ V}}{24\text{ V}}\right)$	Substitute the correct values with units.	1 mark
$N_s = 720$	Calculate the final value.	1 mark

b	Step-up transformer	The output emf is higher than the input, and the secondary coil has more turns than the primary coil. This must be a step-up transformer.	1 mark
c	$P_{out} = P_{in}$	Apply conservation of energy.	1 mark
	$P_{out} = I_s V_s = P_{in}$	Relate current to power.	1 mark
	$I_s = \frac{P_{in}}{V_s}$	Rearrange for current.	1 mark
	$I_s = \frac{120\ \text{W}}{72\ \text{V}}$	Substitute the correct values with units.	1 mark
	$I_s = 1.7\ \text{A}$	Calculate the final value.	1 mark

Try these yourself

1 A transformer has a primary coil with 500 turns and a secondary coil with 150 turns. The input potential difference is 240 V. What is the output emf? (2 marks)

2 A transformer needs to supply 12 V and 1.0 A to a laptop from the mains power supply.

  a What is the ratio of turns on the secondary coil to the primary coil? (2 marks)

  b What is the input current to the transformer? (2 marks)

EXPERIMENT 5.2

TRANSFORMERS

Aims

To observe the effect of flux changes in the secondary coil of a transformer caused by changing potential difference across the primary

To compare the ratio of the potential differences to the turns ratio in both step-up and step-down transformers

Materials

- 2 air-core solenoids with known number of turns
- iron rod
- AC/DC power supply
- momentary switch
- oscilloscope (CRO)

What are the risks in doing this experiment?	How can you manage these risks to stay safe?
Power supplies can cause electric shocks.	Do not use or induce potential differences above 15 V. Make sure your equipment is checked by your teacher before use.

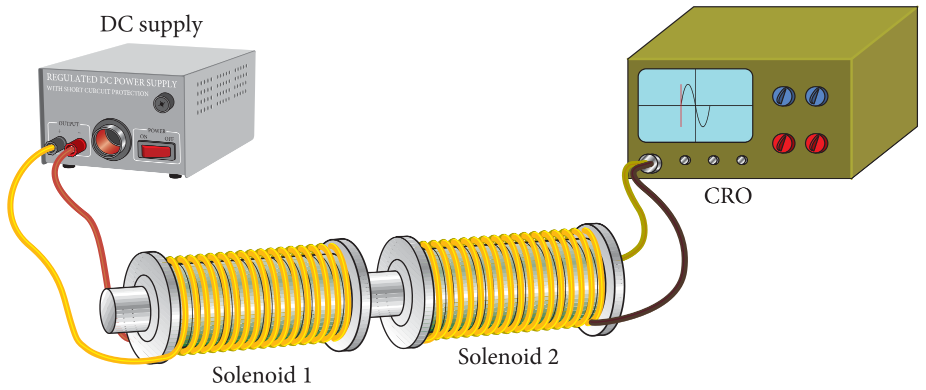
In your write-up, add any more risks you can think of, as well as ways to manage them.

Part A

Procedure

1 Place the air-core solenoids end to end with the iron core placed inside them as shown in Figure 5.27.
2 Connect one solenoid to the DC power supply set to 2 V.
3 Connect the other solenoid to a CRO.
4 Tap the switch to turn the DC power off and on.
5 Record any change in the trace on the CRO. You may need to adjust the time base.
6 Describe what happens.

7 Turn the power supply on and observe what happens.
8 Now turn it off and observe what happens.
9 Record your results.



▲ **Figure 5.27** Experimental set-up for part A

Results

Draw a diagram showing what happened when you tapped the switch.

Analysis of results

How does tapping the switch like this compare to the switch simply being turned on?

Part B

Procedure

1 Change the power supply to 2 V 50 Hz AC.
2 Turn on the power supply and, by looking at the CRO trace, observe what happens in the second solenoid.
3 Exchange the two solenoids and repeat your measurements.

Results

Sketch the CRO trace.

Analysis of results

1 Describe how this differs from part A in which DC potential difference was used.
2 Sketch the input and output voltages on the same set of axes for each arrangement of solenoids.
3 Calculate the output voltage predicted by the transformer equations. Does it agree with your measurements?

Discussion

1 For a DC potential difference applied across the primary, when does the flux change? What happens in the secondary?
2 How well did the transformer equations predict the results of your experiment? How can you explain any discrepancies?
3 How could you reduce experimental uncertainties in this experiment?

Transformers can be used to improve safety of people and equipment. Isolation transformers are used to ensure that there is no direct electrical connection between electronic medical equipment and patients. They are also used to protect other medical instruments when defibrillators are used to restart a patient's heart. Sometimes isolation transformers are used in laboratories to prevent electronic 'noise' interference from the mains power supply.

QUESTION SET 5.3

Remembering

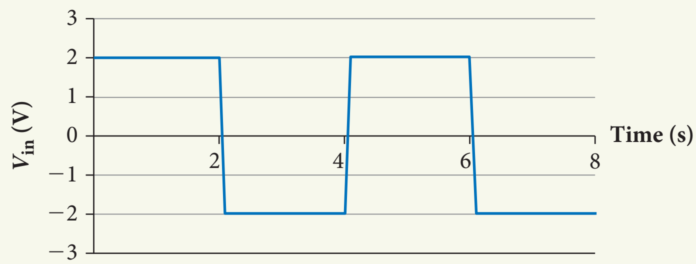
1 What is the difference between a step-up and a step-down transformer?

Understanding

2 Why is an alternating current necessary for a transformer?

3 Why must the output current in a step-up transformer be less than the input current?

4 Figure 5.28 shows the input potential difference to a transformer. Sketch the output potential difference.



◀Figure 5.28

Applying

5 For a transformer with a 1000-turn primary, what potential difference is available at the 200-turn secondary when the primary coil is supplied with 240 V AC?

6 A transformer is connected to an AC source that can deliver 30 A. The secondary coil of the transformer can deliver a maximum current of 10 A.

- a What type of transformer is this?
- b Calculate the ratio between the number of turns in the primary and the number of turns in the secondary.

7 A step-up transformer is connected to an AC generator that delivers 8.0 A at 120 V. The ratio of the number of turns in the secondary coil to the number of turns in the primary is 500.

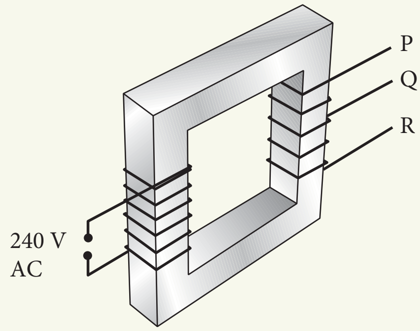
- a What is the emf in the secondary coil?
- b What is the power input?
- c What is the maximum power output?
- d What is the maximum current in the secondary?

Analysing

8 Transformers are not 100% efficient. Their efficiency can be improved by using a laminated core; that is, a core made from thin slices of iron sandwiched together with glue. Explain how this reduces energy loss in a transformer.

9 A step-down transformer is connected to a 240 V 50 Hz AC mains supply. There are 1200 turns in the primary coil. The secondary coil has three terminals, P, Q and R, as shown in Figure 5.29. When a 10 Ω resistor is connected between terminals P and Q, the current flow in the resistor is 0.60 A. When the same resistor is connected between terminals Q and R, the current flow is 1.0 A.

- a What is the potential difference between P and Q?
- b What is the potential difference between Q and R?
- c Find the number of turns in the secondary coil between terminals:
  - i P and Q.
  - ii Q and R.
- d What is the potential difference between P and R?
- e What is the current in the 10 Ω resistor if it is connected between terminals P and R?



▲Figure 5.29

Electric motors

Electric motors are essential in industry, transport and everyday appliances such as refrigerators, washing machines, air conditioners and dishwashers. Many electrical measuring devices also use the rotational effects of current elements in magnetic fields.

Faraday came up with the idea of motors in 1821 as an application of the 'lines of force'. The energy transformations in a motor are the reverse of those in the generator.

Generator: kinetic energy is converted to electric potential energy.
Motor: electric potential energy is converted to kinetic energy.

DC motors

A pair of equal but opposite forces like the forces acting in Figure 5.30 is sometimes called a couple. The resulting torque is a couple moment.

We saw in the previous chapter that a current experiences a force in a magnetic field. The magnitude of the force is given by $F = I\ell B\sin\theta$.

Consider a loop of current-carrying wire in a magnetic field such as that shown in Figure 5.30(a). The length of wire between points N and M has a force, $F = I\ell B\sin\theta$ acting upwards on it. The wire between P and Q has an equal force acting downwards on it. The net force is therefore zero, and we do not expect the loop to go either up or down. However, the loop will move because it experiences a torque. A torque is a push or pull that acts to rotate an object.

As you saw when you studied forces in Chapter 1, torque is the rotational equivalent of force. The torque exerted by any force, F, acting at some distance, r, from an axis of rotation is

$$\tau = rF\sin\theta$$

where θ is the angle between the force, $\vec{F}$, and the vector $\vec{r}$, which points from the axis to the point of application of the force.

The torque is a maximum when $\vec{r}$ and $\vec{F}$ are perpendicular, as in Figure 5.30. Torque is zero when r and F are parallel. The expression for torque is sometimes written as

$$\tau = rF_{\perp}$$

Torque is more correctly given as the vector cross product between $\vec{r}$ and $\vec{F}$:

$$\tau = \vec{r} \times \vec{F}$$

This is another example of the use of the cross product in physics.

where the $\perp$ symbol indicates that it is the product of the vector $\vec{r}$ and the component of the vector $\vec{F}$ which is perpendicular to $\vec{r}$.

Like force, torque is a vector. We use the right-hand rule from Chapter 4, Figure 4.3, to find its direction. Point the fingers of your right hand in the direction of r, then curl them towards F. Your thumb now points in the direction of the torque. Note that the direction of the torque is parallel to the axis of rotation. The curl of your fingers gives the sense of rotation about the axis.

In Figure 5.30(a) the two forces acting on the loop result in a torque that acts to flip the loop over in a clockwise direction. The current has now reversed direction, and so the direction of the forces also changes. The new torque acts to push the loop back in the opposite direction, as in Figure 5.30(b).

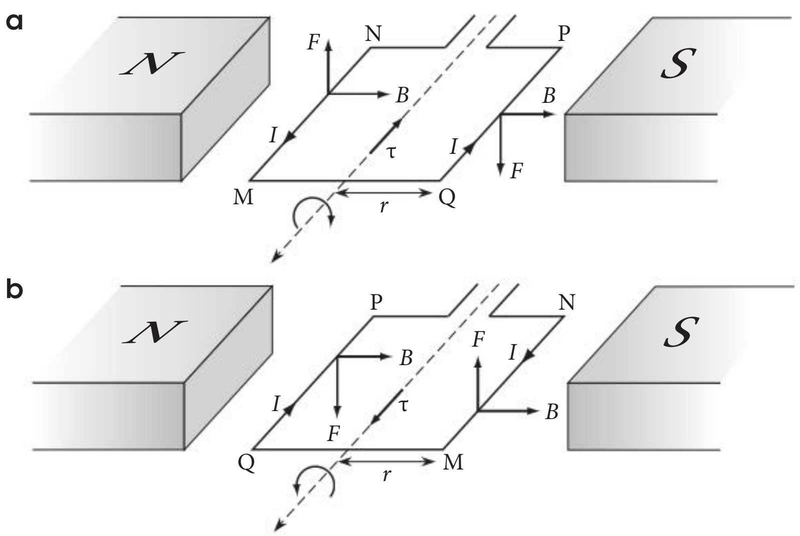
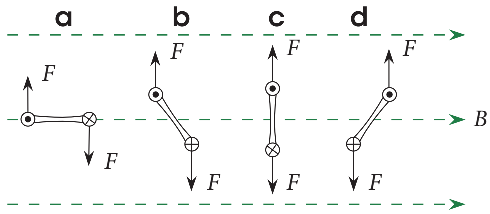


Figure 5.30 ▶ A current in a loop in a magnetic field experiences forces on the arms that are perpendicular to the field. a) The forces act in opposite directions, causing rotation. b) After a half-flip the forces are reversed and the loop flips back the other way.

Seen from the side, the coil starts in the position shown in Figure 5.31(a). It rotates through position (b) and continues to rotate. At position (c) the forces on each side act along the same wire and the net torque is zero. The coil has sufficient rotational momentum to carry it past this balance point. Once past this balance point the net turning effect is reversed, as shown in Figure 5.31(d). The coil now reverses direction and oscillates back and forth about the balance point. The oscillations get smaller due to friction. After a short time no rotation takes place and the coil comes to a stop.

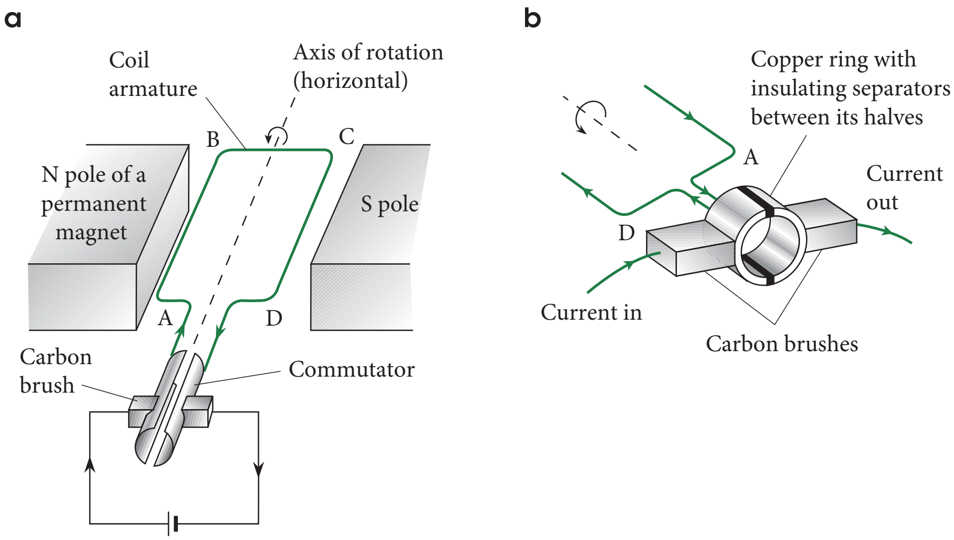


▲ **Figure 5.31**
Side view of a loop rotating about an axis that is at right angles to a magnetic field. a) Maximum net torque, b) decreasing net torque, c) no torque and d) torque reversed.

The current in the coil must be stopped or reversed each half-cycle of rotation to keep the coil rotating. Commutators or reversing switches do this.

Commutators

The commutator changes the electrical contacts on the wires as the coil's momentum carries it past its balance point. Brushes made of graphite or carbon blocks usually provide the sliding contact, as shown in Figure 5.32.

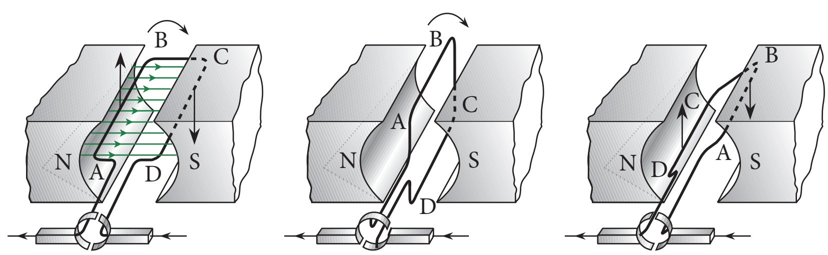


◀ **Figure 5.32**
a) The DC supply is connected to the carbon brushes. The commutator is free to move against the brushes. The insulating separators keep the two sides of the coil electrically separated. As the coil flips over, the current in the coil is reversed, which allows the torque to act in one direction over the full cycle. b) A close-up view of the commutator.

BUILD A MOTOR
Here is a simple design for a DC motor you can build.

HOMOPOLAR MOTOR
This is an even simpler motor, with only one magnet.

Practical motors use shaped magnets (Figure 5.33) or electromagnets and multiple coils at different angles around an iron core. These increase the magnetic field, the torques and the smoothness of the rotations.



◀ **Figure 5.33**
Shaped magnets ensure constant torque.

The speed of a DC motor is controlled by the supplied emf. The energy is being constantly supplied by the source of emf, and is being converted into kinetic energy, so it may seem that the coil should go faster and faster and faster as long as the motor is turned on. This would not violate conservation of energy, as the emf is continuing to supply energy. But this doesn't happen. Instead, the motor accelerates until it reaches a steady rotational speed. The effect of friction alone does not account for this.

We know from Lenz's law that the changing magnetic flux through a loop, due to the movement of the loop, induces an emf in the loop. This is called a **back emf**. The back emf must be in the opposite direction to the applied potential difference that caused the movement. The faster the coil spins, the more rapidly the flux changes and the bigger the back emf. This limits the current in the coil and hence the rate at which it spins.

WORKED EXAMPLE 5.6

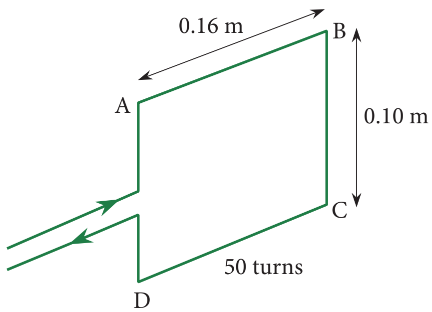
A motor uses a 50-turn coil that has dimensions 0.10 m by 0.16 m. A current of 2.0 A flows through the coil. The coil is vertical and is in a magnetic field of 0.12 T, directed upwards.

a What is the magnitude and direction of the force exerted on side AB? (5 marks)

b What is the magnitude and direction of the force exerted on side BC? (2 marks)

c What is the total torque acting on the coil? (8 marks)

d In which direction will the coil begin to rotate? (1 mark)



▲**Figure 5.34** A 50-turn current-carrying coil

Answers	Logic	
a $F = i\ell B\sin\theta$	Relate the force per section of wire to the variables given.	1 mark
$F = ni\ell B$	Recognise that the angle θ is 90° and that there are *n* lengths of wire in the section.	1 mark
$F = 50 \times 2.0\,\text{A} \times 0.16\,\text{m} \times 0.12\,\text{T}$	Substitute the correct values with units.	1 mark
$F = 1.9\,\text{N}$	Calculate the final value.	1 mark
$F = 1.9\,\text{N}$, to the right	Use the right-hand rule to find the direction.	1 mark
b $F = i\ell B\sin\theta$	Relate the force per section of wire to the variables given.	1 mark
$F = 0$	Recognise that the angle θ is 0° and hence $\sin\theta = 0$.	1 mark
c $\tau = rF\sin\theta$	Relate torque to force.	1 mark
$\tau_{AB} = rF_{AB}$	Find the torque acting on side AB.	1 mark
$\tau_{AB} = 0.05\,\text{m} \times 1.9\,\text{N}$	Substitute the correct values with units.	1 mark
$\tau_{AB} = 0.095\,\text{N m}$	Calculate the final value.	1 mark
$\tau_{BC} = rF_{BC} = 0$	Find torque acting on other sides	3 marks
$\tau_{CD} = \tau_{AB} = 0.095\,\text{N m}$		
$\tau_{DA} = \tau_{BC} = 0$		
$\tau_{total} = \tau_{AB} + \tau_{BC} + \tau_{CD} + \tau_{DA}$	Calculate the total torque.	1 mark
$\tau_{total} = 0.095\,\text{N m} + 0 + 0.095\,\text{N m} + 0 = 0.19\,\text{N m}$		
d Clockwise	Using your right hand, the torque is acting along the direction AB, so your fingers curl clockwise.	1 mark

Try this yourself

Repeat the question above but for a magnetic field directed horizontally to the right. (11 marks)

INVESTIGATION 5.1

MAKE YOUR OWN DC MOTOR

You can make your own DC motor with only a few simple parts. Once you have made your motor you can test it by measuring how fast it spins for a given input. You can also measure how powerful and efficient it is by attaching weights and getting your motor to lift them.

What is your aim?

Your first aim is to build a motor. Then you need to determine how well it works. Think about what measurements you can make. What will they tell you about the performance of your motor?

HOW REAL MOTORS WORK

View the insides of several motors, and learn about real motor design.

What will you need?

Start by doing some research on simple motor designs. Consider the materials you will need to make your motor. You will also need some other materials and measuring devices in order to test the performance of your motor.

What are the risks?

Construct a table similar to the one below. Identify specific risks involved in the investigation and ways that you will manage the risks to avoid injuries or damage to equipment. For example, will you be using high voltages or heavy weights? Will you be doing any soldering?

What are the risks in doing this investigation?	How can you manage these risks to stay safe?

How will you carry out your investigation?

Make sure you allow time to troubleshoot any problems with your motor. Often things do not work the first time, and need to be modified. This can take time.

What results will you collect?

Think about how you can measure motor speed and power. You may need to be creative. The motor will almost certainly spin too fast for you to count revolutions by eye. What else could you do? How can you minimise uncertainties in your results?

How will you analyse your results?

Think about what graphs you can plot to analyse your results. How will you determine uncertainties? Will you compare the performance of your motor to that of others?

What do you conclude?

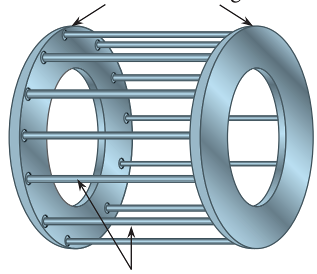
Write a conclusion that is consistent with your results and your aim. Don't forget to write down any ideas you have for improvement or further investigation.

AC induction motors

An AC **induction motor** uses the principle of electromagnetic induction. No current is supplied directly to the rotating coils, but a current is induced in them by using a changing magnetic field. The changing magnetic field is produced by an AC current. The **rotor** (the rotating part) consists of a cylinder with metal rods embedded in it along the length of the cylinder. These are electrically connected at each end of the cylinder to form closed loops. This forms a 'squirrel cage' (Figure 5.35).

The rotor sits between the poles of two electromagnets. The AC supplied to the coils of the electromagnets creates an oscillating magnetic field that induces a current in the rotor coils.

Transformers and induction motors are similar. The primary or stationary coils (**stators**), create the time-varying magnetic field. This produces a time-varying magnetic flux in the secondary coil or rotor. However, a motor is designed to rotate, so the stator coils are



▲ **Figure 5.35** A 'squirrel cage' rotor is used as the coil in an induction motor.

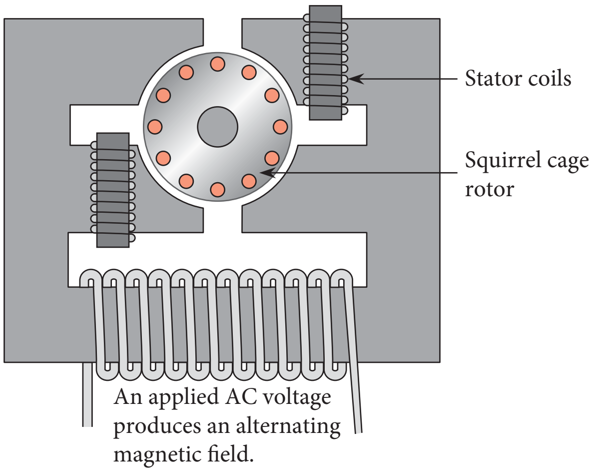


Figure 5.36 ▲
An AC induction motor

arranged to produce a rotating magnetic field. A force acts on the induced current in each connecting bar of the rotor due to the magnetic field from the stators. The connecting rods are arranged symmetrically, so there is no net force on the rotor; however, each rod experiences a torque. The torques are in the same direction and add up to give a net torque on the rotor, causing it to rotate. The basic configuration of an AC induction motor is shown in Figure 5.36.

The AC power supplied through the mains electricity has a frequency of 50 Hz; 50 times each second the current changes direction, and a motor connected to it runs at 50 Hz without a load. When there is a load, the speed is less because the rotor cannot keep up with the magnetic field. This means that AC induction motors generally have a top speed of about 3000 rpm.

Almost all motors in use are AC induction motors. Power tools such as drills and circular saws use AC induction motors, as do large industrial machines. They are widely used because of their simplicity of design and high efficiency. They are also low maintenance because they do not have commutators or brushes to wear out.

ACTIVITY 5.3

A VERY SIMPLE AC INDUCTION MOTOR

Aim

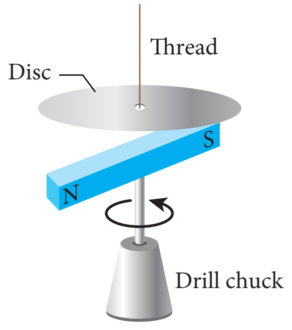
To make a very simple AC induction motor

You will need

- a piece of thread
- disc of metal such as a thin disc of aluminium
- hand drill
- pencil
- bar magnet

What to do

1. Connect the disc of metal to the thread, so that you can hang the disc horizontally.
2. Attach the magnet to one end of the pencil, and insert the other end of the pencil into the chuck of the hand drill.
3. Hold the drill so that the magnet is under, but not touching, the metal disc. Use the hand drill to spin the magnet.



◀ Figure 5.37 A very simple AC induction motor

What did you discover?

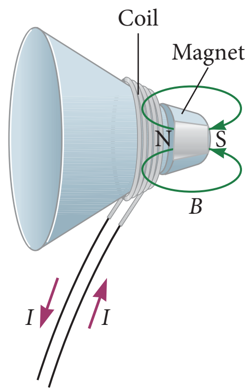
What happened when the magnet was spun? What happens when it is spun the other way?

WOW

Speakers are linear AC motors

Loudspeakers, such as you find attached to stereo systems, are simple linear AC motors. The moving coil is attached to an AC source. This source is the amplifier of the stereo system. The coil is attached to a lightweight cone, often made of paper, held in place by springs. The coil moves in the field of a large magnet. The springs are often made of pleated paper, and allow the cone to move back and forth.

The current coming along the speaker wires and into the coil is AC, with a complicated waveform made up of many frequencies. This is the music converted into a current. This current experiences a force, which makes the coil oscillate backwards and forwards as the current changes direction. The bigger the current, the bigger the force and the bigger the oscillations. The bigger the oscillations, the larger the amplitude of the wave produced by the oscillating cone and the louder the sound. The faster the oscillations, the higher the frequency of the sound.



▲ Figure 5.38 A speaker is a simple linear AC motor.

QUESTION SET 5.4

Remembering

1 State the purpose of the commutator in a DC motor.

2 Why does an AC motor not need a commutator?

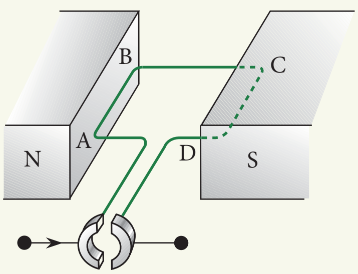
Understanding

3 Briefly explain the principle of operation of an AC induction motor.

4 Why does a current-carrying loop in a magnetic field experience a torque but no net force? Draw a diagram to help explain your answer.

Applying

5 Figure 5.39 shows a simple electric motor. In which direction must the current flow for the coil to rotate in a clockwise direction as seen from the end of the coil AD? Copy the diagram and show the magnetic field, the current, the forces on sides AB and CD, and the torque on the coil.



◀Figure 5.39

6 Figure 5.40 shows a rectangular coil placed between the poles of a strong magnet as part of a DC motor. The coil has 40 turns of wire and each turn is rectangular with length 0.060 m and width of 0.040 m. The magnetic field is uniform and has a magnitude of 0.050 T. The coil is free to rotate when placed between the poles of the magnet. A current of 1.5 A flows in the coil. At the instant represented in the diagram, the plane of the coil is parallel to the magnetic field.

a What is the magnitude and direction of the force on side AB?

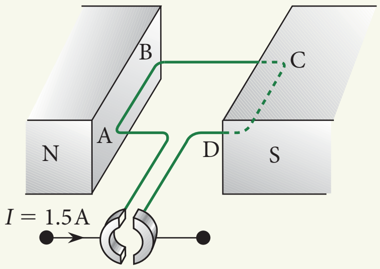
b What is the magnitude and direction of the force on side BC?

c What is the magnitude and direction of the force on side CD?

d What is the magnitude of the torque acting on the coil?

e In what direction – clockwise or anticlockwise – will the coil rotate?

f How could the magnitude of the torque be increased?



◀ Figure 5.40

7 If the coil shown in Figure 5.40 and described in question 6 is to experience a torque of 0.015 N m, how much current must be supplied to it?

Analysing

8 Why do most power tools and appliances such as food processors use AC rather than DC motors?

9 Why do AC induction motors have a top speed of 3000 rpm?

Reflecting

10 Imagine you work for an engineering company and someone tries to sell you a design for a DC motor that runs faster and faster as long as it is connected to a battery, with no limit to its top speed. Would you invest in the patent? What advice would you give to the company director on this?

Electromagnetic waves

Maxwell's most important contribution to physics was to take the equations and experimental results of Faraday and others, and unify them into a single theory of electromagnetism. This theory is summed up in four differential equations:

$\oint \vec{E}.d\vec{A} = \dfrac{q}{\varepsilon_0}$	Gauss's law
$\oint \vec{B}.d\vec{A} = 0$	Gauss's law in magnetism
$\oint \vec{E}.d\vec{s} = -\dfrac{d\Phi_B}{dt}$	Faraday's law
$\oint \vec{B}.d\vec{s} = \mu_0\varepsilon_0\dfrac{d\Phi_E}{dt} + \mu_0 I$	Ampère–Maxwell law

In words, these equations say:

Electric fields are created by charges.

There are no isolated magnetic poles (monopoles).

Electric fields are created by changing magnetic fields.

Magnetic fields are created by moving charges (currents) and changing electric fields.

Once the electric and magnetic fields are known at some point in space, the force acting on a particle at that point is given by:

$$\vec{F} = q\vec{E} + q\vec{v}\vec{B}\sin\theta$$

Maxwell's equations, together with this force law, completely describe all classical electromagnetism.

Maxwell found that when he combined the second pair of equations, they gave a **wave equation** for the electric and magnetic fields. This led him to predict the existence of electromagnetic waves, and to realise that light waves are a form of electromagnetic waves. Electromagnetic waves consist of coupled, self-sustaining oscillating electric and magnetic fields. From the wave equation, he predicted that the speed of light in a vacuum is $c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}} = 3.00 \times 10^8\,\text{m s}^{-1}$. This agreed within uncertainty with the experimentally measured values for the speed of light. In addition, the speed depends only on the constants μ_0 and ε_0, which are properties of empty space. Maxwell's theory did not require a medium, or aether, for the light waves to travel through. This was very important for the development of relativity, as you shall see in Chapter 6.

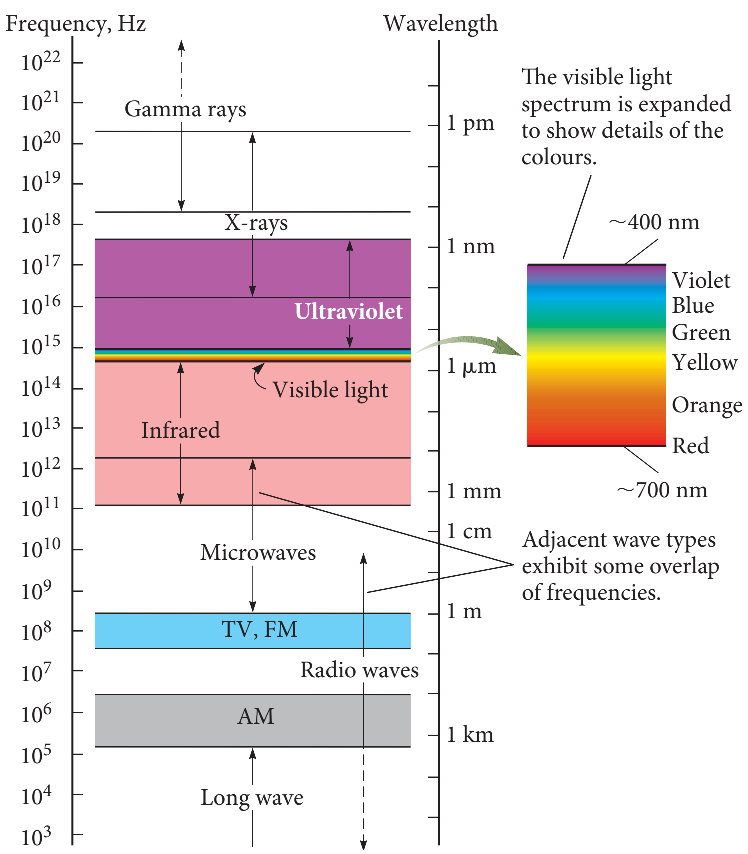
Electromagnetic waves carry energy and momentum, as do all other waves. They also show the wave behaviours that you studied in Unit 2 of *Nelson Physics Units 1 & 2 for the Australian Curriculum*. They obey the superposition principle and show interference and diffraction effects. The **dispersion relation**, $v = f\lambda$, gives us a relationship between speed, frequency and wavelength for light, just as it does for mechanical waves.

For light in a vacuum, $v = c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}} = 3.00 \times 10^8\,\text{m s}^{-1}$. For light in any other medium, the speed is lower and depends on the permittivity and permeability of the medium; $v = \sqrt{\mu\varepsilon}$. Recall from *Nelson Physics Units 1 & 2 for the Australian Curriculum* Chapter 10 that the refractive index of a material is given by $n = \frac{c}{v}$. While the refractive index was defined and measured more than a century before μ_0 and ε_0 were postulated, we can now formally define it as $n = \frac{c}{v} = \frac{\sqrt{\mu\varepsilon}}{\sqrt{\mu_0\varepsilon_0}}$.
For water, for example, $n = \frac{\sqrt{\mu_{\text{water}}\varepsilon_{\text{water}}}}{\sqrt{\mu_0\varepsilon_0}} = 1.33$ leads to values of $\mu_{\text{water}} = 1.27 \times 10^{-6}\,\text{T m A}^{-1}$ and $\varepsilon_{\text{water}} = 1.58 \times 10^{-10}\,\text{C}^2\,\text{N}^{-11}\,\text{T m}^{-2}$.

Maxwell also predicted that a large range of frequencies was possible for electromagnetic waves, well beyond the visible spectrum.

Wave equations are differential equations. They relate the rate of change of displacement in space to the rate of change of displacement in time. Maxwell found that the electric and magnetic fields obey wave equations. Note that v = λ is not a wave equation. It is a dispersion relation.

Both permittivity and permeability depend on frequency of the electromagnetic waves. Hence, different colours travel at different speeds. This is called dispersion. One result of dispersion is rainbows!



◀ **Figure 5.41** The electromagnetic spectrum

You saw in Chapter 2 of Nelson Physics Units 1 & 2 for the Australian Curriculum that heat can be transferred by radiation. This radiation is made of electromagnetic waves, often in the infrared part of the spectrum.

When we look at quantum physics we shall see why objects radiate electromagnetic waves.

Scientific literacy: Australian Synchrotron – seeing the invisible

The Australian Synchrotron is a landmark national research facility that enables the study of materials and biological processes at the nanoscale.

The synchrotron was built next to Monash University in Melbourne. It took five years to build, after several years of planning. The synchrotron cost about $220 million and was largely funded by the Victorian and Commonwealth governments. Various organisations contributed funding for beamlines including Monash and Melbourne Universities, CSIRO (Commonwealth Scientific and Industrial Research Organisation) and ANSTO (Australian Nuclear Science and Technology Organisation).

Courtesy of Australian Synchrotron

▲ Figure 5.42 The Australian Synchrotron

The synchrotron is a circular machine, about the size of a football field, that produces a light source one million times brighter than the Sun. It accelerates electrons to close to the speed of light, under conditions of ultra-high vacuum, and directs their path with powerful electromagnets. The electrons are made to turn some 28 corners as they travel around the 'ring'. As they change path the electrons lose energy, which is given out as synchrotron light. This light is then directed along beamlines to experimental end stations, where it is used in various investigative techniques.

How it works

The metal cathode in an electron gun is heated to ~1000°C and made to emit bunches of electrons in a process known as thermionic emission. Electrons leaving the cathode's surface are forced towards an anode by the electric field that exists between the cathode and the anode. Electrons leave the gun with 90 keV of kinetic energy, and travel at a speed of 59 per cent of the speed of light.

A linear accelerator using radio frequency (RF) electric fields is used to further increase their kinetic energy to 100 MeV.

The electrons next enter the booster ring, where a series of powerful electromagnets are used to force the electrons around a closed loop. The electron energy is again increased, so that each time the electrons complete one lap of the ring, their energy is increased. This results in an increase in the radius of the electron path, so the magnetic field is simultaneously increased to ensure that the electrons return to the RF cavity at the right time for the correct accelerating field to be applied.

AUSTRALIAN SYNCHROTRON

Learn more about the Australian Synchrotron and the experiments being carried out on its beamlines.

Courtesy of Australian Synchrotron

▲ Figure 5.43 Inside the Australian Synchrotron

The final destination of the electrons is the storage ring. Powerful electromagnets are used to steer the electrons around a closed loop – they move at a speed equivalent to travelling 7.5 times around Earth each second. However, a charge that undergoes acceleration also emits electromagnetic radiation, so the storage ring at synchrotron light source serves as a source of intense beams of light, or synchrotron radiation, which ranges from infrared to hard X-rays.

How researchers use the light source

The light is channelled down beamlines to end stations, where researchers measure the interaction of light with their sample. The interactions that occur include X-ray scattering or diffraction, X-ray fluorescence, and X-ray and infrared absorption.

Synchrotron techniques are used worldwide in many important areas, including understanding the capabilities of advanced materials such as carbon nanotubes and graphene – a highly conductive material consisting of a single layer of carbon atoms that could transform a multitude of industries from electronics to renewable energy. Agriculture, biomedics, environmental health, food technology, oil and gas, mining and nanotechnology also have synchrotron science to thank for major innovations in their respective fields.

Text courtesy of the Australian Synchrotron

Questions

1 Explain how charges that have undergone circular acceleration in the synchrotron's storage ring produce electromagnetic radiation.

2 Describe the basic design of an electromagnet and how its magnetic field can be modified.

3 Research and describe the five main benefits of X-rays produced by synchrotron light sources over other conventional laboratory-based sources.

4 X-ray fluorescence is an analytical technique available at both synchrotron light sources and using conventional laboratory-based instruments. Investigate and discuss the technique, what it measures, and the benefits of using synchrotron sources to perform these studies.

5 The Imaging and Medical Beamline (IMBL) at the Australian Synchrotron can produce microbeams of X-rays for radiotherapy that are 25 µm in size. Discuss the benefits of using a microbeam for radiation therapy treatment over conventional radiotherapy treatments.

Making and detecting electromagnetic waves

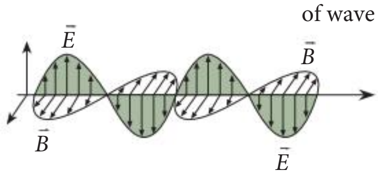
Stationary charges and steady currents cannot produce electromagnetic waves but, whenever the current in a wire changes with time, the wire emits electromagnetic radiation. Whenever a charged particle accelerates, it radiates energy in the form of electromagnetic waves.

A transmitting antenna works by being connected to an AC. This current causes charges to oscillate back and forth along the length of the antenna. This creates a magnetic field. The magnetic field is perpendicular to the axis of the antenna and forms loops around the antenna. Recall the right-hand rule from Chapter 4.

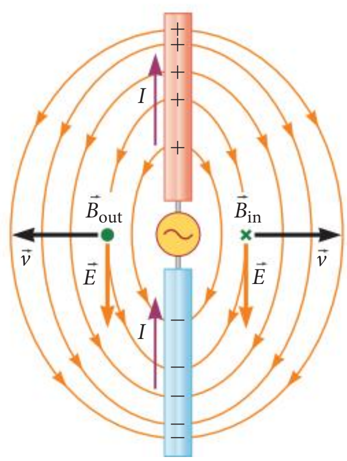
The current oscillates with some frequency, f. The magnetic field also oscillates with this frequency. This time-varying magnetic field creates a time-varying electric field. This electric field also oscillates with frequency f, at right angles to the magnetic field. These two oscillating fields are the electromagnetic wave. The electromagnetic wave travels in a direction perpendicular to both fields that make up the wave, which are themselves mutually perpendicular. This is shown in Figure 5.44. Figure 5.45 shows how such waves are produced by an antenna. The frequency of the wave is f, the same as the current, and it travels at speed c in vacuum or very close to this in air; hence it has wavelength $\lambda = \frac{c}{f}$.

When this travelling electromagnetic wave is incident on free charges, such as electrons in a receiving antenna, it will make them oscillate. The electric field applies a force to the electrons, and an AC with the same frequency, f, is produced. This current can be picked up and converted into another form. For example, a transducer attached to the antenna can convert the current into sound waves. Hence, an antenna can either transmit electromagnetic waves if it is connected to an AC, or receive them and produce an AC. This is how radio, TV and mobile phone antennae work. The length of the antenna depends on the frequency of the waves it must transmit or receive and the current it must carry. Longer antennae are used for lower frequencies.

Direction of motion of wave



▲ Figure 5.44 The electric and magnetic fields in an electromagnetic wave are perpendicular.



▲ Figure 5.45 An antenna produces an electromagnetic wave.

WORKED EXAMPLE 5.7

An antenna for a radio station transmits a signal with frequency of 106.3 MHz. What is the wavelength of this electromagnetic wave? (4 marks)

Answer	Logic	
$c = f\lambda$	Use the dispersion relation to relate frequency to wavelength.	1 mark
$\lambda = \frac{c}{f}$	Rearrange for wavelength.	1 mark
$\lambda = \frac{3.00 \times 10^8\ \text{m s}^{-1}}{106.3 \times 10^6\ \text{Hz}}$	Substitute the correct values with units.	1 mark
$\lambda = 2.82\ \text{m}$	Calculate the final value.	1 mark

Try this yourself

A TV station broadcasts a signal with a wavelength of 0.28 m. What frequency does this correspond to? (3 marks)

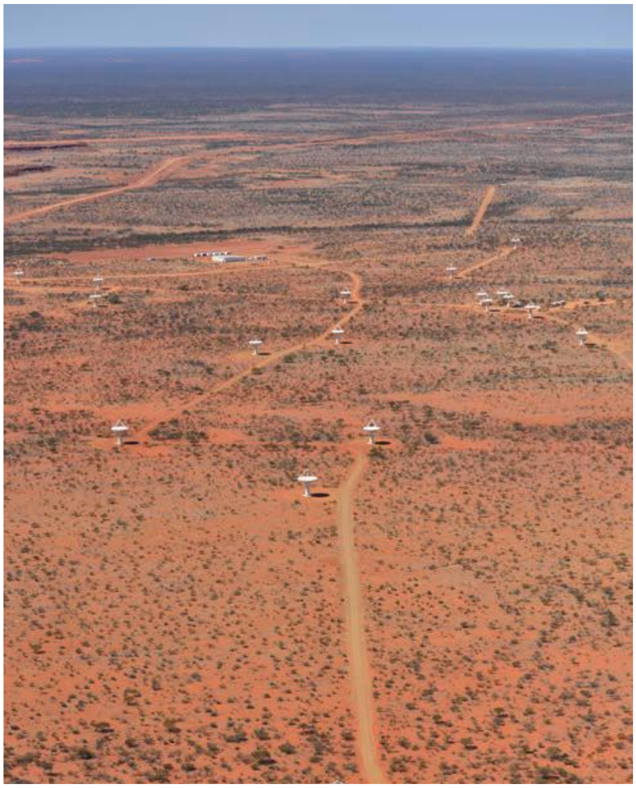
Case study

Dr Maria Cunningham and the Square Kilometre Array (SKA)

Dr Maria Cunningham is an astrophysicist at the University of New South Wales. She works in a field called *astrobiology*.

▲ **Figure 5.46** Dr Maria Cunningham

At school, Dr Cunningham's interest in mathematics was encouraged, but her interest in physics puzzled people. She had various jobs after finishing school before pursuing physics research. She started studying mathematics and physics as a hobby while breastfeeding her children. As she says, 'it takes a lot of time, but you do have a hand free'. She has combined bringing up four children with a PhD in astrophysics and a career that she loves.


CSIRO

▲ **Figure 5.47** CSIRO's Australian SKA Pathfinder telescope at the Murchison Radio-astronomy Observatory in Western Australia will host the core of the SKA.

Dr Cunningham uses large radio telescopes to search for evidence of organic molecules in space. These are detected by the characteristic spectra that they produce. Some of the organic molecules that have already been found in space include methanol (CH_3OH, found in methylated spirits), ethanol, acetic acid or vinegar (CH_3COOH), and the simple sugar ribose.

Dr Cunningham use large radio telescopes, such as the Australia Telescope Compact Array at Narrabri in Australia, the Atacama Large Millimeter Array in the desert in the North of Chile (al norte de Chile), and the Murchison Wide Field Array in the desert and radio quiet zone of Western Australia, inland from Geraldton. These giant telescopes can tune in to the frequency of any molecule. Dr Cunningham uses these telescopes to 'listen' for molecules.

She hopes to find amino acids and more complicated sugars. These are the building blocks for simple living organisms – sugar for food, and amino acids to build the protein structure of organisms. Finding these in space would be particularly exciting because there is a good chance that these complex organic molecules can only form in space. Maybe life comes from outer space via comets and meteorites!

The $1.9 billion Square Kilometre Array (SKA) is a gigantic array of radio telescopes, receiving dishes and antennae funded by Australia, Canada, China, Germany, Italy, New Zealand, South Africa, Sweden, the Netherlands and the United Kingdom. It is to be built mostly in Australia and South Africa. It is a massive international project, far beyond the scope of any individual nation.

The SKA will be at least 10 times more sensitive than any existing radio telescope and be able to survey the sky 10 000 times faster. It will collect massive amounts of data, more than the current global internet load! Development of hardware and software, including algorithms for data compression to deal with this massive amount of data, is one of the major challenges of the SKA project. The World Wide Web started as a way of transferring and linking information at the European Organization for Nuclear Research (CERN); only time will tell what might grow from the developments for the data processing at the SKA.

Astronomers and astrophysicists from all over the world will use the SKA to investigate cosmic magnetism and 'dark matter' and test the general theory of relativity. The SKA will answer fundamental questions about how the universe formed, how it is evolving, and what will happen to it. It may also answer Dr Cunningham's questions about how life began.

THE SQUARE KILOMETRE ARRAY

Discover some amazing facts about what will be the world's largest radiotelescope.

Questions

1 What is astrobiology?
2 Explain briefly what a radiotelescope does.
3 Why do you think the large telescopes that Dr Cunningham uses are typically built in the desert?
4 The SKA will be able to detect electromagnetic waves with frequencies between 70 MHz and 30 GHz. What wavelengths do these correspond to?
5 Why is it necessary for the SKA project to be run by such a large international collaboration? Consider what positive aspects there might be to such a large collaboration. What negative aspects can you think of?
6 Scientists often use posters to present their research at conferences. Research one of the projects that will use the SKA and prepare a poster describing the project. Include the hypothesis to be tested, the sort of data that will be used, and why this research is important.

QUESTION SET 5.5

Remembering

1 Who unified all the information about electromagnetic phenomena into a single theory, described by four equations?
2 Define 'dispersion relation'. Write down the dispersion relation for light.

Understanding

3 If Figure 5.45 represents a transmitting antenna used by a radio station, should you have your car radio antenna orientated horizontally or vertically to listen to this station?
4 When light (or other electromagnetic radiation) travels across a given region, what:
  a oscillates?
  b is transported?
5 What does a radio wave do to the charges in the receiving antenna to provide a signal for your car radio?

Applying

6 The red light emitted by a helium-neon laser has a wavelength of 633 nm.
  a What is the frequency of the light waves?
  b How long would it take for a signal from this laser to travel from the Earth to the Moon and back?

Analysing

7 What experiments can be done to show that light is a wave? Give at least three examples. Hint: Think about what you learnt about wave behaviour in Unit two of *Nelson Physics Units 1 & 2 for the Australian Curriculum*.
8 The human eye is most sensitive to light with a frequency of 5.45×10^{14} Hz, which is in the green-yellow region of the visible electromagnetic spectrum.
  a What is the wavelength of this light?
  b Why do you think the eye is most sensitive to this wavelength?

Reflecting

9 Mark on an electromagnetic spectrum the wavelengths you have used today.

CHAPTER SUMMARY

- An electric field that changes in time creates a magnetic field.
- A magnetic field that changes in time creates an induced electric field.
- A changing magnetic field creates an induced emf, $\varepsilon = -\frac{\Delta\Phi}{\Delta t}$, where $\Phi = \vec{B} \cdot \vec{A}$ is the magnetic flux. This is Faraday's law.
- If there are free charge carriers, an induced emf will cause an induced current to flow.
- An induced current flows in a direction such that it acts to oppose the change in magnetic flux that created it. This is Lenz's law and it is a consequence of energy conservation.
- A generator converts kinetic energy into electric potential energy.
- A generator uses the relative movement of coils of wire and magnets to induce an emf across the coils, to drive a current.
- The maximum emf produced by an AC generator is $\varepsilon_{max} = 2\pi fnBA$.
- AC potential differences and currents are described using *rms* values, where $V_{rms} = \frac{1}{\sqrt{2}}V_{peak}$.
- A transformer acts to change the magnitude of an AC potential difference, either increasing it (step-up) or decreasing it (step-down).
- A transformer consists of two coils of wire arranged such that an alternating current in the primary coil can induce a current in the secondary coil, usually via a ferromagnetic core.
- The emfs, currents and number of turns in the coils of a transformer are related by $\frac{V_s}{V_p} = \frac{I_p}{I_s} = \frac{N_s}{N_p}$.
- A motor converts electric potential energy into kinetic energy.
- A motor uses a current flowing through a coil in a magnetic field to produce forces on the coil.
- The forces on the coil of a motor add to zero but give a net torque on the coil, causing it to rotate.
- A DC motor uses a commutator to change the direction of the current to keep the torque in a constant direction.
- An AC induction motor uses a time-varying magnetic field to create the current in the moving coils, so no electrical connection to the coils is needed.
- Electromagnetic waves consist of coupled, self-sustaining, oscillating electric and magnetic fields.
- Electromagnetic waves are produced by oscillating charges.
- An electromagnetic wave has electric and magnetic fields perpendicular to each other and to the direction of propagation.
- The electric and magnetic fields obey a wave equation, and have speed in vacuum given by $c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}} = 3.00 \times 10^8 \text{ m s}^{-1}$.
- The speed is related to the wavelength and frequency by the dispersion relation $c = f\lambda$.

CHAPTER GLOSSARY

alternating current (AC) a current that varies with time between positive and negative values, usually sinusoidally

armature the frame of the rotating part of a motor or generator, holding one or more coils

back emf an induced emf that opposes the flow of current in a circuit, particularly in a coil of a motor

commutator a device for reversing the direction of current flow in a motor or generator

diode a circuit component that allows current to flow in only one direction

direct current (DC) a current that flows in a single direction

dispersion relation the relationship between the frequency of a wave and its wavelength

eddy current a circular current induced in a conductor due to a changing magnetic field

electromagnetic induction the production of an emf due to a time-varying magnetic field

electromagnetic wave coupled oscillating electric and magnetic fields, which in vacuum form a transverse wave that propagates at speed *c*

generator a device that converts kinetic energy to electric potential energy and produces a current

induced current a current created by a changing magnetic field

induced emf an emf created by a changing magnetic field

induction motor an AC motor in which the torque is due to a current induced by the changing magnetic field created by an AC current in a coil

magnetic braking braking due to the interaction of eddy currents and an external magnetic field

magnetic flux the magnetic field passing through a given area, $\Phi = BA\cos\theta$

motor a device that converts electric potential energy into kinetic energy, usually rotational kinetic energy

rectifier a circuit, typically consisting of diodes, that converts AC to DC

root mean square (rms) the average AC potential difference or current that produces the same power in a load as a DC potential difference or current of the same magnitude

rotor the rotating part, typically coils, of a motor or generator

stator the stationary or non-rotating coils of a motor or generator

step-down transformer a transformer with a lower output potential difference than the input potential difference

step-up transformer a transformer with a higher output potential difference than the input potential difference

transformer a device for changing the magnitude of an AC potential difference

vector dot product the scalar product of two vectors, $C = \vec{A} \cdot \vec{B} = AB\cos\theta$

wave equation a differential equation that describes wave behaviour; its solutions are wave functions that are typically sinusoids

CHAPTER REVIEW QUESTIONS

Remembering

1 What is the function of a rectifier circuit?

2 What is an induced current?

3 Why does an AC induction motor require less maintenance than a DC motor?

4 Name the main components of a generator.

5 How do transmitting and receiving antennas differ?

Understanding

6 Express the unit Wb in fundamental units.

7 Explain the difference between AC and DC generators. Sketch the voltage output as a function of time for both types.

8 Briefly describe how a transmitting antenna works.

9 Why is it necessary for the direction of current in the coil of a DC motor to be reversed every half turn?

10 How does a motor differ from a generator and in what ways are they similar?

11 A battery is connected to the primary coil of a transformer. Why is no current observed in the secondary coil?

12 Explain the difference between V_P, V_{P-P}, V_{ave} and V_{rms} for an AC potential difference. Draw a diagram to help explain your answer.

Applying

13 A generator produces an output current of $I(t) = 3.0A \sin(6.2\ s^{-1}\ t)$. For this generator, what is:

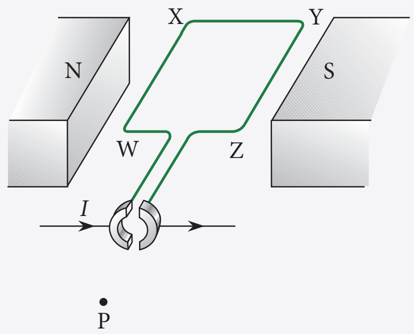
a I_P?

b I_{P-P}?

c I_{ave}?

d I_{rms}?

14 Figure 5.48 shows the basic features of a small DC electric motor. WXYZ is the rotating coil, connected to a DC battery. The direction of the magnetic field is parallel to the plane of the coil. The magnetic field has magnitude of 0.48 T and the coil has side lengths 0.080 m. A current of 2.0 A flows in the coil.



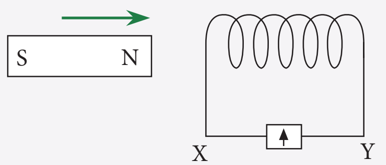
◀ **Figure 5.48**

a With the coil in the position shown in Figure 5.48, and looking from point P, will the coil rotate clockwise or anticlockwise?

b What is the maximum force on the side ZY?

c What is the torque on the coil?

15 Figure 5.49 shows the north pole of a bar magnet being pushed into a solenoid. For each of the following situations, state whether the current flows through the meter from X to Y or from Y to X. If no current flows, write 'N'.



◀ **Figure 5.49**

a The north pole is moved towards the solenoid.

b The magnet is held stationary in the solenoid.

c The north pole of the magnet is withdrawn from the solenoid.

d The south pole of the magnet is moved towards the solenoid.

e The magnet is held stationary and the solenoid is moved to the right.

f Both the magnet and the solenoid are moved to the right with the same speed.

16 The magnetic flux through a wire coil of area 2.0×10^{-3} m^2 is 4.8×10^{-4} Wb. What is the magnetic field (flux density) perpendicular to the plane of the coil?

17 A rectangular coil of length 15 cm and width 12 cm, having 60 turns, rotates in air about its axis at 1200 revolutions per minute in a uniform magnetic field of flux density 0.40 T.

a What is the emf when the plane of the coil is:

i perpendicular to the magnetic field?

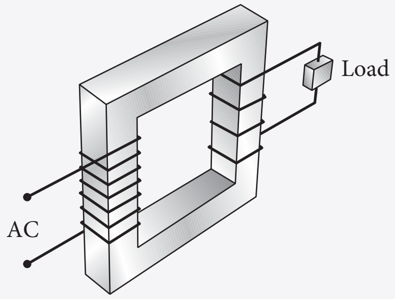
ii parallel to the magnetic field?

b Draw a graph to show the time variation of the emf over an interval of one-twentieth of a second.

c What is the magnitude of the peak emf?

d What is the rms potential difference?

18 Figure 5.50 illustrates a step-down transformer with a primary coil of 500 turns and a secondary coil of 250 turns. The primary coil is connected to a 240 V AC mains supply.



◀ Figure 5.50

- a What is the secondary potential difference?
- b Explain how the transformer operates.
- c If the primary current is 2.4 A, what is the current in the secondary coil?

19 A generator is manufactured to produce an alternating emf of peak value 340 V when the coil rotates at 3000 revolutions per minute in a uniform magnetic field.

- a Why would this be a desirable output for a generator?

The magnetic field strength in which the coil rotates is 0.25 T and the coil has a cross-sectional area of 0.040 m^2.

- b How many turns does the coil have?

20 A mobile phone transmits a signal with frequency 1800 MHz.

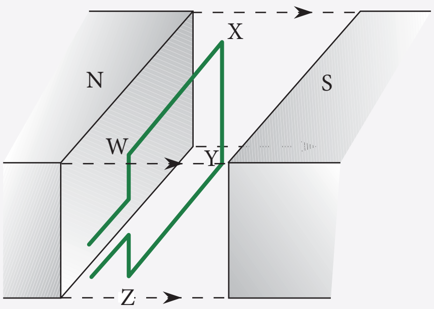
- a What wavelength is this signal?
- b How long does it take for this signal to travel from Perth to Sydney, about 3300 km?
- c Why does it actually take longer than this for a signal to get from a mobile phone in Perth to one in Sydney?

Analysing

21 a Explain the difference between AC and DC motors.

- b If you were designing a remote control car, what sort of motor would you use and why?
- c If you were designing a hair dryer, what sort of motor would you use and why?

22 Figure 5.51 shows a rectangular coil of wire WXYZ in a uniform magnetic field of magnitude 0.06 T. The coil has 160 turns. The area of each turn is 0.025 m^2. A student moves the coil vertically upwards with a uniform speed. The entire coil is in the magnetic field for 0.3 s and leaves the field between 0.3 and 0.5 s.



◀ Figure 5.51

- a Draw a graph of flux through the coil as a function of time. Label your axes clearly, including units.
- b Draw a graph of the potential difference induced across the ends of the coil with time.
- c What is the magnitude of the potential difference produced across the ends of the coil for the time interval from 0.3 to 0.5 s?

23 If the coil shown in Figure 5.51 and described in question 22 has a current of 0.20 A running through it in the direction from W to X, and is in the position shown in Figure 5.51, what is:

a the net force acting on the coil?

b the net torque acting in the coil?

c the direction of rotation of the coil?

If the coil is turned through 90° so that it is parallel with the field and side WX is on the left, what is:

d the net force acting on the coil?

e the net torque acting in the coil?

f the direction of rotation of the coil?

Reflecting

24 Draw a concept map showing how the various ideas, applications and principles covered in this chapter and the previous chapter are related. Express these in words. Then add all the key equations to your concept map. Often these provide the links between concepts.

25 Evaluate the importance of Hertz's 1887 experiment.

UNIT 4
REVOLUTIONS IN MODERN PHYSICS

Shutterstock.com/Lucy Baldwin

CHAPTER 6 EINSTEIN'S SPECIAL RELATIVITY

Shutterstock.com/andrey_l

By the end of this chapter you will have covered the following material.

Science Understanding

- Observations of objects travelling at very high speeds cannot be explained by Newtonian physics (for example, the dilated half-life of high-speed muons created in the upper atmosphere, and the momentum of high-speed particles in particle accelerators) (ACSPH129)
- Einstein's special theory of relativity predicts significantly different results to those of Newtonian physics for velocities approaching the speed of light (ACSPH130)
- The special theory of relativity is based on two postulates: that the speed of light in a vacuum is an absolute constant, and that all inertial reference frames are equivalent (ACSPH131)
- Motion can only be measured relative to an observer; length and time are relative quantities that depend on the observer's frame of reference (ACSPH132)
- Relativistic momentum increases at high relative speed and prevents an object from reaching the speed of light (ACSPH133)
- The concept of mass–energy equivalence emerged from the special theory of relativity and explains the source of the energy produced in nuclear reactions (ACSPH134)

Introduction

Towards the close of the 19th century, most scientists were quite satisfied with the development of mathematical models to explain the results of experiments and observations of the natural world. Newton's earlier work on motion had been thoroughly tested and found to be valid in all circumstances. Scientists felt that they understood interactions between matter and energy quite well.

However, within the space of 30 years there would be a total transformation of these ideas. Small but significant discrepancies were beginning to exist between predictions based on Newton's equations of motion and experimental results.

Refinements to the theory of electromagnetism by James Maxwell, which Hertz later successfully tested by experiment, was one trigger for this change.

The man who had the most impact was Albert Einstein, a German-speaking Swiss patent clerk. He used powerful ideas based on mathematical modelling to convince scientists about a different way of looking at the nature of time, space, matter and energy. This was prior to the period when experiments were able to confirm these ideas or theories; however, measurements of the speed of light by Alfred Michelson and Edward Morley at the end of 19th century were useful in the acceptance of the theory.

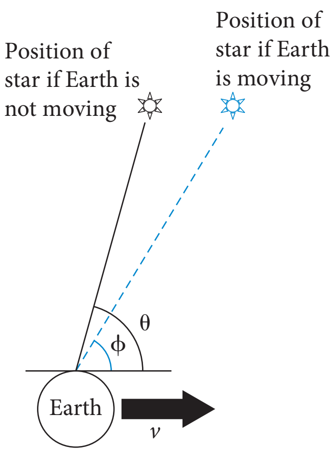


Figure 6.1 ▲ A stellar aberration. A star appears at a different angle to that expected. This is caused by the movement of Earth relative to the speed of light.

The speed of light in a vacuum is 299 792 458 $\mathrm{m\,s^{-1}}$ (approximately $3.0 \times 10^8\ \mathrm{m\,s^{-1}}$). Recently reported studies have suggested that the speed of light has increased. If the speed of light turns out not to be constant, then Einstein's theory will come under pressure and may be superseded.

The nature of light

Many early scientists conjectured that light had a finite speed. This was first demonstrated by Roemer at the end of the 17th century. It was finally confirmed by the astronomer James Bradley in 1728. Bradley measured the difference between where a star was expected to be seen and where it was actually observed. This **stellar aberration** was due to the finite speed of light and the speed of Earth combining to affect the position from which the starlight appeared to come.

He determined the speed to be about 298 000 $\mathrm{km\,s^{-1}}$ (in modern units). This value was later refined in land-based experiments by Fizeau (1849), Foucault (1852) and Michelson (1879, 1883 and 1926). The standard, constant speed of light is now 299 792 458 $\mathrm{m\,s^{-1}}$.

In 1801, Thomas Young demonstrated the effects of interference for light. This indicated that light had wave-like properties. Young's method was used to determine the wavelengths of visible light.

Later, in the same century, James Clerk Maxwell used calculus mathematics to bring all the concepts of electricity and magnetism neatly together into four simple relationships. As well as being a triumph of theoretical physics, it also began the transformation of our understanding of light and matter.

WOW

James Clerk Maxwell (1831–79)

Maxwell is a giant of physics. Born in Scotland, he was educated at the University of Edinburgh, and Trinity College, Cambridge. It was at Cambridge that he began his most important work. In statistical thermodynamics he developed the Maxwell distribution, which correctly predicted the range of molecular speeds in a gas. In optics, he produced the first colour photograph. He improved our understanding of colour perception and its link to colour deficiency. In astronomy, Maxwell used mathematical modelling to show that the rings of Saturn were most likely made of small rock particles.

Maxwell's greatest work was the unification of electric and magnetic theory into electromagnetism. Einstein's reflections on Maxwell's work led to the theory of relativity.

Table 6.1 Maxwell's equations

$\oint \vec{E}.d\vec{A} = \frac{q}{\varepsilon_0}$	From Coulomb's law, relating electric fields (E) to electric charge (q)
$\oint \vec{B}.d\vec{A} = 0$	From Guass's law for magnetism, indicating that there are no single magnetic poles
$\oint \vec{E}.d\vec{s} = -\frac{d\Phi_B}{dt}$	From Faraday's law, in which an electric field is produced by a changing magnetic field
$\oint \vec{B}.d\vec{s} = \mu_0\varepsilon_0\frac{d\Phi_E}{dt} + \mu_0 I$ or $\oint \vec{B}.d\vec{s} = \frac{1}{v^2}\frac{d\Phi_E}{dt} + \mu_0 I$	An extension by Maxwell to Ampère's law, in which a magnetic field ($\vec{B}$) is produced by electric current (I) or changing electric field ($\vec{E}$). Note that v represents the speed of the radiating electric and magnetic fields
Note: ε_0 is a constant, the permittivity of free space, and μ_0 is a constant, the permeability of free space.	

SPEED OF LIGHT MAY HAVE CHANGED

This article looks at some of the debate arising from the discovery that the speed of light may have been lower as recently as two billion years ago.

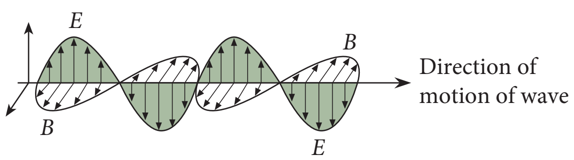


Figure 6.2
A representation of an electromagnetic wave travelling through space. Time-varying electric and magnetic fields (E and B respectively), perpendicular to each other, give rise to an electromagnetic wave that travels perpendicular to both.

Maxwell's equations (Table 6.1) were seen as a great triumph of theoretical physics because they unified two fields – electricity and magnetism. They are as important to electromagnetism as Newton's laws of motion and the law of universal gravitation are to mechanics. Maxwell realised that his latter two equations meant that an electric field produced by the varying magnetic field would itself be varying. It would therefore produce a varying magnetic field, and so on.

Maxwell deduced that these waves could move through space with a fixed velocity, v. The velocity of the electromagnetic wave was only dependent on the constants of **magnetic permeability** (μ_0) and **electrical permittivity** (ε_0). Using modern values, this turns out to be:

$$v = \frac{1}{\sqrt{\mu_0\varepsilon_0}} = \frac{1}{\sqrt{(8.85\times10^{-12})(4\pi\times10^{-7})}}$$
$$= 2.99\times10^{8}\ \text{m s}^{-1}$$

The speed of electromagnetic radiation in a vacuum was the same as that of light. This result, Maxwell argued, suggested two things. First, visible light must be like an electromagnetic wave. Second, the speed of light depends only on the medium. Initially, scientists were sceptical, but it was not long before Maxwell's view was accepted because of the strength of his mathematical arguments and the accumulation of evidence.

The speed of electromagnetic waves and light depend only on properties of the medium:

electrical permittivity, ε, and magnetic permeability, μ

In a vacuum, the speed of electromagnetic waves and the speed of light are the same:

$$c = \frac{1}{\sqrt{\mu_0\varepsilon_0}}$$

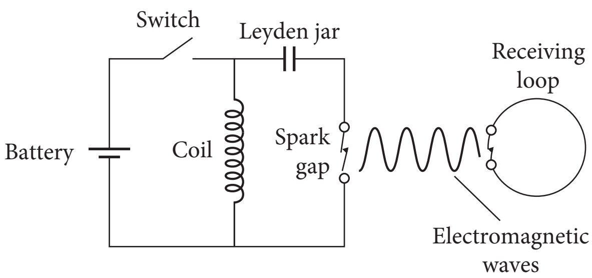
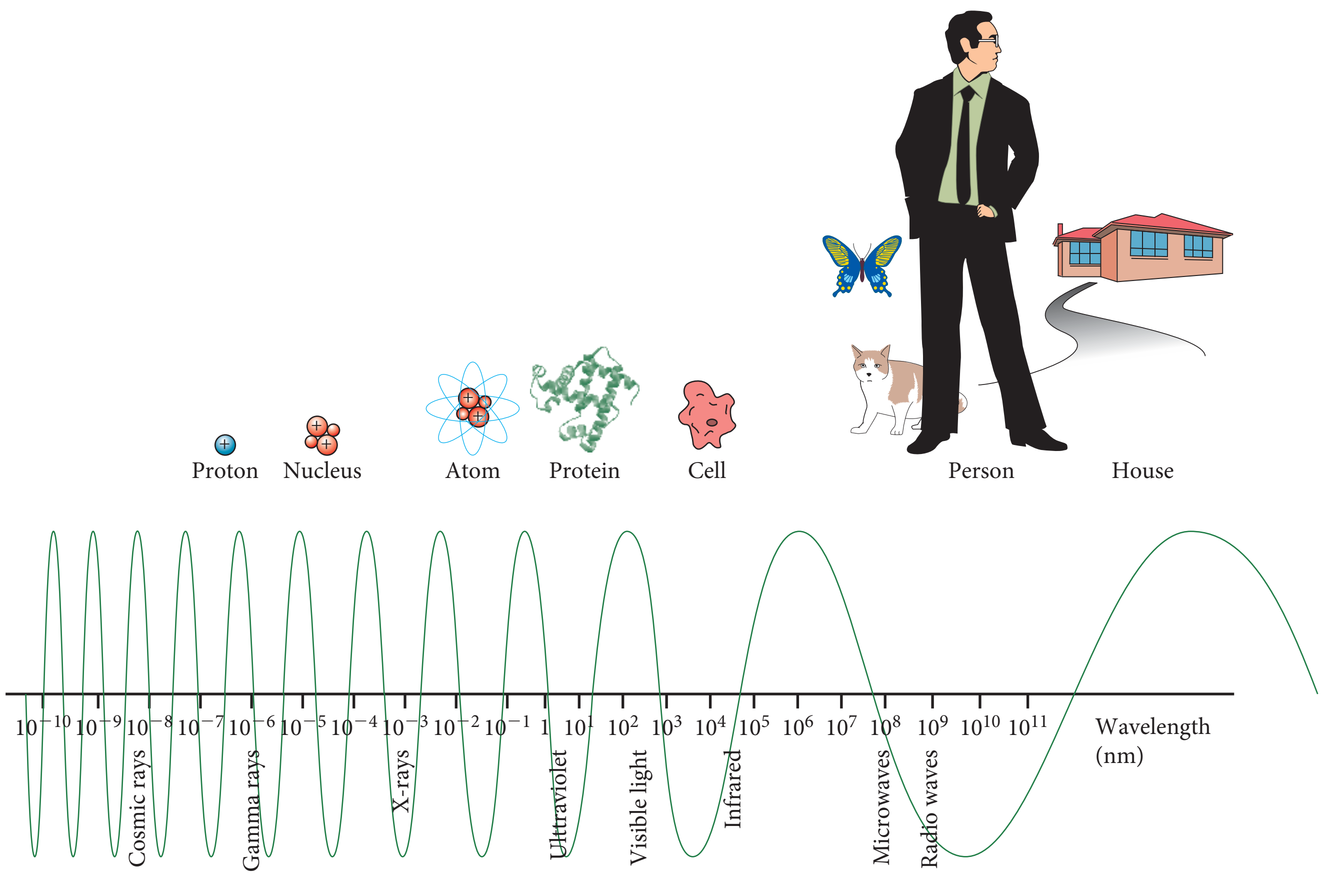


Figure 6.3 ▲ The circuit in Hertz's equipment caused a spark to generate an electromagnetic wave that was received by a simple antenna.

Electromagnetic radiation

The first experiment to confirm the presence of these electromagnetic waves was conducted by Heinrich Hertz in 1886, 7 years after Maxwell's death. Hertz used a high-voltage spark gap to produce some electromagnetic waves that were detected a few metres away. He later showed that these waves could be reflected and refracted, and had a speed of $3.0 \times 10^8\,\mathrm{m\,s^{-1}}$. This was experimental confirmation of predictions based on Maxwell's equations.

We now know that, as well as visible light, the electromagnetic wave spectrum consists of radio waves, microwaves, infrared light, ultraviolet light, X-rays and gamma radiation. They are all produced by the acceleration of charged particles, as predicted by Maxwell's equations.



▲**Figure 6.4** Electromagnetic spectrum

QUESTION SET 6.1

Remembering

1 Name one contribution to the understanding of light by these scientists:
- a Roemer
- b Bradley
- c Young

2 On what does the speed of electromagnetic radiation depend?

Understanding

3 What did Maxwell achieve in his famous equations?

4 Draw a three-dimensional representation of an electromagnetic wave. Describe what it means.

Applying

5 Describe Hertz's experiment. Why was this so important?

Analysing

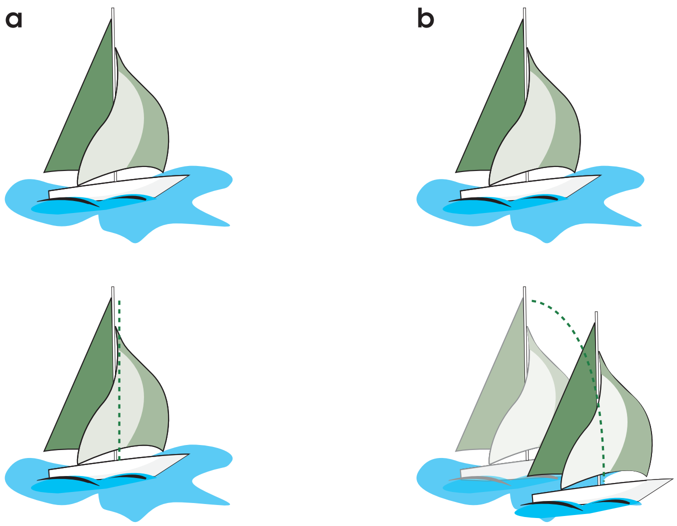
6 How could electromagnetic waves be used to accelerate a small bunch of electrons? (Hint: Recall what you learnt about antennae in Chapter 5.)

Reflecting

7 How does scientific knowledge develop and change? In your answer refer to the speed of light and Maxwell's work.

Frames of reference and relativity

Galileo was one of the earliest thinkers to comment on events observed in different frames of reference. In his book, *Dialogues on the Two Chief Systems of the World*, Galileo described a thought experiment in which a sailor drops an object from the mast of a sailing ship moving at steady velocity. He asked the question: 'Where would the object land relative to the deck of the ship?' In his frame of reference the sailor would see the object fall straight down parallel to the mast; however, a nearby observer who is on land would see from his frame of reference that the object would follow a parabolic path.



◀ **Figure 6.5** Path of a falling object in the reference frame of a) a sailing ship and b) Earth

Later, Newton would also agree with Galileo. He described frames of reference that were stationary or moving at constant velocity as **inertial frames of reference**. They are not accelerating. An inertial frame of reference is an ideal situation in which objects at rest remain at rest and objects travelling at a constant velocity remain travelling at constant velocity. Inertial reference frames could include a spaceship, a table and a cruising aeroplane. Non-inertial frames of reference include merry-go-rounds and aeroplanes taking off or landing.

An inertial frame is one in which Newton's first law applies to a very good approximation. Any departures from the law are negligible.

Galilean transformations

Classical physics, the physics of Galileo and Newton, relies on the 'sensible' idea that inertial coordinate systems are equivalent. That is, there is a set of translation rules to connect measurements in one frame of reference or coordinate system to any other reference frame or coordinate system.

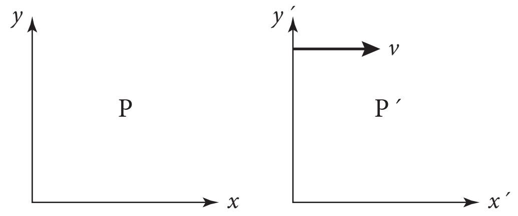


Figure 6.6 ▲ Two inertial frames that are moving relative to each other. One frame is regarded as stationary.

Let us analyse this assumption for the case of two frames of reference in two dimensions. The 'privileged' frame of reference, P, has coordinates, (x, y). It is stationary with respect to a moving frame of reference, P′, with coordinates, (x', y'). The two coordinates can be made to coincide at the same time, but then P′ moves further and further away from P. This motion depends on the relative speed, v, of P′ with respect to P and the time elapsed, Δt, after the two frames coincided.

The time interval in each frame of reference is the same. That is, for classical physics, time intervals are measured in the same way and occur at the same rate. Clocks in all frames of reference are identical in their timekeeping. Time is **invariant**. Thus, Δt in P and $\Delta t'$ in P′ are equal – we shall denote the time interval as Δt.

Suppose an observer located at (0, 0) in frame P sees a rocket some distance y away. It is travelling parallel to the x axis at speed v. Initially (at $t = 0$) the rocket's coordinates in P′ coincide with those in P. But as the rocket travels in P′ its x' coordinates increase in the time interval, Δt, by $v\Delta t$. This is just the distance between the two coordinate systems. Thus, the x coordinates in P relate to the x' coordinates in P′ by the **Galilean transformation**:

$$x = x' + v\Delta t$$

Similar arguments can be used to show that if the rocket is travelling at speed, v, parallel to the y and y' axes:

$$y = y' + v\Delta t$$

The transformation equations rely on the relative speeds of the two frames of reference. If P′ had been chosen as the stationary (or privileged) frame, then P would be travelling at a negative speed with respect to P′. This needs to be taken into account in the Galilean transformation equations. Finally, if the rocket were travelling with vector velocity $\vec{v}$, the transformation equations in the x and y directions require the use of the x and y components of the velocity respectively.

For a two dimensional inertial frame of reference, P′ moving at constant speed with respect to another inertial frame, P:

$$x = x' + v\Delta t$$
$$y = y' + v\Delta t$$

WORKED EXAMPLE 6.1

A car travels at 12 m s^{-1} through an intersection. After 10 s what is the position of the car with respect to coordinates based on the:

a intersection? (2 marks)

b car? (1 mark)

Answers	Logic	
a $x = x' + v\Delta t$		
$\Rightarrow x = 0\text{ m} + 12\text{ m s}^{-1} \times 10\text{ s}$	Substitute the correct values with units.	1 mark
$\Rightarrow x = 120\text{ m}$	Calculate the answer.	1 mark

▶

b $x' = x - v\Delta t$

$\Rightarrow x' = 120\,\text{m} - 12\,\text{m s}^{-1} \times 10\,\text{s}$ — Recognise that the car is not moving with respect to the reference frame. — 1 mark

$\Rightarrow x' = 0\,\text{m}$

Try these yourself

A train passes through a station at $30\,\text{m s}^{-1}$. A person on the train walks towards the front of the train at $2\,\text{m s}^{-1}$. After 5.0 s, the person has moved a distance from the common origin on train and station. What are the coordinates with respect to:

a the station? (3 marks)

b the train? (3 marks)

Principles of classical relativity

As well as considering a sailor dropping an object from the mast of a sailing ship, Galileo also discussed the situation of a person walking within the cabin of a ship. If the sailing ship moves forwards at a velocity of $5\,\text{m s}^{-1}$, a person moving forwards at a velocity of $1\,\text{m s}^{-1}$ relative to the cabin will be moving forwards at a velocity of $6\,\text{m s}^{-1}$ relative to Earth. The position and velocity of the person is different in each frame of reference.

However, according to Galileo and Newton, acceleration of a body will be the same in each frame of reference, providing they are inertial frames. In the sailing ship's cabin, for example, the person may have accelerated from 0 to $1\,\text{m s}^{-1}$ in 0.5 s in the cabin, which is an acceleration of $2\,\text{m s}^{-2}$. From the perspective of a nearby land-based observer, the person will have accelerated from 5 to $6\,\text{m s}^{-1}$ in 0.5 s, again an acceleration of $2\,\text{m s}^{-2}$.

Galileo and Newton argued that the laws of motion are the same in all inertial frames of reference (**relativity principle**). It can be shown that the laws of motion, including those for energy and momentum, are the same in all inertial reference frames. Observers in different inertial frames would record different velocities, and therefore determine different values of energy and momentum. Nevertheless, they would agree that there had been no net change of either energy or momentum. Consequently, they would agree on conservation of energy and conservation of momentum.

This suggests that there is no privileged inertial frame. No inertial reference frame is better than any other. If you are on a train travelling at a steady velocity of $80\,\text{km h}^{-1}$ west across the Nullarbor Plain, it is quite valid for you to argue that, from your reference frame, you are stationary and Earth is moving at $80\,\text{km h}^{-1}$ east. Providing your ride is smooth and at steady velocity, there is no experiment you can perform to test whether you are moving or stationary.

Galileo wrote about this in his book, in which he discussed the example of a person observing a range of movements in the cabin of a sailing ship that was sailing steadily at constant velocity. He argued that the person would not be able to tell if the ship was moving or not. He believed that there is no absolute frame of reference against which the velocities of all other frames can be measured. In this he differed from Newton, who believed that Earth–Sun system could be considered to be the absolute frame of reference.

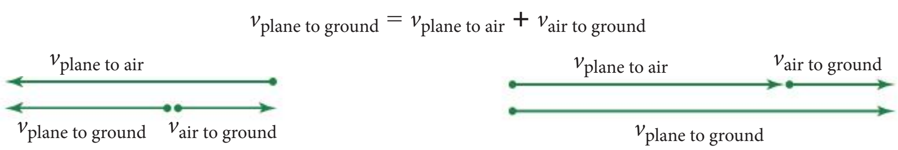
The laws of motion are the same in all inertial frames of reference.
- **The laws of conservation of energy and conservation of momentum apply in all inertial frames of reference.**
- **All inertial frames are equivalent. All are equally valid.**

Relative velocities

It is straightforward to calculate the velocity of an object in one inertial frame of reference compared to another inertial frame of reference, providing the object is travelling along the same direction, such as in the ship's cabin in the example discussed earlier. However, how do we generalise to two or even three dimensions? To do that we use vectors to compare the relative motions of an object in different inertial frames of reference.

Take the example of an aeroplane flying from Sydney to Perth. At the cruising altitude of passenger aircraft, about 10 000 m, there is nearly always a strong westerly wind, which varies from 50 km h^{-1} to more than 300 km h^{-1}. The direction of the aircraft's movement is almost exactly parallel to the direction of the wind. The westbound flight is scheduled at 4 hours 25 minutes (Figure 6.7(a)), while the eastbound flight is 3 hours 50 minutes (Figure 6.7(b)). The aeroplane's velocity, relative to the ground, is higher on the way to Sydney than on the way to Perth.

a **b**



▲**Figure 6.7** Vector addition of the relative velocities of an aeroplane flying from a) Sydney to Perth and b) Perth to Sydney

> **The velocity of A relative to B is the velocity of A relative to C plus (vector addition) the velocity of C relative to B:**
>
> $\vec{v}_{AB} = \vec{v}_{AC} + \vec{v}_{CB}$

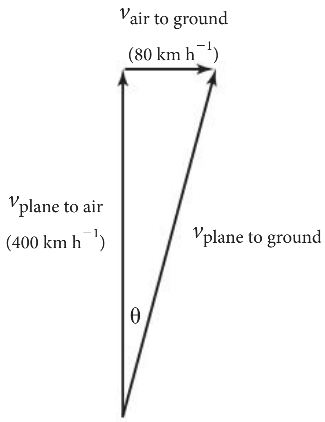
WORKED EXAMPLE 6.2

An aeroplane is headed due north at a speed of 400 km h^{-1}. There is a westerly wind (i.e. coming from the west) of 80 km h^{-1}.

a Draw a vector diagram to show the velocity of the plane relative to the ground. (2 marks)

b What is the speed of the plane relative to the ground? (1 mark)

c What is the resultant velocity of the plane relative to the ground? (2 marks)

Answers	Logic	
a  ▲**Figure 6.8**	Construct a correct vector diagram.	2 marks
b From Figure 6.8 we can use Pythagoras' theorem to find: $v^2 = (400\,\text{km h}^{-1})^2 + (80\,\text{km h}^{-1})^2$ $\Rightarrow v = 408\,\text{km h}^{-1}$	Calculate the answer.	1 mark

c $\tan\theta = \dfrac{80\text{ km h}^{-1}}{400\text{ km h}^{-1}} = 0.2$

$\Rightarrow \theta = 11.3°$	Find the correct angle.	1 mark
The velocity of the plane relative to the ground is $408\,\text{km h}^{-1}$, N11°E.	State the velocity with direction.	1 mark

Try these yourself

A person rows at $1.0\,\text{m s}^{-1}$ through the water of a river that is flowing at $0.5\,\text{m s}^{-1}$. The rower keeps the boat moving perpendicular to the bank.

a Draw a vector diagram to show the velocity of the rower from the reference frame of a person on the bank. (3 marks)

b Using your diagram, specify completely the velocity of the boat relative to the bank. (4 marks)

EXPERIMENT 6.1

RELATIVE MOTION AND FRAMES OF REFERENCE

There are two types of reference frame: inertial and non-inertial (accelerating). The path travelled by an object will be different in these two frames of reference.

Aim

To observe motion in different frames of reference

Materials

- several sheets of white paper
- carbon paper
- large metal ball bearing
- ramp
- potter's wheel, lazy susan or record turntable
- 2 dissecting boards or similar-sized flat surfaces
- tape
- video recording device
- ripple tank or frame to support video recorder

What are the risks in doing this experiment?	How can you manage these risks to stay safe
Large metal ball bearings can cause injuries to people.	Keep the ball bearing safely on the table and use only in the experiment.

In your write-up, add any more risks you can think of, as well as ways to manage them.

Procedure

Part A: Inertial frame of reference

1. Place the carbon paper face down on one white sheet of paper.
2. Place a sheet of graph paper (facing up) on top to form a sandwich. The graph paper establishes the grid system.
3. Tape the sandwich to one dissecting board (P), with the carbonised side down.

4 Abut a second dissecting board, Q, next to P.
5 Secure the ramp on Q.
6 Place the ripple tank over P and Q and secure the camera in position to view the arrangement.
7 Move P at constant speed relative to Q. (This may take a little practice.)
8 While P is moving at constant speed, release the ball bearing from the ramp. Use several different release heights.
9 Video the motion in each case.

Part B: Accelerating frame of reference

1 Place the carbon paper face down on one white sheet of paper.
2 Place a sheet of graph paper (facing up) on top to form a sandwich.
3 Cut the sandwich to fit the circular turntable.
4 Tape the sandwich to the turntable, with the carbonised side down.
5 Arrange the turntable and ramp so that the ball bearing rolls seamlessly onto the turntable.
6 Rotate the turntable at a constant rate.
7 Release the ball bearing from the ramp. Use several different release heights.
8 Video the motion in each case.

Results

Identify each path and relate it to the height of release of the ball bearing.

Analysis of results

1 From the video recording find, for the camera frame of reference, the speed of:
   a the ball bearing.
   b board Q.
2 Qualitatively compare the tracks for:
   a Part A.
   b Part B.
   c Part A relative to Part B.
3 For Part A:
   a establish coordinate systems for the camera frame (x, t) and the frame moving at constant speed (x', t').
   b produce data tables for both coordinate systems.
   c use the Galilean transformations to quantitatively compare the data from the two frames of reference.

Discussion

1 How well did your results for Part A demonstrate the Galilean transformations?
2 What differences were there between the observations in the inertial frame and those in the non-inertial frame?
3 How can the observations in Part B be reconciled with Newton's laws?

QUESTION SET 6.2

Remembering

1 Define 'inertial frame of reference'.
2 Write the Galilean relativity transformations.

Understanding

3 The path followed by an object dropped from the mast of a moving ship can appear to be both a straight line and a parabola. Explain why.
4 Can you tell if you are in an inertial frame of reference? Give a reason.

Applying

5 A bus is travelling at 12 m s^{-1} when a passenger starts to walk at 2.0 m s^{-1} towards the front. A stationary person outside the bus also observes the situation.

- a What is the speed of the person relative to the:
  - i bus?
  - ii person outside the bus?
- b After 5.0 s, what is the position of the person relative to the:
  - i bus?
  - ii person outside the bus?

6 The tide is running south at 3.0 m s^{-1}. At the same time, a yacht is steering at 4.0 m s^{-1} directly towards a buoy in the east. What is the velocity of the yacht relative to the water? Show your answer in vector form.

Analysing

7 Describe each of Galileo's thought experiments in his *Dialogues*, which were discussed above. How did he explain the observed motion?

Reflecting

8 Summarise Galilean relativity.

The aether and problems with classical relativity

The Galilean principle of relativity states that there is no absolute frame of reference and that all velocities are relative. This principle became known as classical relativity. This was the predominant view of physicists through the 18th and 19th centuries.

Maxwell's equations of electromagnetism presented a problem to this view of the laws of physics. His equations indicated that light acts like an electromagnetic wave, and that its speed is a constant in a particular medium, irrespective of the frame of reference of the observer or the source of light. Maxwell's equations make no mention of relativistic adjustments, only of two simple constants, the permeability and permittivity of free space.

Physicists in the mid-19th century thought that light required a medium, just like sound waves require air molecules as a vibrating material to move through. They called this medium the **aether**. It was thought to be a transparent weightless substance that enabled light or electromagnetic waves to vibrate through space. This aether was also thought to be fixed in space; the Sun and the planets moved relative to this medium. Could c be exactly 3.00×10^8 m s^{-1} in this aether and faster or slower in other reference frames, depending on the relative speed of the reference frame to this aether?

Would the presence of this aether allow electromagnetism to join with the laws of motion in conforming to classical relativity? It would depend on the experimental discovery of the aether and, if it existed, on its properties.

The Michelson–Morley experiment

In America, Albert Michelson had established a reputation for careful experimentation. By the early 1880s he had successfully measured the speed of light by using a rotating mirror to reflect pulses of light off a second distant mirror. Michelson related the angular rotation of the mirror to the total optical path to determine the speed.

In 1887, Michelson teamed up with Edward Morley to attempt to detect the presence of the aether by measuring the speed of light in different reference frames. Their apparatus consisted of a large stone block floating on a bath of mercury, so that it could be rotated smoothly. The apparatus had a number of mirrors designed to create two light paths at right angles to each other.

WOW

Albert Michelson (1852–1931)

Albert Michelson was born in Poland and emigrated to the United States when very young. He graduated in 1873 from the US Naval Academy, where he later taught chemistry and physics. While studying in Europe, Michelson became fascinated with the problem of measuring the speed of light, *c*.

Michelson was a first-rate measurement expert. In 1883, using rotating mirrors, he measured *c* more accurately than anyone had before.

Michelson's interferometer used light interference to resolve small path differences between two light sources. He used the interferometer to measure the Standard International Metre against the wavelength of cadmium light. With Morley, Michelson used it to find evidence for the luminiferous aether, without success. This is the most famous of all null results! Michelson and Francis Pease used an interferometer to measure the diameter of the star Betelgeuse, another first. Michelson was the first American scientist to win the Nobel Prize (1907).

Figure 6.9 shows a simplified diagram of the apparatus Michelson and Morley used.

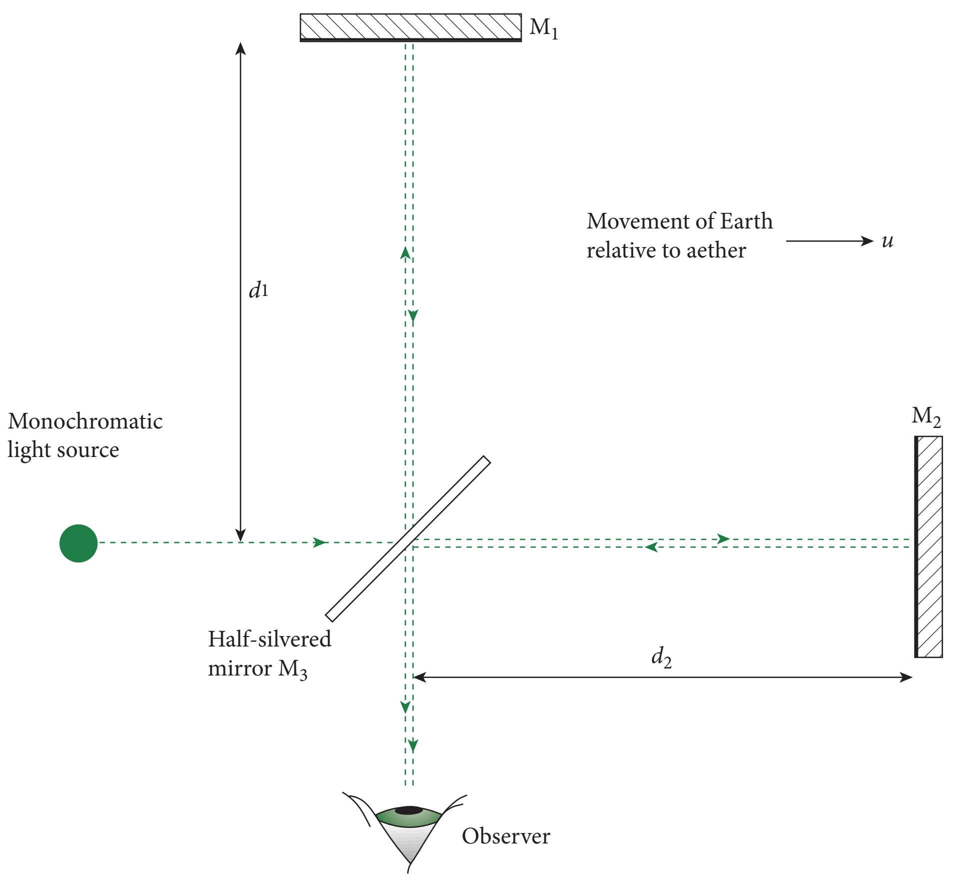
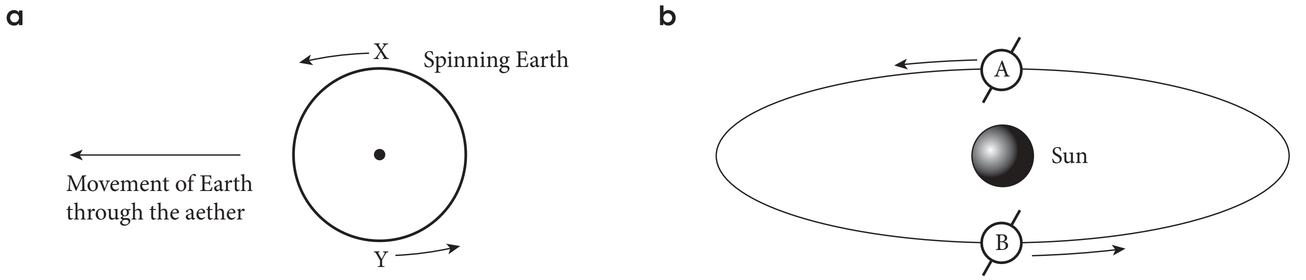


Figure 6.9 ▶ Schematic diagram of the Michelson–Morley apparatus. Lengths d_1 and d_2 were made equal to very high precision.

Part of a beam of light from the source would be transmitted through the half-silvered mirror M_3 and travel to, where it would be reflected; some of that beam would then reflect off M_3 towards the observer. The other part of the beam from the source would be reflected off M_3 and travel to M_1, where it would be reflected; some of that beam would then be transmitted through M_3 towards the observer. The two beams of light travel the same distance, arrive together, superimpose and produce an interference pattern.

If Earth were travelling through the aether at a velocity *u* towards the right, then a 90° rotation of the apparatus would create a path difference that could be observed in the interferometer. This path difference would be due to the assistance of the moving aether. The result of this path difference should be evident in a slight change to the location of the interference fringes.

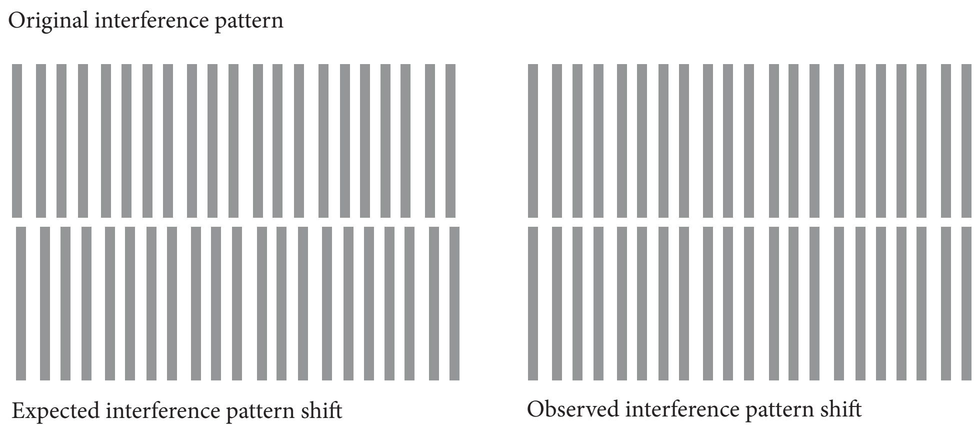


Michelson and Morley took a reading at position X and then another reading 12 hours later at position Y, to compare measurements at different speeds relative to the 'aether'.

By taking measurements 6 months apart, at positions A and then B, Michelson and Morley were again able to compare situations when Earth is moving at very different speeds relative to the 'aether'.

▲**Figure 6.10** Earth–Sun diagrams

Michelson and Morley took many readings – some twice in a 24-hour period (Figure 6.10(a)), some 6 months apart (Figure 6.10(b)) – in an attempt to measure the speed of light from a distant star. If the aether existed, then these positions (travelling towards the star and then away from the star 6 months later) should have produced interference shifts of up to 0.4 of a fringe. This is difficult to see but, in the hands of an excellent observer, is not impossible.



▲**Figure 6.11** Typical interference patterns. If the aether existed, there should be a shift in the pattern.

Much to their surprise their measurements failed to detect any evidence of the presence of the aether. The maximum shift observed was less than 0.01 of a fringe, well below the expected result and within the experimental error. Michelson and Morley were confident that their experiment showed that the aether did not exist. They had achieved a null result, which pointed to the non-existence of the aether. Michelson and Morley were not convinced. They made many further attempts to find the aether, always without success.

This null result was not what many scientists expected. Over the next 50 years, other scientists used a variety of techniques to repeat the experiment, but no evidence for the aether was found. Some scientists argued that perhaps Earth dragged the aether along with it and this would explain Michelson and Morley's results.

Astronomers later found that the revolution of Earth around the Sun caused aberration of observed starlight from distant stars. This indicated that the aether, if it existed, was not dragged along with Earth, ruling out this explanation of the Michelson–Morley experiment.

MICHELSON–MORLEY ANIMATION

Change the speed of light relative to the aether wind and find out how the time taken to traverse each path differs.

QUESTION SET 6.3

Remembering

1 What is the aether? Why did people go looking for the aether?

Understanding

2 What was the point of the Michelson–Morley experiment? What did they find?

3 Why did Michelson and Morley use light interference in their eponymous experiment?

Applying

4 In their experiment, how did Michelson and Morley use Earth's:

- a orbital motion?
- b rotation?

Explain in terms of aether drag.

Analysing

5 In the Michelson–Morley experiment, how would the shift of the interference pattern confirm the existence of the aether?

Reflecting

6 What conclusions can you draw from the Michelson–Morley experiment about the way science is conducted?

Einstein's theory of special relativity

Albert Einstein reflected long and hard on Maxwell's equations. He wondered what an electromagnetic wave would look like if you travelled along with it. He concluded that you would see no time change – the wave would be stuck in time – but that it would continue to oscillate in space. How could something change, yet no time pass? In order to resolve this problem, Einstein decided to investigate the nature of space and time. The result was the theory of special relativity. This was a bold step but a well-trodden path to discovery: questioning taken-for-granted assumptions leads to fresh, more powerful ways of understanding the world.

In 1905, Einstein published four papers that would change the way we view the universe. One of these papers introduced relativity. It took a totally fresh look at the nature of the laws of physics. It did away with the idea of the existence of the aether, and argued that there was no absolute frame of reference.

His theory was based on two clear propositions:

- *First postulate of special relativity* – The laws of physics are the same in all inertial frames of reference – the principle of special relativity.
- *Second postulate of special relativity* – The speed of light has the same value, *c*, in all inertial frames. It does not depend on the speed of either the source or the observer.

In all inertial frames of reference:

1 **the laws of physics are the same.**

2 **the speed of light is constant.**

Relativity did away with the aether. The aether was assumed to be the one, privileged reference frame, which violated the first postulate. Thus, the aether could not be uniquely detected. Either it did not exist, or its existence could not be demonstrated by experiments conducted completely within an observer's inertial reference frame. How could such an observer demonstrate the motion of their own reference frame?

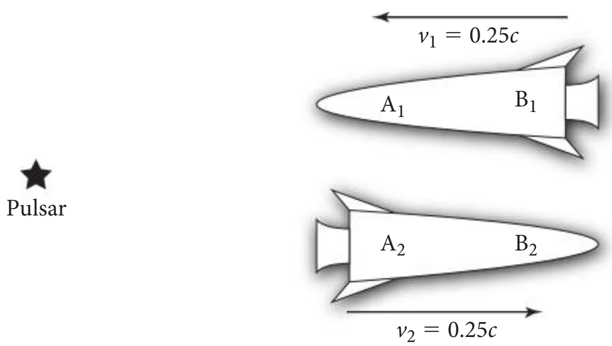
The second postulate completely contradicts Galilean and Newtonian relativity, and goes against apparent common sense. We would expect light to appear faster or slower depending on whether we are moving towards or away from the source. All experiments show that this is not the case. Indeed, the second postulate is consistent with the results of the Michelson–Morley experiment. The speed of light in a vacuum is $2.998 \times 10^8\,\text{m s}^{-1}$ regardless of the speed of either the observer or of the source.

Einstein was not sure whether he knew of the Michelson–Morley experiment when he came up with relativity. He was focused on Maxwell's equations when he gradually came to realise that the aether was not necessary for these equations to work. The Michelson–Morley results were, however, highly influential in the relatively rapid acceptance of the theory of special relativity.

Time dilation

The theory of special relativity asks us to give up our Newtonian view of space and time, and accept some very strange and puzzling ideas. To illustrate this, we will use a technique that Einstein used himself: simple thought experiments (*Gedanken* experiments, in German) that are based on the two postulates of special relativity.

Assume two rockets, each travelling at $0.25c$ but in opposite directions, pass each other, at the instant they receive a light pulse from a distant pulsar (see Figure 6.12). Observers on each rocket attempt to measure the speed of this pulse of light by using two light-sensitive cells set on the outside of the rocket (A_1 and B_1, and A_2 and B_2, respectively).



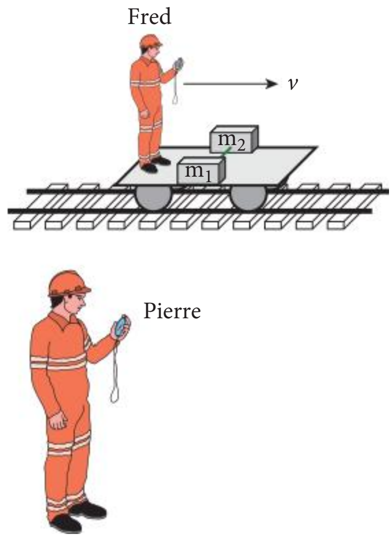
◀**Figure 6.12**
Passing rockets both view light from a pulsar.

Each rocket uses a timer and the distance between the sensors to measure the speed of light relative to its own reference frame. Rocket 1 measures the speed of light to be $3.00 \times 10^8\,\text{m s}^{-1}$. The second rocket also determines it to be $3.00 \times 10^8\,\text{m s}^{-1}$ relative to its reference frame.

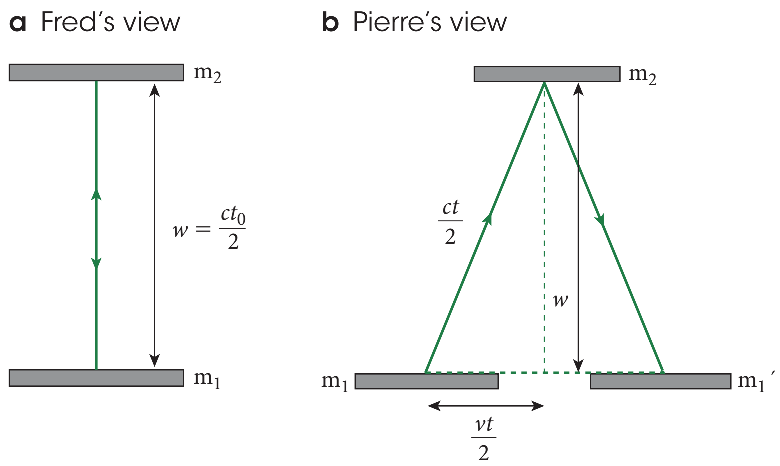
Newtonian relativity would have argued that the rocket heading towards the pulsar would have registered a light speed of $1.25c$ and the rocket moving away from the pulsar would have registered a light speed of $0.75c$. Of course, in practice it would be difficult to complete the experiment, given that we do not have rockets that can travel at $0.25c$. However, we can conceive of such a scenario and use logic and Einstein's postulates to relate what each observer would see or measure.

Let us imagine another experiment using a train, Einstein's favourite thought experiment apparatus. The train is running on a smooth track at a high velocity of v relative to the ground. On one carriage is a set of two mirrors (m_1 and m_2), set up to allow a series of light pulses to bounce forwards and back between them. Pierre, standing on the ground nearby, watches the train pass (see Figure 6.13). His watch is capable of measuring very small periods of time.

Fred, an observer on the train (also using a very accurate watch), measures the time (t_0) it takes for a pulse of light to travel from m_1 to m_2 and back again, from his viewpoint. Pierre and Fred agree that the distance between the mirrors is w.



▲**Figure 6.13**
A passing high-speed rail cart



▲Figure 6.14 An event as viewed from a) Fred's frame and b) Pierre's frame of reference. The time measured by Fred is the proper time.

Fred sees the situation as a simple path of light between two mirrors, and therefore the time it takes for the pulse to move from m_1 to m_2 and back again to m_1 (Figure 6.14(a)) will be:

$$t_0 = \frac{2w}{c} \text{ (relative to the train)}$$

The train is moving very fast and Pierre sees the situation quite differently. From his viewpoint, the velocity of the mirrors results in the light pulse forming a triangle (Figure 6.14(b)), because both mirrors are moving to the right with a velocity of v as the light pulse moves from one mirror to the other and back again. Pierre measures the time for one pulse to move from m_1 to m_2 to be t. One half of the journey forms a right-angled triangle with sides of length w, $\frac{vt}{2}$ and $\frac{ct}{2}$.

Proper time

Proper time is the time interval between two events occurring at the same place in an inertial frame, as measured by an observer in that inertial frame. Thus, from Fred's viewpoint, proper time is:

$$t_0 = \frac{2w}{c}$$

The time measured by Pierre is not the proper time, because Pierre is not travelling with the mirrors.

From Pythagoras' theorem we see that:

$$\left(c\frac{t}{2}\right)^2 = \left(v\frac{t}{2}\right)^2 + w^2$$

$$\Rightarrow \frac{t^2}{4}(c^2 - v^2) = w^2$$

$$\Rightarrow \frac{t^2}{4}\left(1 - \frac{v^2}{c^2}\right) = \frac{w^2}{c^2}$$

$$\Rightarrow \frac{t}{2}\sqrt{1 - \frac{v^2}{c^2}} = \frac{w}{c}$$

$$\Rightarrow t\sqrt{1 - \frac{v^2}{c^2}} = \frac{2w}{c} = t_0$$

$$\Rightarrow t = \frac{t_0}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Pierre tells Fred that the time for a single reflected pulse, as recorded by his watch (which is identical in every way to Fred's), is longer than that recorded by Fred's watch. Pierre's time measurement has been dilated. Einstein argued that the postulates of special relativity led to the understanding that observers in different inertial frames would not agree on time measurements.

Of course, **time dilation** will only be observed when inertial reference frames are moving relative to each other at speeds close to the speed of light. If v is small, such as the speeds we normally encounter every day, then $t = t_0$ because the ratio $\frac{v^2}{c^2}$ approaches zero.

The factor $\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ appears repeatedly in relativity. It is given the symbol γ. It is also common to denote the ratio $\frac{v}{c}$ as β.

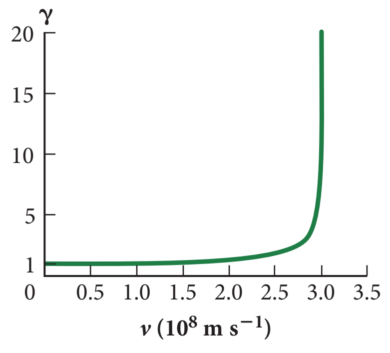
A clock in an inertial frame that is moving relative to a clock in a stationary frame will be regarded as running slow. The time interval for the moving clock is greater than the time interval for the stationary clock (time dilation):

$$t = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} t_0$$

$$\text{where } \gamma = \frac{1}{\sqrt{1-\beta^2}}$$

Table 6.2 Approximate values of γ for at various values of β

$\beta = \frac{v}{c}$	γ
0	1
0.0010	1.0000005
0.010	1.00005
0.10	1.005
0.20	1.021
0.50	1.155
0.80	1.667
0.90	2.294
0.94	2.931
0.99	7.089
0.999	22.37



▲Figure 6.15 Graph of γ vs c. As speed approaches c, γ increases rapidly.

Let us refer back to Pierre and Fred. What would happen if the mirrors were on the ground with Pierre and the experiment was repeated? Fred could argue that his train was stationary and that Pierre was moving at $-v$ with respect to him. This means he would find that Pierre's clock is slow compared to his. How do we reconcile these two observations?

Time dilation is about measurement in *different* inertial frames; it is a result of relative movement between the frames. The clocks do not physically change. Time dilation is about what an observer in one frame observes about an event in another, and it must be reciprocal because there is no absolute frame of reference.

WORKED EXAMPLE 6.3

A pilot in a rocket travelling with a velocity of 0.250*c* presses a button to flash a 'Hello' sign for 5.00 s at a space station as the rocket passes.

a For how long is the flash seen by an observer on the space station? (2 marks)

b Explain your reasoning (1 mark)

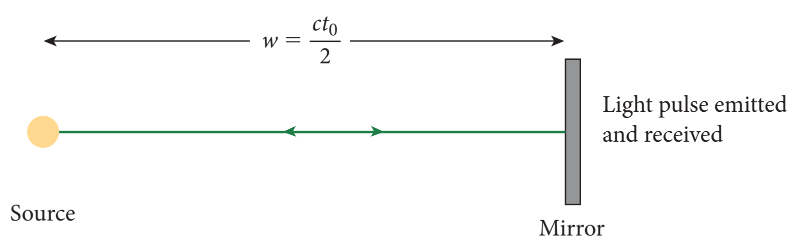
Answer	Logic	
a $t = \dfrac{t_0}{\sqrt{1 - \dfrac{v^2}{c^2}}}$		
$= \dfrac{5.00\,\text{s}}{\sqrt{1 - 0.25^2}}$	Substitute the correct values.	1 mark
$= \dfrac{5.00\,\text{s}}{\sqrt{0.9375}}$		
$= 5.16\,\text{s}$	Calculate the answer.	1 mark
b The observer in the space station views the rocket travelling towards it at a velocity of 0.250*c*. From this viewpoint, the clock on the rocket will be slow compared to the one in the space station. Any observer regards a clock that is moving relative to their frame of reference as running slow.	Recognise that time dilation is important at this relative speed.	1 mark

Length contraction

If time measurements are different relative to different inertial frames, what happens to measurements of length? Let us look at another thought experiment.

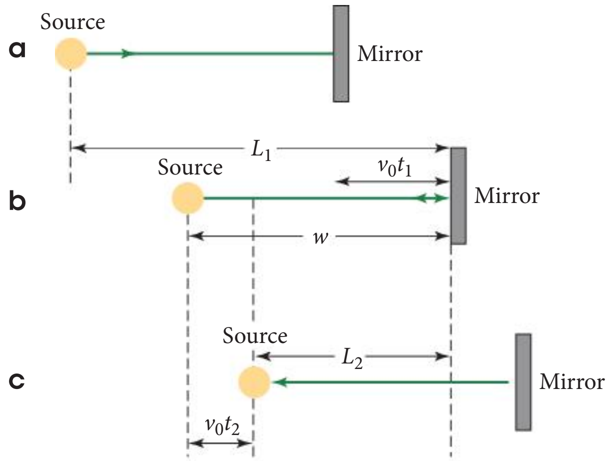
A moving train travels at a velocity of v_0. A source of light in a reference frame emits a short pulse of light that travels to a mirror and back (see Figure 6.16). An observer in the same frame of reference as the mirrors measures the time (t_0) for the pulse. If w_0 is the distance from the source to the mirror, the **proper length**, and c is the speed of light, then t_0 is equal to:

$$t_0 = \frac{2w}{c}$$



▲**Figure 6.16** An event as viewed within an inertial reference frame.

On the ground, an observer, who is stationary relative to the moving frame of the train, measures the time of transit of the light pulse to the mirror to be t_1, the initial distance between the source and mirror to be w and the distance between the original location of the source and the mirror to be L_1, as shown in Figure 6.17(a).



◀**Figure 6.17** The event as viewed in the stationary frame of reference

In this stationary reference frame, the observer notes that during the transit of the light pulse to the mirror the mirror has moved forwards a distant of v_0t_1, as shown in Figure 6.17(b). Then:

$$L_1 = w + v_0t_1$$

However, L_1 is also equal to ct_1, as the speed of light is the same in all reference frames. Therefore:

$$ct_1 = w + v_0t_1$$

$$\text{So, } t_1 = \frac{w}{c - v_0}$$

In Figure 6.17(c), the light pulse bounces off the mirror and travels back to the source in a time measured in the stationary reference frame to be t_2. So:

$$L_2 = w - v_0t_2$$

However, L_2 is also equal to ct_2, so we have:

$$ct_2 = w - v_0t_2$$

$$t_2 = \frac{w}{c + v_0}$$

From the perspective of the observer in the stationary reference frame, the total time for the journey is $t_1 + t_2$.

$$t_1 + t_2 = \frac{w}{c - v_0} + \frac{w}{c + v_0} = \frac{w(c + v_0) + w(c - v_0)}{c^2 - v_0^2} = \frac{2wc}{c^2 - v_0^2}$$

From the time dilation equation we know that:

$$t_1 + t_2 = \frac{t_0}{\sqrt{1 - \frac{v_0^2}{c^2}}}$$

Therefore, we have:

$$\frac{2wc}{c^2 - v_0^2} = \frac{t_0}{\sqrt{1 - \frac{v_0^2}{c^2}}}$$

$$\Rightarrow \frac{2wc}{c^2 - v_0^2} = \frac{2\frac{w_0}{c}}{\sqrt{1 - \frac{v_0^2}{c^2}}}$$

$$\Rightarrow \frac{wc^2}{c^2 - v_0^2} = \frac{w_0}{\sqrt{1 - \frac{v_0^2}{c^2}}}$$

$$\Rightarrow \frac{w}{1 - \frac{v_0^2}{c^2}} = \frac{w_0}{\sqrt{1 - \frac{v_0^2}{c^2}}}$$

$$\Rightarrow w = w_0\sqrt{1 - \frac{v_0^2}{c^2}}$$

This represents **length contraction**, because whatever the value of v_0, the factor $\frac{1}{\gamma}$ will always be less than one. This means that an observer in a stationary reference frame will regard the length of an object moving relative to them to be shorter than when it is at rest. The length of an object at rest is called its proper length. Like time dilation, length contraction will be reciprocal.

All observers will measure a moving object as being shorter or contracted in the direction of relative motion than when the object is at rest:

$$L = L_0\sqrt{1 - \frac{v^2}{c^2}}$$

$$\Rightarrow L = \frac{L_0}{\gamma}$$

where L_0 is the length of an object at rest, and L is the length of the object measured by an observer who is moving at relative velocity to the object's inertial frame.

It is important to note that the contraction is only in the direction of relative motion. There is no contraction in length at right angles to the motion. Just like time dilation, length contraction is only observable at relatively very high speeds.

Lorentz contractions and transformations

Two scientists, G. Fitzgerald from England and H. Lorentz from Holland, independently suggested an alternative explanation for the null result for the aether from Michelson and Morley's experiment that allowed Maxwell's equations to be applied to moving charges. Their explanation was that all bodies shrink in the direction of motion relative to the stationary aether by a modifying factor of:

$$\sqrt{1 - \frac{v^2}{c^2}}$$

Lorentz saw the consequence of the equation as a physical contraction of objects in space. However, Einstein argued that there was no physical change in the length of high-speed objects, but a change in the properties of time and space itself. The modern version of this factor, $\left(1 - \frac{v^2}{c^2}\right)^{\frac{1}{2}}$, is called the **Lorentz factor**.

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Time dilation: $t = \gamma t_0$

Length contraction: $L = \frac{L_0}{\gamma}$

WORKED EXAMPLE 6.4

1 An observer on the Moon notices a spacecraft travelling past at a speed of $2.08 \times 10^8\,\mathrm{m\,s^{-1}}$. The spacecraft has a proper length of 120 m.

What length will the observer on the Moon measure? (2 marks)

2 A crewed mission is to be sent to a newly discovered exoplanet 8 light-years away. It will travel at a velocity of $0.5c$ to get there.

- a According to the mission crew:
  - i how far away from Earth is the exoplanet? (2 marks)
  - ii how long will the journey take? (1 mark)
- b According to the mission command on Earth, how long will the journey take? (1 mark)

Answers	Logic	
1 $L = \frac{L_0}{\gamma}$		
$L = L_0\sqrt{1 - \frac{v^2}{c^2}}$		
$\Rightarrow L = 120\text{ m} \times \sqrt{1 - \frac{(2.08 \times 10^8\,\mathrm{m\,s^{-1}})^2}{(3.00 \times 10^8\,\mathrm{m\,s^{-1}})^2}}$	Substitute the correct values.	1 mark
$= 120\text{ m} \times \sqrt{1 - 0.481}$		
$= 120\text{ m} \times 0.72$		
$= 86.5\text{ m}$	Calculate the answer.	1 mark
2 a i $L = \frac{L_0}{\gamma}$		
$L = L_0\sqrt{1 - \frac{v^2}{c^2}}$		
$= 8\text{ light-years} \times \sqrt{1 - \frac{(0.5c)^2}{c^2}}$	Substitute the correct values.	1 mark
$= 8\text{ light-years} \times \sqrt{1 - 0.25}$		
$= 8 \times 0.866\text{ light-years}$		
$= 6.9\text{ light-years}$	Calculate the answer.	1 mark
ii $t_{crew} = \frac{s'}{v} = \frac{6.9\text{ light-years}}{0.5c} = 14\text{ years}$	Substitute the correct values and calculate the answer.	1 mark

b $t_{Earth} = \frac{s}{v} = \frac{8 \text{ light-years}}{0.5c} = 16$ years | Substitute the correct values and calculate the answer. | 1 mark

Try these yourself

A spaceship travels at an average velocity of 0.4*c* to an exoplanet 4.4 light-years away.

a What is the distance from Earth to the exoplanet in the frame of reference of the spacecraft? (2 marks)

b What is the difference between the time taken from Earth's perspective and the perspective of the spacecraft? (2 marks)

Several classes of subatomic particle have been identified as part of the Standard Model (Chapter 10), the current best model of matter and interactions. Muons are leptons, a class that includes electrons and neutrinos.

PARTICLE ADVENTURE

Fermions, quarks, hadrons, baryons, leptons, bosons and more.

The mean lifetime for particle decay comes from the exponential statistics of half-life. It is related to half-life by the equation $t_{mean} = 1.44t_{\frac{1}{2}}$.

DECAY AND GROWTH

Find out about a range of statistics that are useful for decay and growth.

Experimental evidence for relativistic effects: muon decay

Muons are a class of particles with a mass about 200 times greater than an electron. They are produced naturally by cosmic ray bombardment of the upper atmosphere. They travel at close to the speed of light and have a **mean lifetime** of 2.2×10^{-6} s (2.2 μs) as measured in the laboratory. The mean lifetime, t_{mean}, is a different measure from the half-life you met in Unit 1 Chapter 3. It is used here because it is a better measure of the probability that all the muons in a sample will have decayed in this time.

Clearly, muons do not exist for very long but, because they have relativistic speeds, they are one of the few available particles with which to demonstrate relativistic effects. This is what D.H. Frisch and J.H. Smith did more than 50 years ago. In an elegant experiment involving muons, they were able to demonstrate relativistic time dilation. Muons travelling down through the atmosphere were collected on a mountaintop at an altitude of 1907 m above sea level. They were travelling at 0.995c. This speed is measured from the muon relative to Earth or from Earth relative to the muon. Both agree about the relative speed.

At that speed, an observer travelling on the muon, that is, in the rest frame of the muon, would expect to reach a second detector at sea level in a proper time:

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

$$\Rightarrow t = \frac{1907}{0.995 \times 2.998 \times 10^8}$$

$$\Rightarrow t = 6.39\ \mu s$$

In terms of the mean lifetime, t_{mean}, this is $\frac{6.39\ \mu s}{2.2\ \mu s} = 2.90$ mean lifetimes. The number of muons reaching the ground would be negligible.

However, Frisch and Smith predicted that an observer on the ground (in Earth's rest frame) should observe a time dilation effect relative to an observer travelling with the muon (in the muon's rest frame). Muons should be observed in Earth's frame of reference to have a much longer mean lifetime than the mean lifetime measured in the muon's frame of reference. The number of lifetimes elapsed in the muon's frame of reference should be larger than the number of lifetimes elapsed as measured in Earth's frame of reference. The number of elapsed mean lifetimes measured in Earth's reference frame, $(t_{mean})_E$, should be related to the lifetime in the muon's reference frame, $(t_{mean})_\mu = 2.2\ \mu s$ according to the relativistic time dilation formula:

$$(t_{mean})_E = \gamma (t_{mean})_\mu$$

$$\Rightarrow (t_{mean})_E = \frac{2.2\ \mu s}{\sqrt{1 - (0.995)^2}}$$

$$(t_{mean})_E = 22.0\ \mu s$$

In this case, the time of travel to Earth would be $\frac{6.39\ \mu s}{22\ \mu s} = 0.29$ of the mean lifetime. There is plenty of time for muons to arrive at the collector on Earth.

These were the kind of predictions that Frisch and Smith made. When they did their experiment, they found that their relativistically corrected predictions were confirmed.

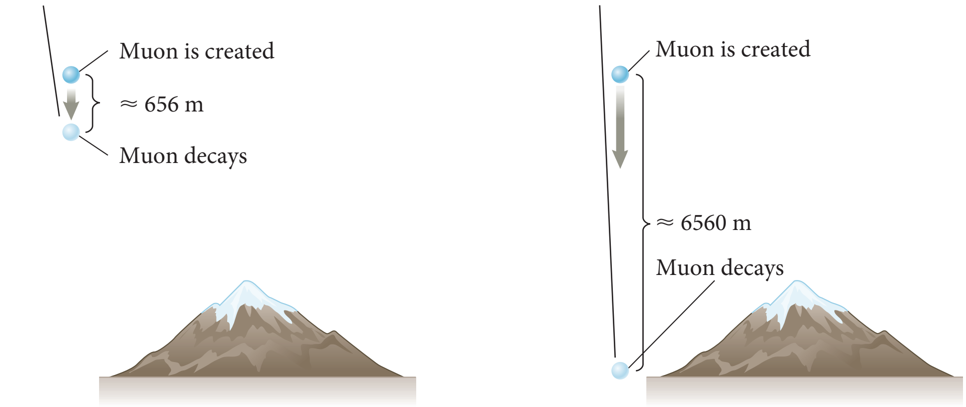
In the muon's frame of reference, almost all muons should have been extinguished in 2.2 μs. At a speed of 0.995*c*, this is a distance of 656 m. But in Earth's reference frame, the mean lifetime is 10 times longer. The muons should travel 6560 m. This is more than three times the height of Frisch and Smith's experiment. In Earth's frame, then, a significant number of muons should be, and were, detected. Relativistic effects were shown to have real consequences for measurement.

a

Without relativistic considerations, according to an observer on Earth, muons created in the atmosphere and travelling downwards with a speed close to *c* travel only about 656 m before decaying with an average lifetime of 2.2 μs. Therefore, very few muons would reach the surface of Earth.

b

With relativistic considerations, the muon's lifetime is dilated according to an observer on Earth. Hence, according to the observer, the muon can travel about 6560 m before decaying. The result is many of them arriving at the surface.



◀**Figure 6.18** Travel of muons according to an observer a) on the muon and b) on Earth

Thus, an observer on a moving muon (muon frame of reference) and an observer on Earth (Earth frame of reference) will agree about the:

- relative speed, 0.995*c*
- number of physical decays of muons
- number of elapsed mean lifetimes.

They will disagree about the:

- mean lifetime
- distance between the collectors.

MUON DETECTION

Watch the installation of a compact muon solenoid (CMS) in the large hadron collider (LHC) at the European Organization for Nuclear Research (CERN).

WOW

Muons

Frisch and Smith measured an average of 563 muons per hour in their upper detector. Assuming that other muons were passing the detector at the same rate, and assuming no muons are created between detectors, the number of muons arriving per hour at the second detector should be almost impossible to count. In fact, they measured 412 muons per hour – Far too many to be ignored and close to their predicted count rate.

WORKED EXAMPLE 6.5

Muons with a mean lifetime of 2.2×10^{-6} s and speed $0.999c$ were observed at a height of 10 000 m above Earth.

a From the viewpoint of a stationary observer on Earth:

i how long would it take for a muon to travel to the ground (assuming it does not decay)? (2 marks)

ii how many mean lifetimes will elapse before this muon arrives at the ground? (1 mark)

iii how likely is it that this muon will not have decayed before reaching the ground? (1 mark)

b From the viewpoint of an observer travelling with the muon (muon reference frame), how many mean lifetimes will elapse before the muon arrives at the ground? (4 marks)

Answers	Logic	
a i $t = \frac{S}{v} = \frac{10\,000\ \text{m}}{0.999 \times 3.0 \times 10^{8}\ \text{m s}^{-1}}$	Substitute the correct values with units.	1 mark
$\Rightarrow t = 3.3 \times 10^{-5}$ s	Calculate the answer.	1 mark
ii No. of mean lifetimes = $n = \frac{3.3 \times 10^{-5}\text{S}}{2.2 \times 10^{-6}\text{S}} = 15.2$	Calculate the answer.	1 mark
iii The probability is finite but extremely small.	Correct use of probability ideas.	1 mark
b $(t_{\text{mean}})_E = \gamma (t_{\text{mean}})_\mu$		
$\Rightarrow (t_{\text{mean}})_E = 22.37 \times 2.2\,\mu\text{s}$	Substitute the correct values and	1 mark
$\Rightarrow (t_{\text{mean}})_E = 49.2\,\mu\text{s}$	calculate the answer.	1 mark
$n_E = \frac{t}{(t_{\text{muon}})E} = \frac{3.3 \times 10^{-5}\text{s}}{49.2\,\mu\text{s}} = 0.67$	Calculate the final answer.	2 marks

Try these yourself

Muons are formed 3.0 km above Earth. They travel at $0.996c$ and have a mean lifetime, in the laboratory, of 2.2 μs. (2 marks)

a For an observer on a muon:

i how long does it take for the muon to travel to the ground?

ii how many mean lifetimes will elapse before the muon arrives at the ground? (1 mark)

iii explain why it is possible for the muon to arrive at the ground. (1 mark)

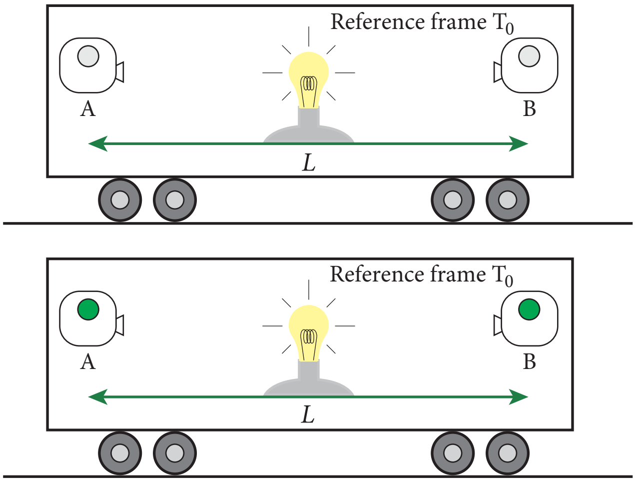
b For an observer on the ground, how many lifetimes will elapse before the muon arrives at the ground? (1 mark)

Simultaneity

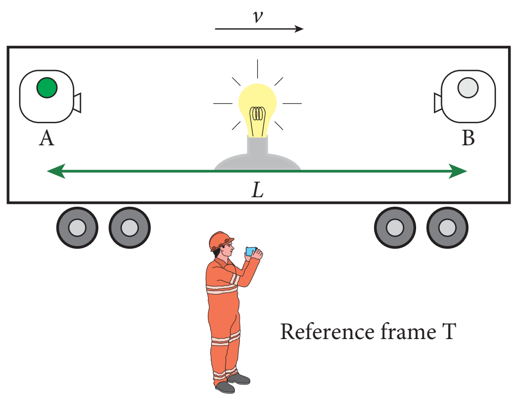
Measurement of time in different reference frames produces interesting results, not the least being that of simultaneous events. What may be seen as simultaneous events in one inertial reference frame may not be seen as simultaneous in another. This argues against the existence of a universal time frame.

What is meant by simultaneous events? Let us consider the example of a light that is positioned midway between each end of a train carriage. Two sensors A and B are placed at each of end of the carriage so that the distance to the light from each sensor is the same, $\frac{L}{2}$.

The globe is switched on and light radiates out so that, a short time later, it reaches each sensor, triggering a green indicator. In this reference frame, T_0, the indicator lights come on simultaneously, as the light took the same time to cover the same distance, $\frac{L}{2}$.



▲**Figure 6.19** Reference frame, T_0, inside the moving train



▲**Figure 6.20** View from the ground: reference frame, T

However, the train carriage is moving at very high constant velocity, v, relative to the ground. An observer in this reference frame, T, watches the train go past and is able to see the light globe and each sensor.

As the train carriage moves past, the observer notices that the globe and the trailing sensor, A, are both on, but the leading sensor, B, is off. In this reference frame, the observer sees the trailing sensor A rushing forwards to meet the light waves moving outwards from the globe, so A is triggered first. The leading sensor B is moving away from the light waves and takes longer to be turned triggered. The simultaneous events in reference frame T_0 are not simultaneous in reference frame T.

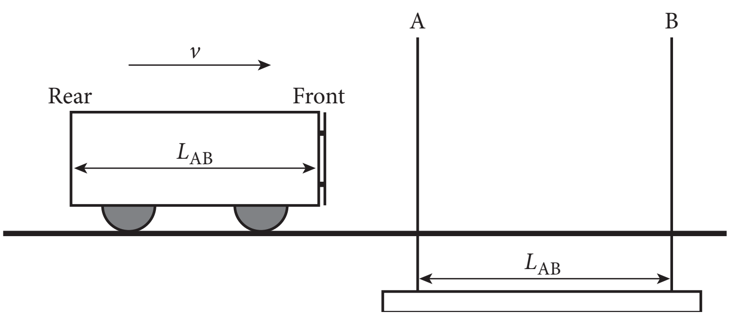
The events described above occur because **simultaneity** depends on agreement on time measurements, and time is relative. If Einstein's second postulate was not true, then simultaneity could be agreed on across different reference frames. This is seems to be against common sense or experience. Because the speed of light is so large compared to ordinary speeds, we normally do not see evidence of problems of timing in everyday life.

Two events are simultaneous when they occur at the same time.

Simultaneous events in one inertial frame may not be simultaneous events in another inertial frame of reference.

WORKED EXAMPLE 6.6

A rapidly moving rail cart (Figure 6.21) approaches a pair of markers at a speed of v.



◀**Figure 6.21** A rapidly moving rail cart from the reference frame of a ground-based observer

An observer on the side of the track measures the cart to be L_{AB} in length. As it speeds past the markers, this land-based observer notices that the front of the cart coincides with marker B and the rear of the cart coincides exactly with marker A. This means that this observer sees the two events as simultaneous.

Does an observer on the rail cart view these events as simultaneous? Give a qualitative answer. (3 marks)

Answer	Logic	
The cart-based observer sees the two markers as moving towards the cart at a velocity of v. This observer will measure L, the distance between A and B as being shorter due to length contraction: $L = \frac{L_{AB}}{\gamma}$	Recognise there will be length contraction.	1 mark
Hence, from their reference frame, the cart will be longer than the distance between the markers, and the front will coincide with B before the rear will line up with A.	State the correct conclusion.	1 mark
The cart-based observer will not see the two events as simultaneous.	State the correct answer.	1 mark

Try this yourself

The rail cart in Figure 6.21 has a proper length of 8.0 m and a speed of 0.9c. What is the difference between the proper length and the length AB measured by an observer on board the rail cart? (3 marks)

QUESTION SET 6.4

Remembering

1 Define these terms.

- a Proper time
- b Proper length

2 Write down Einstein's two postulates of relativity.

3 Define 'simultaneity'.

Understanding

4 How do Einstein's two relativity postulates relate to the conservation laws of physics?

5 Define these terms.

- a Time dilation
- b Length contraction

Applying

6 A spacecraft has a proper length of 50 m. It travels past an observer at 0.6c. What was its length in the observer's frame of reference?

7 A train carriage travels at 0.75c. To an observer outside the carriage, it fits exactly between two markers, P and Q. To an observer on the train, does the carriage appear to fit exactly between P and Q. If not, what can you conclude about events in the different frames of reference?

Analysing

8 A pion has a mean lifetime of 26 ns in Earth's inertial frame of reference. What is its mean lifetime in the pion's frame of reference if it has velocity $v = 0.75c$?

Reflecting

9 Describe your reaction to the discussion of the muon. How have you resolved any immediate concerns you had about the explanation?

10 Do you feel you understand relativity? What more do you need to understand to get a good grasp of the ideas in special relativity?

Einstein's famous equation: $E = mc^2$

Length and time have been shown to have values that depend on relativity. The third fundamental physical quantity is mass. Einstein argued that the mass of an object will be relative, and will be dependent on its relative speed:

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}$$

$$\Rightarrow m = \gamma m_0$$

The term m_0 is known as the **rest mass**, as measured when the mass is stationary in an inertial reference frame. m is the measurement of its mass in a reference frame moving in relation to a stationary frame, and is called the **relativistic mass** or **relativistically corrected mass**.

- **Rest or proper mass, m_0: mass as measured when the mass is stationary in an inertial reference frame. Proper mass never changes.**
- **Relativistic mass or relativistically corrected mass: the greater the relative velocity, the greater the relativistic mass: $m = \gamma m_0$**

This means that, as the velocity of an object increases and approaches the speed of light, the mass of the object will greatly increase. In fact, there is an ultimate velocity. As v approaches c in size then the term $\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$ approaches zero and m will be very large in size. If v is greater than c then the term would result in the square root of a negative number, an invalid result. This implies that the speed of an object with non-zero rest mass cannot be equal to or exceed the speed of light. It appears that the speed of light is the ultimate speed in the physical world.

WOW

Tachyons

According to some physicists and many science fiction writers, particles travelling faster than the speed of light – tachyons – may exist. Einstein's special theory of relativity does not completely rule out the existence of tachyons. Unlike photons and neutrinos, which are particle-like but with zero or very small rest mass, tachyons have imaginary mass. No tachyons have been discovered as yet.

WORKED EXAMPLE 6.7

1 What is the relativistically corrected mass of an electron whose speed is measured to be $1.8 \times 10^8\,m\,s^{-1}$? (1 mark)

Rest mass of an electron is 9.109×10^{-31}.

2 At what speed is a particle moving if its relativistic mass is five times larger than its rest mass? (4 marks)

Answers

Logic

1 Relativistic mass of electron:

$$m = \gamma m_0 = \frac{9.109 \times 10^{-31 kg}}{\sqrt{1 - \frac{(1.8 \times 10^8 \text{ m s}^{-1})^2}{(3.0 \times 10^8 \text{ m s}^{-1})^2}}}$$

$$= 1.1 \times 10^{-30} \text{ kg}$$

Substitute the correct values and calculate. 1 mark

2 $m = \gamma m_0$

$$\Rightarrow \gamma = \frac{m}{m_0}$$

Write the correct ratio. 1 mark

$$\Rightarrow \gamma = 5$$

$$\Rightarrow \sqrt{1 - \frac{v^2}{c^2}} = \frac{1}{5}$$

$$\Rightarrow 1 - \frac{v^2}{c^2} = 0.04$$

Find the square. 1 mark

$$\Rightarrow \frac{v^2}{c^2} = 0.96$$

Write the correct ratio. 1 mark

$$\Rightarrow \frac{v}{c} = 0.9798$$

$$\Rightarrow v = 0.9798 \times 2.998 \times 10^8 \text{ m s}^{-1}$$

$$\Rightarrow v = 2.94 \times 10^8 \text{ m s}^{-1}$$

Substitute the correct values and calculate. 1 mark

Try these yourself

1 What is the relativistically corrected mass of a proton whose speed is 0.75*c*? (2 marks)

2 A neutron has a relativistic mass that is 2.5 times greater than its rest mass. What is the speed of the neutron? (4 marks)

Momentum

Now that we have a definition for relativistic mass, what does this mean about momentum? Special relativity uses the classic equation for momentum:

$$p = mv$$

However, m is now the relativistic mass whose magnitude depends on the velocity of the mass. This means that the magnitude of an object's relativistically corrected momentum is given by:

$$p = \frac{m_0 v}{\sqrt{1 - \frac{v^2}{c^2}}}; \quad p = \frac{p_0}{\sqrt{1 - \frac{v^2}{c^2}}}; \quad p = \gamma p_0$$

Mass and energy

To explore the relationship between matter and energy, let us discuss a thought experiment involving the momentum of a photon that we know to be $\frac{E}{c}$. Imagine a stationary spacecraft in distant space.

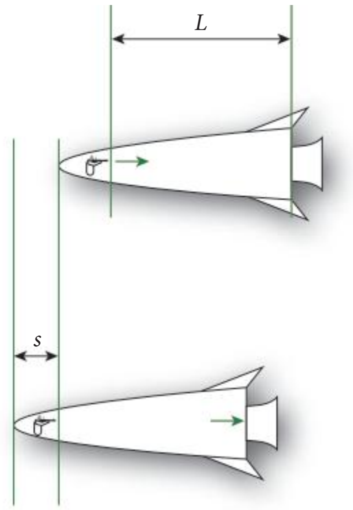
The pilot of the spacecraft is practising her aim with a laser gun that is capable of firing a single photon or a high-energy beam. She fires a single photon towards the rear of the spacecraft. The photon has momentum of $\frac{E}{c}$ (where E is its energy) and since momentum is conserved in this isolated system, the pilot and, in turn, the spacecraft (as she is seated in a chair at the front of the craft), will recoil forwards with an equal but opposite momentum. The distance the craft moves, s, will be equal to vt, the velocity of the craft multiplied by the time the photon takes to reach the rear of the spacecraft.

The magnitude of the momentum the spacecraft receives will be mv, and M is the mass of the spacecraft. Of course, the pilot will probably not notice this movement because of the small size of the momentum of a photon. However, as small as it is, the movement will occur. The time it takes for the photon to reach the rear of the craft, t, will be equal to $\frac{L}{c}$, where L is the length of the spacecraft cabin. Momentum is conserved, so the momentum of the photon $\left(\frac{E}{c}\right)$ will equal the size of the momentum of the spacecraft (Mv):

$$\frac{E}{c} = Mv$$

$$\Rightarrow v = \frac{E}{Mc}$$

Now, $s = vt$, so: $$s = \left(\frac{E}{Mc}\right)\left(\frac{L}{c}\right) = \frac{EL}{Mc^2}$$



▲**Figure 6.22** Photon fired in spacecraft. L = distance from front to rear of spacecraft, t = time for photon to reach rear of spacecraft, s = distance of recoil (with velocity v) = vt

As the photon reaches the rear of the spacecraft, its momentum will be transferred, causing the craft to stop. There is no net external force applied to the system; the total momentum of the system has not changed, yet the system has moved. It is the photon that has changed position and caused this re-distribution. In this interaction with the photon, the spacecraft has behaved exactly as would be expected to behave if there had been a re-distribution of mass (Einstein's 'mass equivalent' of energy hypothesis).

If we suppose the photon to have a (relativistically corrected) mass, m, then its momentum would be $p = mc$. Thus, the speed of the photon as it travels to the rear of the spacecraft is $c = \frac{L}{t}$. Therefore, again equating momenta:

$$\text{momentum of craft} = \text{momentum of photon}$$

$$\frac{Ms}{t} = \frac{mL}{t}$$

$$s = \frac{mL}{M}$$

We now have two equations for s, so equating them we have:

$$\frac{EL}{Mc^2} = \frac{mL}{M}$$

$$E = mc^2$$

Of course, this is only the discussion of a thought experiment and does not necessarily prove anything. However, the above equation has been derived in many other ways and, more importantly, has been successfully tested by experiment, not the least being the exploding of atomic bombs.

The energy associated with mass at rest is called the **rest energy** of the mass, and is given by the equation:

$$E = m_0c^2$$

The relativistic total energy of an object that is moving is defined by:

$$\text{total energy} = \text{rest energy} + \text{relativistic kinetic energy}$$

Einstein's famous equation is usually stated as $E = mc^2$. It really means the energy difference between the energy equivalence of zero rest mass and a non-zero rest mass. That is, the equation refers to a change of energy from zero to a rest mass energy: $E = mc^2$. This is why Einstein's mass–energy equation is often written as $E = \Delta mc^2$, which is the energy change from rest mass energy to the energy of the relativistically corrected mass.

One way of interpreting the above statements is to consider that the *change* in the **relativistic kinetic energy** of a body is related to a change in its mass:

$$\Delta E_k = (\Delta m)c^2$$
$$\Rightarrow \Delta E_k = (m - m_0)c^2$$
$$\Rightarrow \Delta E_k = mc^2 - m_0c^2$$
$$\Rightarrow \Delta E_k = \gamma m_0c^2 - m_0c^2$$
$$\Rightarrow \Delta E_k = (\gamma - 1)m_0c^2$$

Mass and energy are equivalent. Mass is a manifestation of energy.

Rest mass–energy: $E = m_0c^2$

- $\Delta E = \Delta mc^2$
- $\Delta E = (\gamma - 1)m_0c^2$

WORKED EXAMPLE 6.8

An electron is accelerated from rest by a potential difference of 500 kV.

a What speed does the electron attain? (8 marks)

b What speed would the electron attain if non-relativistic mechanics were used? (2 marks)

Answer	Logic	
a $\Delta E = qV$ and $\Delta E_k = (\gamma - 1)m_0c^2$	Use the correct equality.	1 mark
$\Rightarrow (\gamma - 1)m_0c^2 = qV$	Use the correct transformation.	1 mark
$\Rightarrow (\gamma - 1) = \frac{qV}{m_0c^2}$		
$\Rightarrow \gamma = \frac{qV}{m_0c^2} + 1$		
$\Rightarrow \gamma = \frac{(1.602 \times 10^{-19}\text{ C})(500 \times 10^3\text{ V})}{(9.109 \times 10^{-31}\text{ kg})(2.998 \times 10^8\text{ m s}^{-1})^2} + 1$	Substitute the correct values with units.	1 mark
$\Rightarrow \gamma = 1.9784$	Calculate the value.	1 mark
$\Rightarrow \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} = 1.9784$	Substitute the correct values and calculate. Calculate the value.	1 mark 1 mark
$\Rightarrow \sqrt{1 - \frac{v^2}{c^2}} = \frac{1}{1.9784} = 0.5055$	Find the inverse.	1 mark
$\Rightarrow 1 - \frac{v^2}{c^2} = (0.5055)^2 = 0.2555$		
$\Rightarrow \frac{v^2}{c^2} = 1 - 0.2555 = 0.7445$		
$\Rightarrow \frac{v}{c} = 0.8628$		
$\Rightarrow v = 0.8628 \times 2.998 \times 10^8\text{ m s}^{-1}$		
$\Rightarrow v = 2.587 \times 10^8\text{ m s}^{-1}$	Calculate the answer.	1 mark

b Non-relativistically:

$\Delta E_k = qV$ and $\Delta E_k = \frac{1}{2}mv^2$ — Use the correct equality and transformation. 1 mark

$\Rightarrow \frac{1}{2}mv^2 = qV$

$\Rightarrow v = \sqrt{\frac{2qV}{m}}$

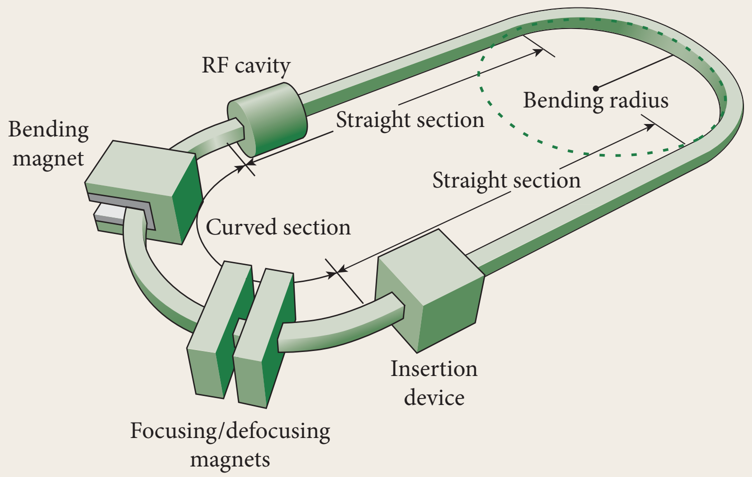
$\Rightarrow v = \sqrt{\frac{2 \times (1.602 \times 10^{-19})(500 \times 10^3)}{9.109 \times 10^{-31}}}$

$\Rightarrow v = 4.194 \times 10^8\,\mathrm{m\,s^{-1}}$ — Calculate the correct answer. 1 mark

Scientific literacy: Energy and momentum at the Australian Synchrotron

The storage ring at the Australian Synchrotron has a radius of about 35 m but it should fit on a kitchen table, from the viewpoint of the particles inside it. The reason has to do with Einstein's relativity. Electrons travel in the storage ring at very nearly the speed of light. For the 3 GeV Australian synchrotron, electrons travel at 99.999995% of the speed of light. This means that their masses must be relativistically corrected. Relativistic corrections become necessary when the speed of the electron is about 10% of the speed of light.

Electrons do not follow a perfect circular path. They travel along straight sections and are then subjected to curved sections.



▲**Figure 6.23** A synchrotron storage ring comprises straight and curved sections.

When a 3 GeV electron is affected by bending magnets that produce a 1.3 T field, it follows a circular path. This path has a radius wider than the kitchen table. In the straight sections it is affected by insertion devices called wigglers and undulators. These produce radiation in the nanometre range. This is a result of the combined effect of the relativistically corrected Doppler shift and length contraction in the electron's frame of reference.

Synchrotron light is emitted when electrons are accelerated by the bending magnets. The electrons are moving at speeds close to the speed of light and the light arrives at the speed of light at an observer. The wavelength, λ, of the radiation in the observer's frame of reference is corrected relative to the frequency observed in the frame of reference of the source of radiation, λ_0, via the relativistic Doppler equation:

$$\lambda = \lambda_0\sqrt{\frac{1-\beta}{1+\beta}}$$

Typically, wigglers produce wavelengths of about 10 cm in the frame of reference of the electron; however, this must be length corrected and Doppler shifted. This produces wavelengths in the nanometre range as expected.

SYNCHROTRON

Read about the links between Einstein's special relativity and the synchrotron.

Questions

1 For the Australian Synchrotron what is:

  a the maximum energy available?

  b its overall radius?

c the speed of the electrons?

d the magnitude of the magnetic field produced by the bending magnets?

e the value of the Lorentz factor, γ, when the speed of the electron is travelling at:

i 10% of c?

ii maximum speed?

(Hint: You will need to use a value for c that has sufficient significant figures.)

2 For the electrons in the storage ring at the Australian Synchrotron, if $\beta = \frac{v}{c}$, what is the value of:

a β?

b v?

3 Copy and complete the table below to compare kinetic energy and momentum for an electron travelling at 0.999 999 95c according to both classical and relativistic physics.

	Kinetic energy (J)	Momentum (N s)
Classical		
Relativistic		

On a single set of axes, plot E_k vs β for kinetic energies up to 1000 keV for both classical and relativistic energies. Use a spreadsheet to produce values of v from values of β; hence find γ and E_k. Compare and contrast these two graph lines.

4 a What is the wavelength of electromagnetic radiation in the:

i wiggler's reference frame?

ii electron's reference frame (Lorentz contraction)?

iii electron's reference frame (Doppler corrected)?

b Is your result consistent with the claim in the article? Give a reason for your answer.

5 Calculate the radius of the electron path in the field of the bending magnets. How does this compare with the claim in the article about the radius?

6 a Summarise the article in 70–75 words.

b Present the summary as a dialogue that starts with a question from an interested Year 8 student. Limit the dialogue to messages that are less than 140 characters in length.

RELATIVITY IN 10 MINUTES!

Use this Weblink to summarise special relativity.

Mass defect in nuclear physics

For a full discussion see Nelson Physics Units 1 & 2 for the Australian Curriculum, Chapter 4.

It has been observed that in any nuclear event – radioactive decay, nuclear reaction, fusion and fission – there is a loss of mass. This **mass defect** appears as energy in amounts that are predicted by Einstein's mass–energy equation.

Nucleons may have their masses reported in different ways: in kilogram (kg), in unified mass units (u) and energy (MeV).

Table 6.3 Mass of nucleons in kilogram, unified mass units and MeV

	Mass (kg $\times 10^{-27}$)	Mass (u)	Energy equivalent (MeV)
	1.6605	1.0000	931.49
Proton	1.6726	1.0073	938.30
Neutron	1.6749	1.0086	939.51

WORKED EXAMPLE 6.9

Calculate the energy of the γ-ray in the following nuclear reaction. (Assume all the energy is carried by the γ-ray):

$^{113}_{48}Cd + {}^{1}_{0}n \rightarrow {}^{114}_{48}Cd + \gamma$ (6 marks)

Answer	**Logic**	
$^{113}_{48}Cd + {}^{1}_{0}n \rightarrow {}^{114}_{48}Cd + \gamma$		
$112.9044\ u + 1.0086\ u \rightarrow 113.9034\ u + \gamma$		
$\Rightarrow (\Delta m) = (112.9044 + 1.0086)\ u - 113.9034\ u$		
$\Rightarrow (\Delta m) = 9.6000 \times 10^{-3}\ u$	Calculate Δm.	(2 marks)
$\Rightarrow$ energy of $\gamma = 9.6000 \times 10^{-3}\ u \times 1.661 \times 10^{-27}\ kg\ u^{-1} \times (2.998 \times 10^{8}\ m\ s^{-1})^2$		
$\Rightarrow$ energy of $\gamma = 1.433 \times 10^{-12}$ J	Calculate γ-ray energy (J).	(2 marks)
but 1 eV = 1.602×10^{-19} J		
$\Rightarrow$ energy of $\gamma = \dfrac{1.447 \times 10^{-12}\ J}{1.602 \times 10^{-19}\ J\ ev^{-1}}$		
$\Rightarrow$ energy of $\gamma = 8.95 \times 10^{6}$ eV		
$\Rightarrow$ energy of $\gamma = 8.95$ MeV	Calculate γ-ray energy (MeV).	(2 marks)

Try this yourself

Find the energy carried by the alpha-particle for the following radioactive decay: (6 marks)

$^{238}_{92}U \rightarrow {}^{234}_{90}Th + {}^{4}_{2}He$

(Assume all the energy is carried by the alpha-particle.)

Nuclear fission and fusion

Almost all of the energy released in a nuclear fission reaction is carried by the fast neutrons that are released. The average energy for a thermal neutron-induced fission reaction is about 200 MeV. This means that the mass defect is about 200 MeV, which amounts to 0.215 u or 0.35×10^{-27} kg. These are very small amounts of energy, but the number of nuclides in a reactor is enormous. Power reactors may have up to 190 tonnes of fuel. If this were all uranium-235 nuclides, and if all of the nuclides underwent fission, the amount of energy available would be enormous (approx. 10^{16} J). Nuclear fusion, which is still in the experimental stage, is predicted to be even more potent.

QUESTION SET 6.5

Remembering

1 Write down relativistic equations for:
   a mass.
   b momentum.
   c rest energy.
   d energy transfer.

2 Define 'mass defect'. How is it related to energy?

Understanding

3 What is the importance of mass defect to nuclear physics?

Applying

4 How much energy does a proton have when it is travelling at 0.80*c*?

Analysing

5 An electron gains 600 keV of energy in an electron gun. What speed does it attain?

6 Find the mass defect for the alpha-decay (4.003 u) of Po-214 (213.995 u) to Pb-210 (209.984 u). Is this consistent with the alpha-particle having an energy of 7.68 MeV? Explain your answer.

7 Confirm that 190 kg of pure uranium-235 could produce ~10^{16} J of energy? (order of magnitude argument only!).

Reflecting

8 What previous knowledge did you apply to the discussion of relativity and its importance for nuclear physics?

9 How can you use *Gedanken* experiments in your learning?

CHAPTER SUMMARY

- The speed of light in a vacuum has a value of 299 792 458 $m\,s^{-1}$ or an approximate value of $3.0 \times 10^8\,m\,s^{-1}$.
- The speed of light in a medium depends only on the magnetic and electric properties of that medium.
- An inertial frame of reference is one that is not accelerating; that is, it is stationary or moving at constant velocity.
- The principle of relativity states that the laws of motion are the same in all inertial frames of reference.
- The proper length of an object is the length measured by an observer at rest relative to the object.
- Relative velocities – the velocity of A relative to B is the velocity of A relative to C plus (vector addition) the velocity of C relative to B, or $\vec{v}_{AB} = \vec{v}_{AC} + \vec{v}_{CB}$.
- Michelson and Morley found no evidence for the existence of the aether.
- *First postulate of special relativity*: The laws of physics are the same in all inertial frames of reference.
- *Second postulate of special relativity*: The speed of light has the same value in all inertial frames.
- The time interval between two events occurring at one place in an inertial frame, as measured by an observer in that inertial frame, is called proper time, t_0 (sometimes called local time).
- Time dilation: The time interval measured between two events occurring at the different locations in an inertial frame will be greater than proper time.

$$t = \frac{t_0}{\sqrt{1 - \frac{v^2}{c^2}}}$$

- Length contraction: All observers will measure a moving object as being shorter or contracted in the direction of relative motion than when the object is at rest.

$$L = L_0\sqrt{1 - \frac{v^2}{c^2}}$$

where L_0 is the length of an object at rest, and L is the length of the object measured by an observer who is moving at relative velocity to the object's inertial frame.

- The Lorentz factor (γ) is defined as:

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

- Simultaneous events in one inertial frame may not be simultaneous events in another inertial frame of reference. This is known as the principle of non-simultaneity.
- Rest mass or proper mass is mass as measured in an intertial reference frame in which the object is stationary.
- Relativistic mass increases as the velocity of the object increases in its reference frame:

$$m = \gamma m_0 c^2$$

- The rest energy of a mass is the energy associated with a mass at rest and is given by:

$$E_{rest} = m_0 c^2$$

where m_0 is the rest mass and c is the speed of light.

- The relativistic total energy of an object that is moving is defined by:

total energy = rest energy + relativistic kinetic energy

- The change in the kinetic energy of a body is related to a change in its mass:

$$\Delta E_k = (m - m_0)c^2 = (\gamma - 1)m_0 c^2$$

CHAPTER GLOSSARY

aether a medium through which light waves were thought to propagate

electrical permittivity a physical property of a medium associated with electricity

Galilean transformation equations relating coordinates in one inertial frame to those in another inertial frame; classical relativistic transformations

inertial frame of reference a non-accelerating frame of reference

invariant same in all frames of reference

length contraction length appears shorter in a reference frame that is moving relative to a stationary frame

Lorentz factor $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$

magnetic permeability physical property of a medium associated with magnetism

mass defect difference between energy before and after a nuclear decay or reaction

mean lifetime statistical measure of the time after which a decaying sample is effectively extinguished; mean lifetime = 1.44 × half-life

muon exotic particle formed by cosmic rays in the upper atmosphere

proper length length measured in an inertial frame of reference in which the object is stationary

proper time time measured between two events occurring at the same location in an inertial stationary frame of reference

relativistic kinetic energy $\Delta E_k = (\Delta m)c^2 \Rightarrow \Delta E_k = (\gamma - 1)m_0c^2$

relativistic mass, relativistically corrected mass mass measured in a reference frame that is moving relative to the reference frame in which the rest mass was measured

relativity principle the laws of physics are the same in all inertial frames of reference

rest energy $E = m_0c^2$

rest mass mass measured in an inertial reference frame in which the object is at rest

simultaneity idea that the same event will be seen to occur at the same time in different reference frames

stellar aberration position of star due to relative motion of Earth

time dilation a longer time is measured

CHAPTER REVIEW QUESTIONS

Remembering

1 Write down Einstein's two postulates for relativistic motion.

2 Define these terms.
  - a Inertial frame of reference
  - b Relative motion
  - c Proper time
  - d Proper motion
  - e Time dilation
  - f Length contraction
  - g Mass defect

3 List all the relativistic equations.

Understanding

4 What is the importance of Maxwell's equation for:
  - a electromagnetism?
  - b relativity?

5 A passenger walks quickly (at 3.0 m s^{-1}) towards the rear of a train carriage that is travelling forwards at 16 m s^{-1}. Relative to an observer on the ground nearby:
  - a what is the speed of the passenger?
  - b where is the passenger after 5.0 s?

6 Describe the main features of the Michelson–Morley experiment. Why is the null result in this experiment significant?

7 What is simultaneity? How is simultaneity understood in relativistic terms? Give an example.

Applying

8 A space probe travels into deep space at an average velocity of 0.6*c*. When it is 1.8×10^9 m away as measured from Earth, It sends back to Earth a pulse of light of duration 2.00 s.

a What is the speed of the light pulse:

i according to the space probe?

ii according to Earth-based observers?

b According to Earth-based observers:

i how long does the pulse take to reach Earth?

ii what is the duration of the pulse?

9 Electrons in an electron gun are about 0.5% heavier than electrons at rest. What is the velocity of the electrons in the electron gun?

10 Two identical clocks are synchronised. One clock is sent off in a spaceship travelling with a speed $v = 0.70c$. After 49 years on the Earth clock, what is the time on the spaceship, as observed from Earth?

11 A pion has a mean lifetime of 26 ns in Earth's frame of reference. What is its mean lifetime in its own frame of reference if it is travelling at 0.5*c*?

12 Muons, are created by the collision of cosmic rays with air molecules high in Earth's atmosphere (10000 m). They approach the ground with a velocity of 0.999*c*. From the viewpoint of an observer travelling at the velocity of the muons, how high do they measure Earth's atmosphere to be?

13 The main nuclear fusion reaction that creates the energy in a hydrogen bomb is:

$$4\,{}^{1}_{1}\mathrm{H} \rightarrow {}^{4}_{2}\mathrm{He} + 2\,{}^{0}_{+1}\mathrm{e} + \text{energy}$$

The mass of H is 1.673×10^{-27} kg, the mass of He is 6.644×10^{-27} kg and the mass of a positron is 9.109×10^{-31} kg.

How much energy, in MeV and joule, is produced from:

a four hydrogen atoms?

b 10 kg of hydrogen?

Analysing

14 The Twin Paradox illustrates the effects of time dilation. Fred and Pierre are twins. Fred takes an extraterrestrial journey that takes him to a distant stellar system and then returns. The average speed of Fred's spacecraft is 0.6*c*. Pierre remains on Earth. When Fred returns, Pierre is 20 years older. How much has Fred aged according to his space watch?

15 Xavier, Ignatius and Maxine are triplets. Xavier pilots a spacecraft away from Earth towards the centre of the Milky Way for a distance of 7 light-years relative to Earth. Maxine pilots a similar spacecraft in the opposite direction, away from the centre of the Milky Way. Her craft also only travels outwards for a distance of 7 light-years. Their average speed is 0.61*c*. Ignatius remains at home. Both return at the same time.

a Will Xavier and Maxine appear to be the same age on their return?

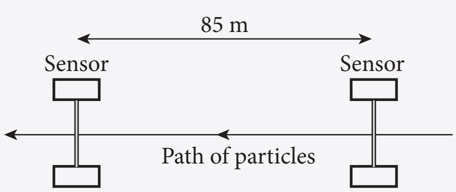
b How much older than the travellers will Ignatius be when the travellers return?

16 The pilot of a non-accelerating spacecraft, moving away from Earth at great speed, celebrates the passing of six birthdays. Earth-bound observers measure this elapsed time to be 10 years. Relative to Earth:

a What is the speed of the spacecraft?

b How far does the spacecraft travel over these six birthdays?

17 A short pulse of particles from a linear accelerator passes through two sensors that are spaced 85 m apart (Figure 6.24). The particles are timed to travel the distance in 0.377 μs.



◀Figure 6.24

a What is the velocity of the particles, as a proportion of the speed of light?

b In the reference frame of the particles, what is the spacing of the detectors?

18 A positron is the antimatter equivalent of an electron. Both have a rest mass of 9.109×10^{-31} kg. When an electron and a positron collide, they annihilate each other and produce energy in the form of electromagnetic radiation. How much energy is released by the collision?

19 Muons have a speed of $0.995c$ when they are recorded at 5000 m above Earth's surface. They have a mean lifetime of 2.2 μs in their rest frame.

In the muon reference frame:

- a i How far do muons travel in a mean lifetime?
  - ii Is it possible for muons to reach Earth? Explain your answer.
- b How high above Earth's surface are they created?
- c Is it likely that muons will reach Earth?

20 The ultimate speed is c, but there does not appear to be an ultimate kinetic energy, $E = mc^2$. Explain this apparent contradiction.

21 In a synchrotron, 2.5 GeV electrons travelling at $0.9996c$ are subjected to a 2.8 T magnetic field.

- a What is the momentum of the electrons?
- b What is the radius of their path?

Reflecting

22 Create a visual representation of your understanding of special relativity.

CHAPTER 7
QUANTUM THEORY AND LIGHT

By the end of this chapter you will have covered the following material.

Science Understanding

- Atomic phenomena and the interaction of light with matter indicate that states of matter and energy are quantised into discrete values (ACSPH135)
- On the atomic level, electromagnetic radiation is emitted or absorbed in discrete packets called photons; the energy of a photon is proportional to its frequency; and the constant of proportionality, Planck's constant, can be determined experimentally (for example, from the photoelectric effect or the threshold voltage of coloured LEDs) (ACSPH136)
- A wide range of phenomena, including black body radiation and the photoelectric effect, are explained using the concept of light quanta (ACSPH137)
- Atoms of an element emit and absorb specific wavelengths of light that are unique to that element; this is the basis of spectral analysis (ACSPH138)

Introduction

You have met several scientific models already in your study of physics. A model is generally considered successful when it has both good explanatory and predictive power. In Chapters 8 and 9 of *Nelson Physics Units 1 & 2 for the Australian Curriculum* you studied a model of forces and motion called Newtonian mechanics. Newtonian mechanics provides a very powerful model that helps us to understand and allows us to predict how objects will behave when forces are exerted on them. A second model is the classical wave model, which you have studied in Chapter 10 of *Nelson Physics Units 1 & 2 for the Australian Curriculum*. You have seen that this model, coupled with the electromagnetic field model in Chapter 5, can explain many of the behaviours of light. Light is an electromagnetic wave. Light behaves just like other waves – it reflects, refracts, shows diffraction and interference, just as do sound and other mechanical waves.

Each of these models replaced earlier, less successful models. Newtonian mechanics replaced Aristotelian mechanics, and the electromagnetic wave model replaced Newton's particle model of light.

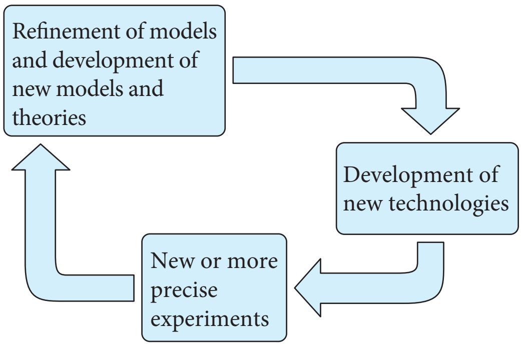


Figure 7.1 ▲ The interactions between experiment and theory – each leads to advances in the other.

For hundreds of years it seemed as if these models could explain the behaviour of all physical systems. Then experiments started to show results that could *not* be explained by these models. You may wonder why it took so long for people to be able to do these experiments. Often in science new developments occur and new theories have to be constructed because improvements in technology allow for more precise measurements or different sorts of measurements. There is a strong interaction between science and technology. New understanding in science leads to the development of new technologies. The development and improvement of technology allows new experiments to be done. If the results of these experiments disagree with the existing models and theories, then new models and theories need to be developed. These, in turn, lead to new understandings and new technologies.

Experiments from the 19th and early 20th century showed results that did not agree with the classical models. Gustav Kirchhoff, whose laws for analysing circuits you learnt about in Chapter 5 of *Nelson Physics Units 1 & 2 for the Australian Curriculum*, introduced the term **black body radiation** in 1859. This refers to the infrared, visible or higher frequency light emitted by an object due to its thermal energy. The spectrum emitted by a black body was not able to be explained in terms of the classical models of waves. The classical wave model was also unable to explain the **photoelectric effect** or atomic spectral lines. These mismatches between theory and experiment suggested that physicists needed to rethink the very nature of light itself. This was the beginning of quantum mechanics.

In Chapter 2 of Nelson Physics Units 1 & 2 for the Australian Curriculum you studied heat transfer and saw that any object with a temperature above absolute zero radiates energy. This energy is quantised and we shall see how it is produced in this chapter.

The development of the quantum theory was a major revolution in science. It changed the way that we understand the behaviour of matter and the nature of the universe. It caused divisions among scientists, and vigorous argument and debate. There is still debate among physicists as to the interpretation of aspects of the quantum theory. However, for all the argument, it has been a spectacularly successful theory. Quantum theory has had a bigger impact on our daily lives than any other theory in science. The quantum theory led to the development of semiconductors and semiconductor devices, including diodes and transistors. All of our modern information and communication technology is based on these developments, and you use them constantly. Every time you use a computer or a smart phone or watch TV you are using an application of quantum mechanics. It has been estimated that 90% or more of the wealth of the world is directly related to quantum mechanics!

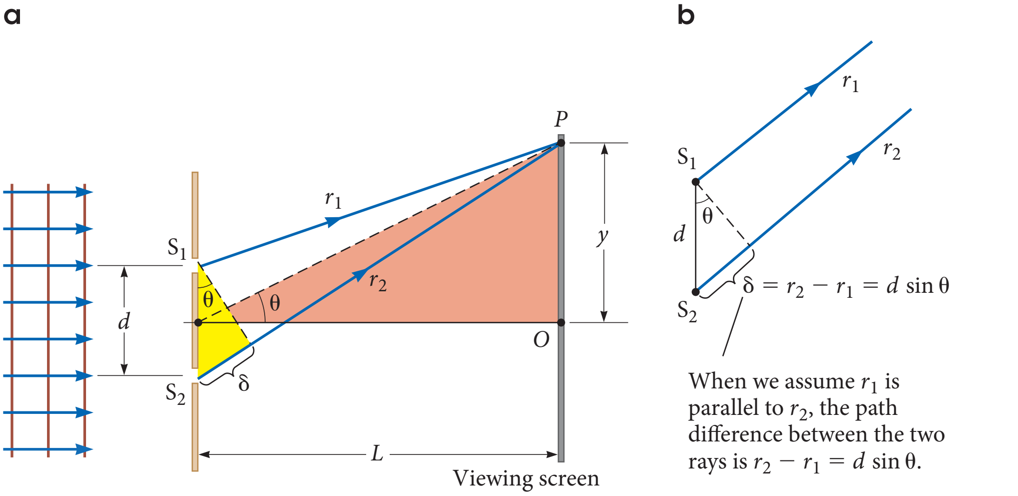
The nature of light

In Chapter 10 of Nelson Physics Units 1 & 2 for the Australian Curriculum you studied the wave model of light, and a simplified wave model, called the ray model.

In Newton's time (the 17th century) there were two competing models for light – the 'undulatory' or wave model and the 'corpuscular' or particle model. Newton himself was a proponent of the particle model. At around the same time Christian Huygens was working on his wave model, as described in Chapter 11 of *Nelson Physics Units 1 & 2 for the Australian Curriculum*. In the early 19th century experiments such as Young's double-slit experiment (sometimes called the twin-slit experiment) provided convincing evidence that light acts like a wave.

The wave nature of light

In the double-slit experiment, light is shone through a pair of narrow, closely spaced slits. The resulting interference pattern is observed on a distant screen. The interference pattern results from the path difference between the light waves coming from the two different slits. The path difference results in areas of constructive and destructive interference – giving a pattern of bright and dark fringes. The path difference is calculated as shown in Figure 7.2.



◀ **Figure 7.2** The double-slit experiment. Bright fringes occur when $d\sin\theta = m\lambda$, which is when $y = \frac{Lm\lambda}{d}$ for $L >> d$.

When the screen is a long way from the slits, then the two rays, r_1 and r_2, are approximately parallel. Figure 7.2(b) shows a close up of the approximately parallel rays at the slits. The interference pattern produced at the screen is due to the two rays travelling different distances to reach a given point on the screen. The difference in distance travelled is called the **path difference**. The path difference is shown as δ in Figure 7.2, and is given by:

$$\delta = r_2 - r_1.$$

We define an angle, θ, as the angle between the normal to the line joining the two slits, and the point of interest, P, on the distant screen. We assume that the screen is also perpendicular to this normal line. This allows us to write the path difference in terms of the angle θ:

$$\delta = r_2 - r_1 = d\sin\theta$$

where d is the distance between the slits.

Recall from your previous studies of waves that constructive interference occurs when two waves have the same phase, in other words when peaks line up with peaks and troughs line up with troughs. At these points there is an antinode or bright spot in the pattern. This occurs whenever the path difference is equal to a whole number of wavelengths, or when:

$$d\sin\theta = m\lambda, \quad m = 0, 1, 2, \ldots$$

At the positions where peaks meet troughs – the waves are always half a cycle out of phase – a node appears. This occurs when:

$$d\sin\theta = \left(m + \frac{1}{2}\right)\lambda, \quad m = 0, 1, 2, \ldots$$

The angle, θ, can be related to the height, y, above the point where the normal line reaches the screen (point O in Figure 7.2(a)) and the distance L to the screen. For small angles,

$$\sin\theta \approx \tan\theta = \frac{y}{L}$$

Therefore, the path difference can also be written as:

$$\delta = \frac{dy}{L}$$

Relating this to the conditions for constructive and destructive interference tells us where the bright and dark spots in the pattern will occur.

The points of constructive interference (bright spots) in a double-slit interference pattern occur at

$$y = \frac{Lm\lambda}{d}$$

The points of destructive interference (dark spots) occur at

$$y = \frac{L(m + \frac{1}{2})\lambda}{d}$$

where $m = 0, 1, 2, \ldots$.

Diffraction and refraction were discussed in Chapter 11 of Nelson Physics Units 1 & 2 for the Australian Curriculum.

The double-slit experiment is not the only experiment that is explained by the wave model of light. The wave model also correctly predicts the diffraction of light by small apertures and around small objects. The interference pattern produced as a result of diffraction through an aperture can be modelled as the interference of waves from different parts of the aperture. The interference pattern due to diffraction around a small object is explained by the interference of waves coming from each side of the object.

All of these interference effects are explained by the wave model of light. In this model, light propagates as a wave and, hence, is spread out over a region of space, just as a water wave is spread out over a surface. It is this spreading out or *delocalisation* that allows diffraction and interference to occur. When more than one light wave exists in a region of space, superposition means the waves can add to give points of constructive and destructive interference.

Mechanical waves such as sound need a medium through which to propagate. They cannot propagate through vacuum. Light waves do not need a medium, so they can propagate through vacuum.

In *Nelson Physics Units 1 & 2 for the Australian Curriculum* you also studied refraction of waves. Light waves refract as do other waves, and the electromagnetic wave model of light (Chapter 5) explains why this occurs. Unlike mechanical waves, light does not need a medium through which to propagate. The oscillating electric and magnetic fields that make up the light wave are coupled. As the electric field varies, it creates a varying magnetic field, which in turn creates a varying electric field, and so on, and thus the wave can propagate through empty space. When the light wave meets a medium other than vacuum, the electric and magnetic fields interact with the atoms and electrons in the material. This slows the light wave down, and causes its path to bend. This is refraction, as you saw in *Nelson Physics Units 1 & 2 for the Australian Curriculum* Chapter 11.

The electromagnetic wave model of light also explains polarisation. Polarisation is a result of the electric field component of light only being allowed to oscillate in a particular direction. As described in Chapter 11 of *Nelson Physics Units 1 and 2 for the Australian Curriculum*, this can be achieved by absorbing and rotating components of the electric field using a polaroid sheet, or by reflecting light from a smooth surface.

Maxwell's electromagnetic wave model of light successfully predicts and explains all these behaviours of light. It also correctly predicted the speed of light, as we saw in Chapter 5.

Hence, the wave model of light became the mainstream, accepted model. It was a very successful model for more than a hundred years, and is still very useful.

However, this electromagnetic wave model did not always correctly predict the outcome of experiments. Two experiments in particular showed that a new model was needed. These were the black body radiation and the photoelectric effect experiments.

EXPERIMENT 7.1

THE WAVE NATURE OF LIGHT

When a wave with wavelength λ is passed through a pair of closely spaced narrow slits it forms an interference pattern. A pattern of maxima and minima resulting from constructive and destructive interference can be observed. If the pattern is observed at some distance, *L*, from the slits, where *L* is much greater than the slit separation *d*, the distance between any two adjacent maxima is given by:

$$\Delta y = \frac{\lambda L}{d}$$

Aim

To observe the behaviour of light incident on various arrangements of narrow slits

Materials

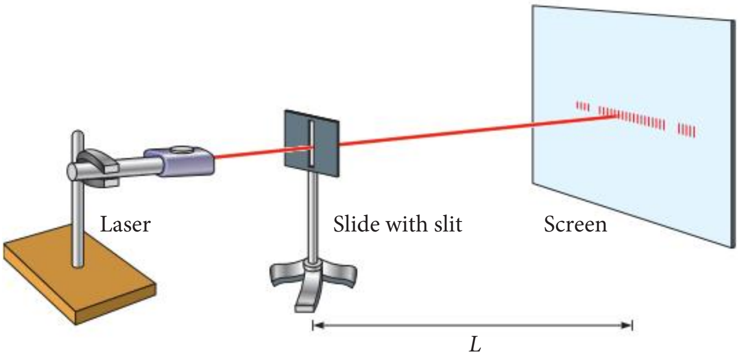
- laser or laser pointer
- slides with a single slit, two closely spaced slits with known slit separation, and a diffraction grating
- a screen

What are the risks in doing this experiment?	How can you manage these risks to stay safe?
Lasers can cause serious eye damage.	Make sure the laser is always pointing away from people. Never look into the laser.

In your write-up, add any more risks you can think of, as well as ways to manage them.

Procedure

1 Arrange the laser, single slit slide and screen as shown in Figure 7.3.
2 Measure the distance from the slide to the screen.
3 Turn on the laser and observe the pattern formed on the screen.
4 Replace the slide with the double-slit slide and observe the pattern formed.
5 Measure the distance from the central maximum to the farthest bright spot that you can clearly see. Count how many spots there are between the central spot and the one you measure to.
6 Replace the slide with the diffraction grating and observe the pattern formed.



◀ **Figure 7.3** Experimental set-up

Results

1 Sketch each of the patterns you observed. Note the relative brightness of the various spots.
2 For the double-slit slide, record the number of spots to each side of the central bright spot. Record the distance from the central spot to the furthest one. Don't forget to include the **uncertainty** in your measurement.

Analysis of results

1 Refer back you what you learned about waves in Unit 1. How can the patterns you observed be explained?
2 For the double-slit slide, use your measurements and the known slit separation to calculate the wavelength of the laser light.

Discussion

1 In what way is the interference pattern from the double-slit slide different from the interference pattern from the grating?
2 Why do we need to use the wave model of light to explain the results of this experiment?
3 Why does a single slit produce an interference pattern (called a diffraction pattern)?

QUESTION SET 7.1

Remembering

1 Name two models of light you have studied.
2 Name three phenomena that the wave model of light can explain.

Understanding

3 Why is it only possible for an interference pattern to form in the double-slit experiment if light travels as a wave and hence is delocalised? Draw a diagram to help explain your answer.
4 In the twin slit experiment, state what happens to the spacing of the light and dark fringes (increases, decreases or stays the same) if:
   a a shorter wavelength of light is used.
   b the screen is moved closer to the slits.
   c more closely spaced slits are used.

Applying

5 In a twin slit experiment light of wavelength 630 nm is incident on a pair of slits spaced a distance 15 μm apart. If the screen is a distance 2.0 m from the slits, at what positions do the first three bright spots appear?
6 In a measurement to find the wavelength of a light source, a viewing screen is placed a distance 4.8 m from a pair of slits with separation 0.030 mm. The first dark fringe is a distance of 4.5 cm from the centre line on the screen.
   a Find the wavelength of the light.
   b Find the distance between any two adjacent bright spots.

Analysing

7 When unpolarised light is passed through a polariser, and then through an analyser with its polarising axis at a right angle to the axis of the polariser, the intensity of the remaining light is approximately zero. When a second polarising sheet is placed between the polariser and the analyser, the final intensity is much brighter and clearly non-zero. Explain how this can happen.
8 A pair of slits spaced 0.015 mm apart is illuminated with light of two wavelengths at the same time: λ_1 = 630 nm and λ_2 = 420 nm. The viewing screen is a distance 3.0 m from the slits. At what position on the screen, other than at $y = 0$, do the maxima from the two interference patterns first line up? Give the m values for the bright spot for each wavelength at this position.

Reflecting

9 Draw a spider diagram summarising the wave model of light. Include on your diagram all the behaviours explained by the electromagnetic wave model.

Black body radiation

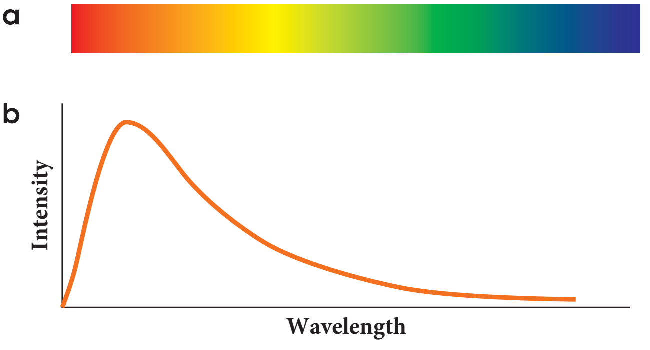
All objects continuously radiate energy in the form of electromagnetic waves. At any non-zero temperature, a body emits radiation of all wavelengths, but the distribution or **spectrum** of wavelengths depends on its temperature. If an object is very hot you can see the light that is being emitted; for example, you can see the glowing coals in a fire or the filament of a light globe. At low temperatures, the wavelengths of the emitted radiation are mainly in the low frequency, infrared region and cannot be seen, although you may still be able to feel the radiation as heat with your skin. Measurements show that the hotter an object is, the more electromagnetic radiation is emitted, and the more of that radiation is at shorter wavelengths. Hence, if you take a piece of metal and heat it up slowly, it will first glow a dim red, then bright yellow and eventually very bright white.

Measurement of the intensity of emitted radiation as a function of wavelength shows that it is a **continuous** distribution of wavelengths from the infrared, through the visible and ultraviolet. This distribution is called a **continuous spectrum**. The shape of the spectrum depends only on the temperature of the object, and not on any of its other properties. A continuous spectrum is shown in Figure 7.5.

Shutterstock.com/Pablo Hidalgo

▲ Figure 7.4
Lava glows red-yellow.

In Nelson Physics 1 & 2 for the Australian Curriculum, Chapter 1, you looked at models of energy transfer, including radiation. In Chapter 5 we saw that this radiation, mostly infrared light, is a form of electromagnetic waves.



▲ Figure 7.5
a) A continuous spectrum. b) Intensity is a function of wavelength for a continuous black body spectrum.

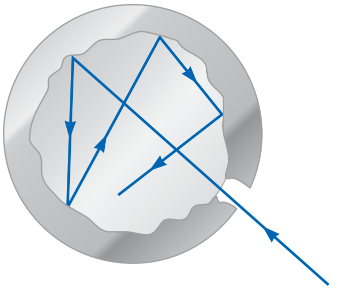
What is a black body?

A **black body** is an ideal surface that completely absorbs all wavelengths of electromagnetic radiation incident on it. Hence it is a *black* body. Such a surface will also be a perfect emitter of electromagnetic radiation at all wavelengths. The radiation is characteristic of the temperature of the black body and, when the body is at room temperature, is mainly in the infrared part of the electromagnetic spectrum when the body is at room temperature.

Although a true black body is only a theoretical concept, it can be closely simulated in a laboratory. Consider a cavity (hollow space) that has the interior walls blackened and which is kept at a constant temperature (Figure 7.6). If a small hole is made in the wall of the cavity, it will act like a black body radiator. Any radiation that falls on the hole from the outside will pass through it. After multiple reflections, the radiation will be absorbed by the interior surfaces. As the cavity is in thermal equilibrium with its surrounds, the interior surfaces will emit radiation at the same rate at which it is absorbed. The radiation that escapes depends only on the temperature of the cavity. It is not affected by the size of the cavity or the material of which it is made.

Remember that this is an idealised object, not a real one. In practice, materials that absorb most of the light incident on them are good approximations of a black body. This is where the term 'black body' comes from – black objects absorb most of the light incident on them, regardless of wavelength.

Recall from Nelson Physics 1 & 2 for the Australian Curriculum that if two substances in physical contact with each other exchange no net heat energy they are in thermal equilibrium. The two substances must be at the same temperature.



▲ Figure 7.6
The opening to a cavity is a good approximation of an ideal black body. Note that it is the *opening* to the cavity that is the black body, not the entire hollow object. The hole acts as a perfect absorber.

The black body spectrum

The black body model is useful because it allows us to determine the temperature of distant objects. For example, we can estimate the surface temperature of the Sun by measuring its electromagnetic spectrum.

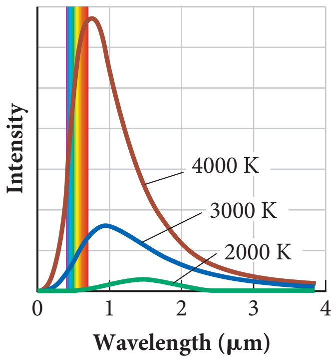


Figure 7.7 ▲ Intensity and distribution of wavelengths of radiation from a black body at different temperatures

Figure 7.7 shows the spectra of emitted radiation for a black body at various temperatures. Note how the peak in the radiation curve gets higher and shifts to shorter wavelengths as the temperature increases.

In 1893 Wilhelm Wien derived a relationship between the position of the peak wavelength at which radiation is emitted and the temperature of a black body. He used the idealised black body cavity model to derive the relationship, now known as Wien's displacement law or simply as Wien's law.

The position of the peak wavelength is given by Wien's law:

$$\lambda_{max} = \frac{b}{T}$$

where λ_{max} = peak wavelength in m, T is the absolute temperature in K, and b is Wein's constant, $b = 2.898 \times 10^{-3}\,\text{m K}$.

ACTIVITY 7.1

BLACK BODY RADIATION FROM A LIGHT GLOBE FILAMENT

Aim

To observe black body radiation from a light globe filament as the temperature of the filament changes

You will need

A 12 V incandescent light globe and a continuously variable 12 V DC power supply

What to do

1. Make sure the power supply is switched off and that the voltage is turned to zero. Connect the light globe across the terminals of the power supply. Turn on the power supply and *very slowly* turn up the voltage from zero until the globe just starts to glow.
2. What colour does it glow at this very low power?
3. Increase the voltage *very slowly*, observing what happens to the globe. Observe how the colour changes. What colour does the filament appear when the voltage is turned up to 12 V?
4. Record your observations. What is happening to the temperature of the filament as the voltage is increased?

WORKED EXAMPLE 7.1

The surface of the Sun has a temperature of approximately 5800 K. If we treat the Sun as a black body, what is the peak wavelength of the radiation emitted? Use Figure 5.41 in Chapter 5 to determine the part of the electromagnetic spectrum to which this wavelength belongs. (4 marks)

Answer	Logic	
$\lambda_{max} = \frac{b}{T}$	Relate wavelength to temperature.	1 mark
$\lambda_{max} = \frac{2.898 \times 10^{-3}\,\text{m K}}{5800\,\text{K}}$	Substitute numbers with correct units.	1 mark
$\lambda_{max} = 5.0 \times 10^{-7}\,\text{m}$	Calculate the final value. Note that we have rounded the answer to two significant figures, as the temperature was given to two significant figures.	1 mark
This is yellow (visible) light.	Locate the wavelength on Figure 5.41.	1 mark

Try this yourself

The black body spectrum shown is for the star Antares. What is the surface temperature of the star Antares? (4 marks)

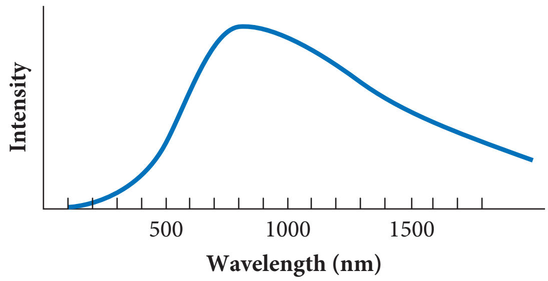


Figure 7.8 Black body spectrum for Antares

Wien's law was a successful model in that it accurately predicts the position of the peak wavelength. However, there were still two problems. First, there was no theory that explained the shape of the curve. Second, Wien's law was based on an idealised system – a cavity with a small hole. It is difficult to see how this theoretical model could represent the surface of a solid piece of material or a star such as the Sun.

Classically, it was thought that the thermal radiation originated from oscillating charged particles near the surface of an object. In the previous chapter we saw that oscillating charges are a source of electromagnetic waves. This is how antennae work. The oscillating charges in the antenna produce an electromagnetic wave of the same frequency as the oscillations. Recall from Chapter 1 of *Nelson Physics Units 1 & 2 for the Australian Curriculum* that the temperature of a material is a measure of the average kinetic energy of the atoms of that material. In a gas or a liquid the particles are free to move. The higher the temperature, the more kinetic energy the particles have and the faster they move. In a solid material, the atoms are not free to move, so this kinetic energy is observed as vibrations, and the higher the temperature, the higher is the frequency of vibration. And as you know, atoms are made up of smaller particles including protons and electrons, which are charged. Hence, this theory provided the oscillating charges needed to produce the electromagnetic radiation.

Now consider again the ideal model of the black body cavity. If the atoms on the inside surface are acting as little antennae, we would see standing waves set up between the walls of the cavity. Recall that you learnt about standing waves in Chapter 10 of *Nelson Physics Units 1 & 2 for the Australian Curriculum*. The waves produced by the vibrating atoms in the inside surface would reflect from the opposite surface. If the waves have the right wavelength, a standing wave is set up, just like standing waves in a string. We call these standing waves **modes of vibration**.

Classically, all the possible modes of vibration would be equally probable, and the total energy would be divided equally between them all. In any cavity, more short wavelength modes would be able to fit in the cavity. This means more short wavelength radiation should be emitted through the hole. As the temperature of the cavity increased, so should the total energy. As the energy increased, the energy associated with the short wavelengths (ultraviolet, X-rays and gamma rays) would approach infinity. According to this theory, even a regular heater should be emitting dangerous amounts of X-rays and gamma rays!

Figure 7.9 shows a comparison of a theoretical spectrum based on this model and a measured spectrum. This mismatch between theory and experiment was called the 'ultraviolet catastrophe'. Of course it was only a catastrophe for the theory that predicted it.

A new theory was needed to solve these problems.

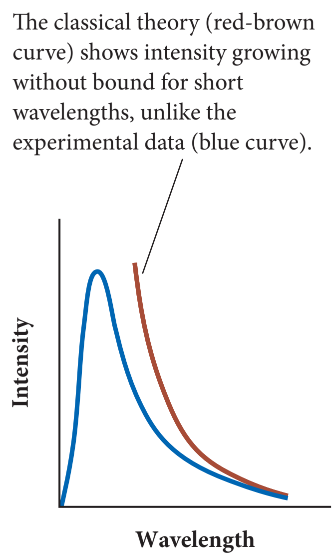


Figure 7.9 Comparison of the classically predicted and experimental black body spectra

Planck's quanta of energy

In 1900 a German physicist named Max Planck used 'lucky guesswork' (as he called it) to derive a formula that correctly matched the experimentally observed spectrum. Planck proposed that the atoms could only oscillate with **discrete** energies, given by

$$E_n = nhf$$

Watch your units! KE and h must have consistent units. If KE is in J, then h must be in Js. If KE is in eV, then use h in eVs. Always write your numbers with units so you don't get mixed up!

where n is an integer, f is the frequency of oscillation and h is a constant. The constant h, now known as the **Planck constant**, is $h = 6.626 \times 10^{-34}$ J s $= 4.14 \times 10^{-15}$ eV s.

This was a radical proposition. It means that the energy of the oscillators is **quantised**. It may only take discrete values, given by the equation above, rather than any possible value in a continuous range.

From this Planck deduced that the oscillators could only emit and absorb electromagnetic radiation (light) in packets of energy of specific sizes. He called these packets of energy 'quanta'. The amount of energy emitted is equal to the amount of energy lost by an oscillator when it goes to a lower energy state. For example, if an oscillator goes from an energy of $E_3 = 3hf$ to $E_2 = 2hf$, the energy lost is $E_3 - E_2 = 3hf - 2hf = hf$.

One quantum of light has energy $E = hf$.

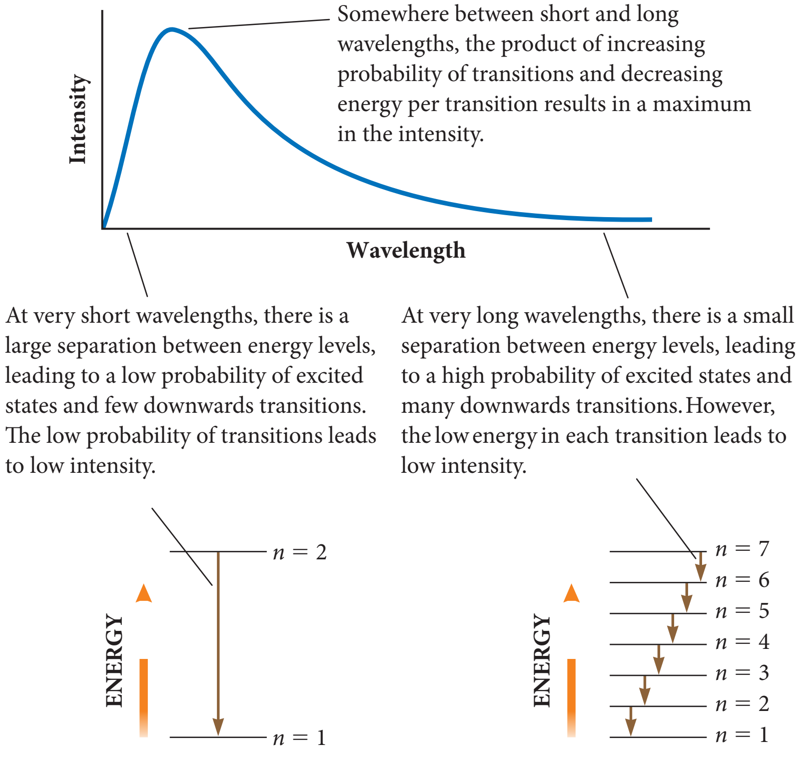


Figure 7.10 ▲ Planck's model for black body radiation

Planck combined this idea of quantisation with two ideas from classical statistical mechanics. First, the probability of an oscillator having a particular energy decreases as the energy increases. Hence, the probability of an atom being in a higher energy state (called an excited state) is lower. This means that the intensity of radiation at high frequencies (short wavelengths) is small. Second, the probability of a *change* in energy decreases with the relative gap between energy levels. The relative gap is larger for lower energies, or long wavelengths, so intensity is again low at these wavelengths. In between, we see the peak observed in the experimental spectra.

The quantisation of energy was such a revolutionary departure from classical physics that even Planck was reluctant to accept his own idea. Although Planck had discovered a mathematical way of explaining the shape of the black body spectrum, he was concerned that there was no physical model for how the energy could be in these discrete packets. It was Einstein who put physical meaning to Planck's **quantum** hypothesis.

WORKED EXAMPLE 7.2

A quantum of energy has wavelength 5.8×10^{-5} m.

a What is the frequency of this quantum? (4 marks)

b What is the energy of this quantum? (3 marks)

Answers	Logic	
a $c = f\lambda$	Use the dispersion relation for light to relate frequency to wavelength.	1 mark
$\Rightarrow f = \dfrac{c}{\lambda}$	Rearrange for frequency.	1 mark
$f = \dfrac{3.00 \times 10^{8}\ \text{m s}^{-1}}{5.8 \times 10^{-5}\ \text{m}}$	Substitute numbers with units.	1 mark
$f = 5.2 \times 10^{12}$ Hz	Calculate final value.	1 mark

b $E = hf$ Relate energy to frequency. 1 mark

$E = (6.626 \times 10^{-34}\,\text{J s})(5.2 \times 10^{12}\,\text{Hz})$ Substitute numbers with correct units. 1 mark

$E = 3.3 \times 10^{-21}\,\text{J}$ Calculate final answer. 1 mark

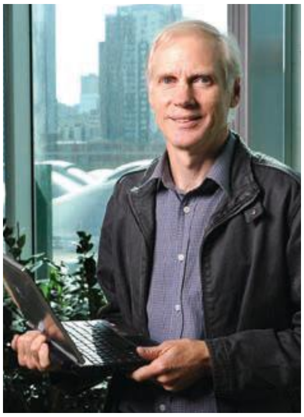
Try these yourself

1 What is the frequency of a quantum of energy with $E = 2.5\,\text{keV}$? (4 marks)

2 What is the wavelength of a quantum of energy with $E = 1.7 \times 10^{-25}\,\text{J}$? (5 marks)

Case study

Dr Robert Colman – Black body radiation and climate change modelling



▲ **Figure 7.11**
Dr Robert Colman

Dr Robert Colman is head of the Climate Change Processes Team in the Centre for Australian Weather and Climate Research at the Bureau of Meteorology in Melbourne. Dr Colman did a PhD in physics at the University of Melbourne, after studying maths and physics as an undergraduate. He is the author of more than 100 scientific articles and was a lead author on the United Nations Intergovernmental Panel on Climate Change (IPCC) 4th climate change assessment. The IPCC assesses scientific, technical and socio-economic information on climate change. The IPCC is a massive international collaboration that reviews the work of thousands of scientists. It then provides rigorous scientific information to decision makers, including governments such as our own.

The Climate Change Processes Team conducts research into the causes of past climate changes and into the changes we can expect in temperature, rainfall and other climate features in the future. The mathematical models used are based on fundamental physical principles, such as the laws of conservation of energy, mass and momentum, as well as a wealth of experimental observations.

Black body radiation is one of the important physical processes that are taken into account. The combined Earth–atmosphere system acts like a black body. It absorbs radiation from the Sun, and also radiates energy out into space.

The amount of energy absorbed depends on the **albedo** (reflectivity) of Earth's surface and atmosphere. The albedo is a measure of how much light is reflected. The more light is reflected, the less energy is absorbed. Snow and sea ice have a much greater albedo – they reflect more light – than rock or ocean. This means that more light is reflected. As polar snow and ice melts, more energy is absorbed by the darker surface beneath. This leads to an increased rate of heating.

The amount of energy radiated by Earth must ultimately match the energy it absorbs from the Sun. However, very little radiation from the surface can escape directly to space. Instead, it is mostly absorbed in the atmosphere by water vapour, CO_2 and other 'greenhouse gases'. It is then re-radiated, until it eventually escapes to space. The altitude of this 'final escape' is about 5 km; it is the temperature of the 'black body radiator' at this level, not at the surface, that would be seen by a space traveller looking at Earth. Here the atmosphere is much cooler than at the surface, so the atmosphere makes the black body radiator much less 'efficient' than if the radiation could escape directly from the surface. This is the 'natural' greenhouse effect, and keeps the surface about 33°C warmer than it would otherwise be, at a comfortable 15°C on average.

Human emissions raise the height at which this black body radiation occurs, because they introduce more greenhouse gases, such as CO_2 and CH_4. This makes Earth an even less efficient radiator, so the whole planet and atmosphere must warm up to restore the balance: greenhouse warming!

The models that Dr Colman and his team use have to include these effects. Climate modelling is a challenging task because there are so many factors that affect the climate. These include changes in the atmosphere, oceans, ice sheets and land surface. Modelling these complex interactions requires programs with millions of lines of code, run on the world's most powerful supercomputers. It also requires major new collaborations between atmospheric physicists, oceanographers, plant biologists, mathematicians and computer scientists.

To test the models, scientists use them to simulate observed features of recent climate and past climate changes. If a model can accurately reproduce measured temperatures and rainfall from the past, it raises confidence that the model will be able to predict temperatures and rainfall in the future. Climate models are a lot like the weather models used to forecast weather, but climate models make projections for decades into the future.

Figure 7.12 shows the results from two sets of models compared to measured data for temperature variations in the last hundred years. The black lines show the temperature anomaly. Temperature anomaly is the difference between recently recorded temperatures and long-term average temperatures. The lower set of models, shown in blue, do not include anthropogenic (human-caused) effects. The upper set of models, shown in orange, does include these effects and fits the data much better. This shows that models do not reproduce the observed warming unless the effects of greenhouse gases emitted as a result of human activities are included. Multiple sources of evidence, including results such as this, have convinced climate scientists that humans have significantly changed Earth's climate, and that even larger changes are to come.

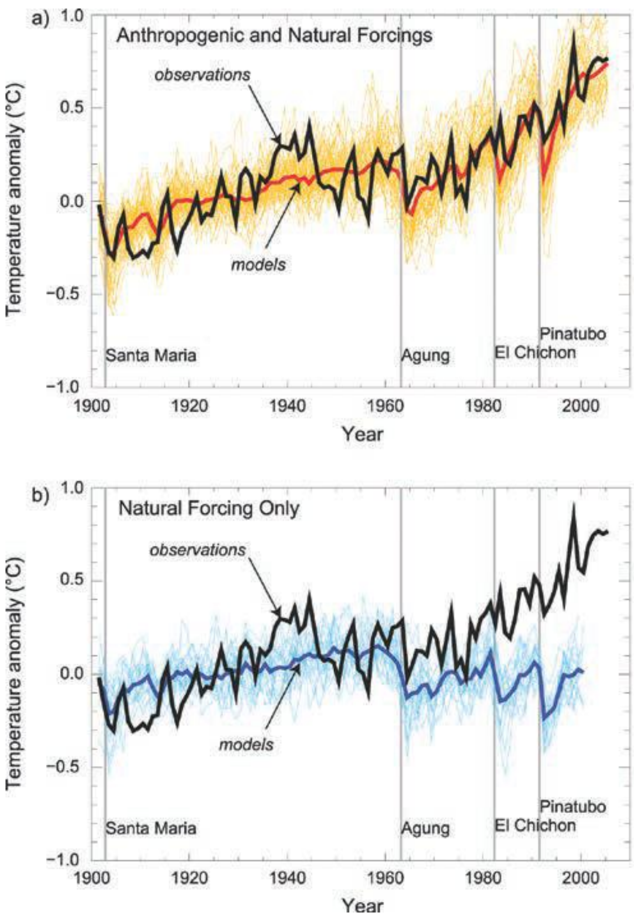


Figure 7.12 ▲
Models of climate change that a) do not include anthropogenic effects and b) include anthropogenic effects. The black lines are measured data and the coloured lines are simulations from the models. The heavy coloured lines are averages from all the individual models. 'Temperature anomaly' is the difference in global temperature from the long-term mean. Grey vertical lines show when prominent volcanoes erupted.

Questions

1. Define 'anthropogenic'.
2. If a year is colder than usual, would you expect the temperature anomaly for that year to be positive or negative? What if it was hotter than usual?
3. How have advances in information technology contributed to improved climate modelling?
4. What effect do you think clouds have on the albedo of Earth? How do you think this affects the solar energy absorbed by Earth? What about the energy radiated?
5. What effect do you think large volcanic eruptions might have on the climate?
6. There is still debate in the media and among politicians about the causes and even the existence of climate change. Research and evaluate the arguments put forward by climate-change sceptics. Reflect on your own views on this topic. Write a brief summary of your position in the debate, and include evidence to support your opinions.

QUESTION SET 7.2

Remembering

1 Name two classical models in physics.

2 Define 'black body'.

Understanding

3 Show that the units for Planck's constant must be J s.

4 Vega is a blue star and Antares is a red star. Which is hotter? Explain your answer.

Applying

5 Imagine an oven being used to bake a cake. What is the most likely wavelength of light in the oven?

6 An atomic oscillator has frequency $f = 6.1 \times 10^{12}$ Hz, and is in the $n = 3$ state.

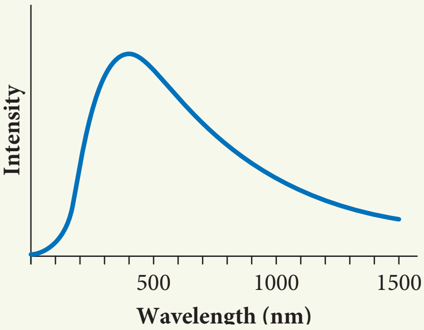
a What is the energy of this oscillator?

b What frequency light will be emitted if it transitions to the $n = 2$ state?

7 Figure 7.13 shows the black body radiation spectrum for the star Vega.

a What is the peak wavelength?

b Calculate the surface temperature of Vega. Give your answer to an appropriate number of significant figures.



▲ Figure 7.13

Analysing

8 The filament of an incandescent light globe can be modelled as a black body. A tungsten filament reaches a temperature of 2900 K.

a What wavelength does it emit most strongly?

b Explain why such light globes emit more radiation in the infrared than in the visible part of the electromagnetic spectrum.

Reflecting

9 Follow the weblink to the IPCC article. Read the summary of at least one article and reflect on how climate change might affect your life in the future. Consider the impact on where you live, and on where your food, water and electricity come from.

10 For climate scientists, the evidence of anthropogenic global warming is clear. However, many people, including some government and industry leaders, do not accept the evidence. What social, cultural and economic factors may be important in decision and policy making about the climate? Why do you think that some politicians are reluctant to make policy decisions that require action to be taken to mitigate the risks associated with climate change?

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE

Read the summary of one or more of the publications listed.

Quantisation and the photoelectric effect

As we have seen, Planck introduced the idea of quantised electromagnetic energy, light, as a mathematical trick to explain the black body spectrum.

It was already known at the time that matter was quantised. Scientists accepted that matter came in discrete quanta, or atoms. The atoms combine to form molecules, and so on. In 1897, J.J. Thompson discovered electrons when he realised that cathode rays were made of tiny negatively charged particles that were pieces of atoms. So, although it was known that atoms could be broken down into smaller components, these components were themselves quantised into discrete bits or particles. This was just three years before Planck published his theory of black body radiation.

In Chapters 9 and 10 you will look at the discovery of many more particles, and how they are related, when you explore the Standard Model.

When the quantum model was being developed, the idea of quantisation of matter was already well established and accepted, but the idea of quantisation of energy was completely new. It contradicted the accepted model of light as a wave.

The photoelectric effect provided the evidence needed for quantisation of energy to be accepted as more than just a mathematical trick.

The photoelectric effect

The photoelectric effect was first observed by Heinrich Hertz in 1887. He observed that when light is shone on a highly polished metal surface, electrons can be emitted from the surface. One of Hertz's assistants, Philipp Lenard, performed experiments to investigate the photoelectric effect in detail. Lenard developed much of the equipment needed to make quantitative measurements of the intensity and energy of the emitted 'cathode rays', as they were called at the time. Other physicists, including Robert Millikan, also investigated the effect.

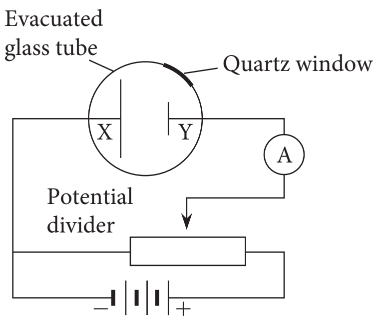


Figure 7.14 ▲ A typical set-up for a photoelectric experiment

Their data showed that:

- no electrons were emitted unless the frequency of the light was above some minimum or critical frequency, regardless of the intensity of the light.
- the number of electrons (the current) was proportional to the intensity. It did not vary with the frequency of the light, as long as the frequency was high enough.

A photoelectric effect apparatus is shown in Figure 7.14.

When light is shone through the quartz window at the polished metal plate, X, **photoelectrons** are emitted. The photoelectrons are attracted to the positively charged metal plate, Y. The ammeter, A, measures the current of photoelectrons produced – the **photocurrent**.

Using this apparatus, experiments show that:

- A photocurrent is only produced when the frequency of the light is above some minimum value, called the **cut-off frequency, f_0**. This implies a cut-off wavelength, $\lambda_0 = \frac{c}{f_0}$, above which no photocurrent is produced.
- The size of the current (the number of photoelectrons produced) depends on the intensity of the light but not the frequency, as long as the frequency is above f_0.
- There is no time delay between light being incident on the metal and photoelectrons being emitted, regardless of intensity.
- Different metals have different characteristic cut-off frequencies.

The voltage divider is used to vary the potential difference between X and Y. When the potential difference is reversed, the maximum kinetic energy of the emitted photoelectrons can be measured. This is called a reverse bias voltage. In this case, plate Y is negative and repels the photoelectrons. As you know from Chapter 3, the electrons lose energy as they move from a point of higher potential to one of lower potential in the direction of an electric field. Hence, the electrons lose kinetic energy as they move in the direction of X to Y.

Recall from Chapter 3 that potential is potential energy per unit charge. As a negatively charged electron moves from a point of higher to lower potential, its potential energy increases so its kinetic energy must decrease.

The reverse bias voltage between X and Y is slowly increased and the current observed. When the current just drops to zero, the potential difference is equal to the maximum energy per unit charge of the electrons. This potential difference is called the **stopping voltage**, V_s. Hence, the potential difference times the electron charge is equal to the maximum kinetic energy of the photoelectrons.

$$KE_{max} = V_s q = V_s e$$

This experiment has the following results:

- The maximum kinetic energy (measured via the stopping voltage) depends on the frequency of light but *not* on the intensity, as shown in Figure 7.15.
- Different metals have different characteristic stopping voltages, which depend on frequency.

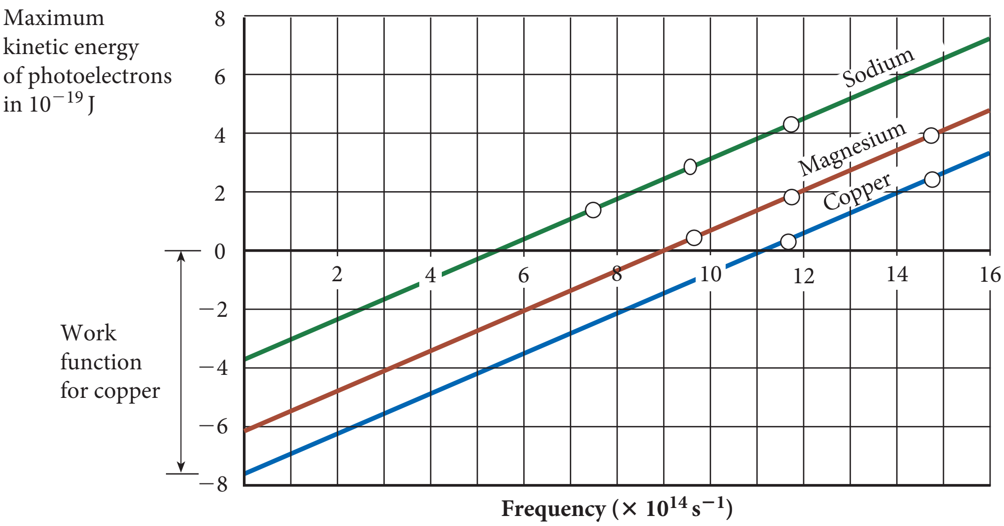


Figure 7.15 Maximum kinetic energy as a function of frequency for three different metals. The data is shown as circles and the lines are extrapolated to find the work functions.

The electromagnetic wave model of light cannot explain all of these observations. Table 7.1 compares the results of these experiments with the predictions of classical electromagnetic wave theory.

Table 7.1 Comparison of experimental results with predictions of classical electromagnetic wave theory

Experimental observation	Prediction from classical electromagnetic wave model
Photocurrent only occurs for frequencies above f_0, and f_0 is characteristic of the material.	Electrons should be emitted at any frequency, as long as the intensity is high enough.
The size of the current depends on intensity but not frequency.	The current should depend on both intensity and frequency.
There is no time delay between the absorption of light and emission of a photoelectron at any intensity.	At low intensities it takes time for enough energy to be absorbed by the atoms. Hence, there should be a delay between the light being turned on and electrons being emitted. The delay should be longer for lower intensities.
The maximum kinetic energy of the electrons depends on the frequency of light but *not* on the intensity.	The kinetic energy should only be related to the intensity, and not to the frequency.

It was for his explanation of the photoelectric effect that Einstein won his Nobel Prize, and not for the development of relativity (Chapter 6). Although relativity is a fascinating theory, in terms of impact on our daily lives it is the quantum theory that has made the most difference.

Just as with black body radiation, a new theory was needed to explain the results.

It was Einstein who came up with a new model in 1905. His explanation combined two ideas – the very familiar one of conservation of energy, and Planck's more recently introduced idea of quantisation.

Conservation of energy

You are already familiar with the idea of conservation of energy. Einstein explained the photoelectric effect by saying that electromagnetic radiation, or light, is quantised, or at least behaves as if it is quantised. When it interacts with matter, such as the metal plate X in Figure 7.14, it can only give up its energy in discrete amounts. Each quantum of light has energy

$$E = hf$$

where h is the Planck constant and f is the frequency of the light. This is the relationship between energy and frequency first introduced by Planck to explain black body radiation.

Recall from chemistry that an ion is a charged particle, usually an atom that has either lost or gained electrons. The ionisation energy is the energy needed to remove an electron from an atom. In the case of the photoelectric effect, electrons are lost from the whole metal lattice rather than from individual atoms.

When an electron in the metal plate X (Figure 7.14) absorbs a **photon**, it gains this energy. However, to leave the metal plate costs it an amount of energy; effectively an ionisation energy. Hence, the cut-off frequency, which is characteristic of the metal, is a measure of this ionisation energy. This energy is called the **work function** of the metal, and is given by

$$W = hf_0$$

where W is the work function, h is Planck's constant and f_0 is the cut-off frequency.

Putting this together with conservation of energy, Einstein said that:

$$KE_{max} = hf - hf_0 = hf - W$$

This is the photoelectric equation. It says that if an electron absorbs light energy hf, and is emitted from the metal, it will have a maximum kinetic energy of hf, minus the energy needed to leave the plate, which is $W = hf_0$.

Looking again at Figure 7.15, we can now see that the gradient of each line must be equal to Planck's constant. The extrapolated straight lines of best fit in Figure 7.15 cross the y axis at the value of the work function. This graphical representation of the experimental data allows us to quickly find values for both the Planck constant and the work function of the metal used.

Table 7.2 Work functions of some metals

Metal	W (eV)	W (J)
Na	2.46	3.94×10^{-19}
Al	4.08	6.53×10^{-19}
Fe	4.50	7.20×10^{-19}
Cu	4.70	7.52×10^{-19}
Zn	4.31	6.90×10^{-19}
Ag	4.73	7.57×10^{-19}
Pt	6.35	1.02×10^{-18}
Pb	4.14	6.62×10^{-19}

Note: these are typical values for these metals. Measured values vary depending on whether the metal is a single crystal or polycrystalline. For single crystals the value also depends on which face of the crystal is illuminated.

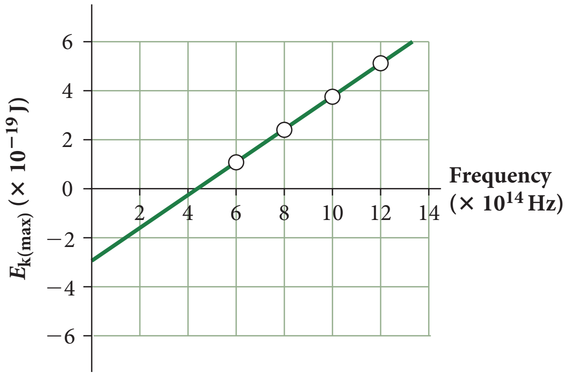
WOW

Millikan's experiments

Einstein's hypothesis that the energy was quantised was confirmed by Robert Millikan. Millikan did not initially believe Einstein's explanation for the photoelectric effect and spent 10 years performing experiments to test it. He was a brilliant experimentalist and had to develop techniques and equipment in order to perform experiments precise enough to really test Einstein's model. Using his equipment, he produced a graph similar to that shown in Figure 7.15. He found that the gradient of his graphs was consistently within uncertainty of Planck's constant. He eventually came to the conclusion that Einstein's model explained the experimental data, and that no other existing model did. He wrote in his autobiography that his data 'scarcely permits of any other interpretation than that which Einstein had originally suggested'. It is interesting to compare Millikan's eventual acceptance of the photon model with the Bohr–Einstein debate over the probabilistic nature of quantum mechanics. This is described later.

WORKED EXAMPLE 7.3

Using the graph in Figure 7.16 below, find the value of the work function for caesium. Give your answer in electron-volts and joules. (5 marks)



◀ **Figure 7.16**
A plot of photoelectron maximum kinetic energy as a function of frequency of incident light for caesium

Answer	Logic	
$f_0 = 4.4 \times 10^{14}$ Hz	Read f_0 from the graph.	1 mark
$W = hf_0$	Relate f_0 to the work function.	1 mark
$W = 6.63 \times 10^{-34}$ J s $\times 4.4 \times 10^{14}$ Hz	Substitute numbers including units.	1 mark
$W = 2.9 \times 10^{-19}$ J	Calculate the final value.	1 mark
$W = \frac{2.9 \times 10^{-19}\ \text{J}}{1.6 \times 10^{-19}\ \text{J eV}^{-1}} = 1.8$ eV	Convert from J to eV.	1 mark

Try this yourself

Use the graph in Figure 7.16 to find a value for Planck's constant, h. (4 marks)

You may have noticed that we have been talking about the *maximum* kinetic energy of the photoelectrons. The photoelectrons have all values of energy *up to* this maximum value. After absorbing the energy hf from the light, the electrons have energy hf. Recall from Chapter 5 of *Nelson Physics Units 1 & 2 for the Australian Curriculum* that in a metal there are lots of conduction electrons. These are electrons that are free to move through the metal and are not bound to any particular atom. It is these conduction electrons that can be ejected as photoelectrons. The valence electrons, which are bound to the atoms, cannot gain enough energy to be ejected in this way.

Electrons that have absorbed $E = hf$ lose a minimum energy of $E = W$ to escape the material, if they are at the very surface and have no interactions with other electrons or nuclei as they escape. However, most of the electrons will lose some of the absorbed energy in collisions with other electrons and with interactions with the atomic nuclei in the metal. Many do not leave the metal at all because they lose all of this energy, which shows up as the increased temperature of the metal. Hence, there is a continuous spectrum of electron energies from zero to the maximum value of $hf - W$.

WORKED EXAMPLE 7.4

Ultraviolet light of wavelength 200 nm is incident on a polished silver plate. The work function for silver is 4.7 eV.

a What is the kinetic energy of the fastest moving electrons? (5 marks)

b What is the kinetic energy of the slowest moving electrons? (1 mark)

c What is the cut-off frequency for silver? (4 marks)

Answers	Logic	
a $KE_{max} = hf - W$	Relate maximum KE to other values given.	1 mark
$f = \frac{c}{\lambda}$	Relate frequency to wavelength.	1 mark
$KE_{max} = \frac{hc}{\lambda} - W$	Substitute the expression for f.	1 mark
$KE_{max} = \frac{(4.14 \times 10^{-15}\,\text{eV s})(3.0 \times 10^{8}\,\text{m s}^{-1})}{(200 \times 10^{-9}\,\text{m})} - 4.7\,\text{eV}$	Substitute values including units. W is in eV, so convert h to eV also.	1 mark
$KE_{max} = 1.5\,\text{eV}$	Calculate the final value.	1 mark
b 0 eV	The slowest moving electrons only just make it out of the metal.	1 mark
c $W = hf_0$	Relate the cut-off frequency to known values.	1 mark
$f_0 = \frac{W}{h}$	Rearrange for f_0.	1 mark
$f_0 = \frac{4.7\,\text{eV}}{4.14 \times 10^{-15}\,\text{eV s}}$	Substitute values with units.	1 mark
$f_0 = 1.1 \times 10^{15}\,\text{Hz}$	Calculate final value.	1 mark

Try this yourself

Draw a graph of maximum photoelectron kinetic energy as a function of frequency of incident light for silver. (4 marks)

EXPERIMENT 7.2

THE PHOTOELECTRIC EFFECT

It is possible to observe the photoelectric effect using some very simple equipment. You can even build some of the equipment yourself! Follow the weblink to find instructions for making your own electroscope.

Aim

To observe the photoelectric effect

Materials

- electroscope
- polished zinc plate
- steel wool or fine sand paper
- glass rod
- polythene rod
- fur or woollen fabric
- ultraviolet light source
- other light sources, for example lasers and torches

SOFT DRINK CAN ELECTROSCOPE

Build your own electroscope using a soft drink can.

What are the risks in doing this experiment?	How can you manage these risks to stay safe?
Ultraviolet light can damage your eyes.	Do not look directly at the ultraviolet light source. Turn it off when not in use.

In your write-up, add any more risks you can think of, as well as ways to manage them.

Procedure

1 Clean the zinc plate *thorough*ly with the steel wool or sand paper, then place it on the electroscope. Try not to leave any fingerprints on it as you do so.

2 Charge up the glass rod by rubbing it vigorously with the fur or wool.

3 Touch the rod to the zinc plate. You should observe the leaves of the electroscope separate.
4 Time how long it takes for the leaves to fall back together. If it takes more than two minutes, just record your result as 'more than two minutes' and discharge the electroscope by touching it with your hand. Remember to avoid touching the zinc plate.
5 Repeat steps 2 to 4, but this time shine one of your light sources on the plate.
6 Repeat steps 2 to 5 with each of your different light sources.
7 Charge up the polythene rod by rubbing it vigorously with the fur or wool.
8 Use the rod to charge the electroscope.
9 Time how long it takes for the leaves to fall back together.
10 Repeat steps 7 to 9, but this time shine a light source on the zinc plate. Repeat this for each light source.
11 Charge the zinc plate with the polythene rod and then shine the ultraviolet light on it from various distances away. This varies the intensity of the light on the plate.

Results

1 Draw a diagram showing your experimental set-up. Label all the parts clearly.
2 Record your results in a table, such as that shown below. The glass rod becomes positively charged. The plate is also positively charged when charged with the glass rod. The polythene rod becomes negatively charged. The plate is also negatively charged when charged with the polythene rod.
3 Don't forget to include units and uncertainties on your data.

Charge on plate	Light used	Time to discharge (s)

4 If you measured the discharge time using the ultraviolet light at various distances (step 11 above), record your results in a table like the one below.

Distance to light (m)	Time to discharge (s)

Analysis of results

1 Explain what is happening in each case you investigated.
2 Plot a graph of time to discharge as a function of distance to light, using the data in your second table. Do you expect this graph to be linear? If not, what shape do you expect it to be? Can you check whether it is?

Discussion

1 Discuss the shape(s) of your graph(s). Are they what you expected? If not, why might this be the case? Think about any assumptions you have made.
2 What could you do to improve the accuracy and precision of this experiment?
3 Can you think of any ways to extend it?

Photons: quantisation of energy

Einstein's explanation of the photoelectric effect gave a physical meaning to the idea of quantisation of energy of electromagnetic radiation. It meant that, in some circumstances, light behaved like particles. The term 'photon' was introduced by the chemist Gilbert Lewis in 1926 to describe these particles. Further experiments by Arthur Holly Compton gave evidence for the existence of photons. Compton scattered single photons from electrons and found that only particular energies were absorbed. Photons are now accepted as particles with zero rest mass, and with energy given by $E = hf$.

WORKED EXAMPLE 7.5

Find the range of energies of photons in the visible spectrum. The visible spectrum ranges from blue light with wavelength approximately 400 nm to red light with wavelength approximately 700 nm. (9 marks)

Answer	Logic	
$E = hf$	Relate energy to other values.	1 mark
$f = \frac{c}{\lambda}$	Relate frequency to wavelength.	1 mark
$E = \frac{hc}{\lambda}$	Substitute expression for f.	1 mark
$E_{max} = \frac{hc}{\lambda_{min}}$	Recognise that the highest energy corresponds to the smallest wavelength.	1 mark
$E_{max} = \frac{(4.14 \times 10^{-15}\ \text{eV s})(3.0 \times 10^{8}\ \text{m s}^{-1})}{400 \times 10^{-9}\ \text{m}}$	Substitute values including units.	1 mark
$E_{max} = 3.1\ \text{eV}$	Calculate the final value.	1 mark
$E_{min} = \frac{hc}{\lambda_{max}}$	Recognise that the lowest energy corresponds to the longest wavelength.	1 mark
$E_{min} = \frac{(4.14 \times 10^{-15}\ \text{eVs})(3.0 \times 10^{8}\ \text{m s}^{-1})}{700 \times 10^{-9}\ \text{m}}$	Substitute values including units.	1 mark
$E_{min} = 1.8\ \text{eV}$	Calculate the final value.	1 mark

Try this yourself

Photons in the visible spectrum have energies ranging from 1.8 eV to 3.1 eV. Convert these energies to J. (3 marks)

QUESTION SET 7.3

Remembering

1 What principle did Einstein use to explain the energies of photoelectrons?

2 What is a photon?

Understanding

3 How is a photoelectron different from any other electron?

4 Explain how you can find Planck's constant from a graph of frequency against stopping voltage from a photoelectric experiment.

5 It requires more energy to remove an electron from the surface of a polished piece of copper than from a polished piece of lithium.

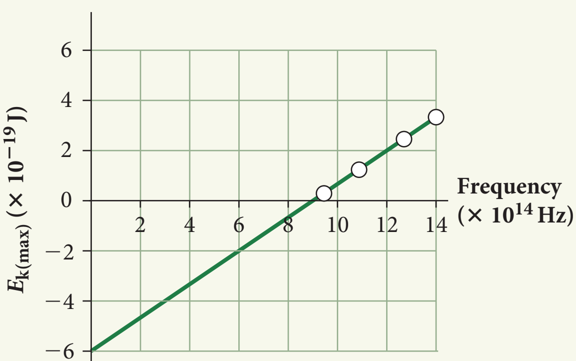
a Which metal has the larger work function?

b Which metal has the greater cut-off wavelength?

Applying

6 If a metal has work function W and is irradiated with light of frequency f, what are the possible energies of any emitted photoelectrons?

7 The graph in Figure 7.17 shows the results of a photoelectric experiment using a magnesium metal plate.



▲ Figure 7.17

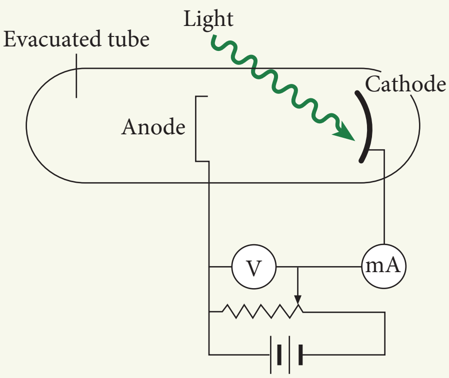
a Find a value for Planck's constant from this graph.

b Find a value for the work function of magnesium.

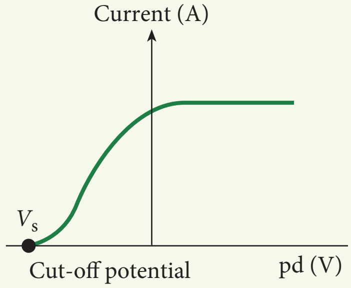
c Imagine that silver had been used in this experiment instead of magnesium. Silver has a work function of 4.7 eV. Draw a line on the graph showing where the data points for silver would be located.

Analysing

8 Figure 7.18 shows a photoelectric tube with light of frequency *f* and intensity *I* incident on a metal cathode. Electrons emitted from the cathode are collected at the anode. The potential difference between the anode and cathode is varied, and the resulting photocurrent is measured. Figure 7.19 shows the results of this experiment.



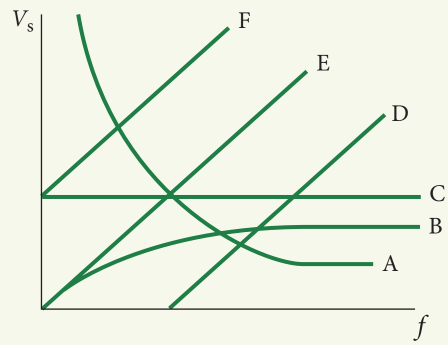
▲ Figure 7.18



▲ Figure 7.19
Photocurrent as a function of applied potential difference

a Why is the photocurrent constant at positive values of pd?

b If the frequency of the light is varied, which of the graphs in Figure 7.20 represents the relationship between the stopping voltage, V_s, and *f*?



◀ Figure 7.20

CHAPTER SUMMARY

- The quantum theory was developed when classical models, such as the wave model of light, were unable to explain experimental results such as the black body spectrum, atomic line spectra and the photoelectric effect.
- The development of quantum theory was a revolution in science, and not a smooth transition between theories. It was surrounded with great argument and debate.
- Black body radiation is emitted by all objects at non-zero temperature.
- Black body radiation has a spectrum characteristic of the temperature of the body. The black body spectrum is a continuous spectrum.
- The peak wavelength in a black body spectrum is given by Wein's displacement law: $\lambda_{max} = \frac{b}{T}$
- Black body radiation can be explained by the quantisation of electromagnetic energy. We call these quanta of energy 'photons'.
- Photons are 'particles' of electromagnetic energy. Each photon has energy $E = hf$.
- Earth behaves as a black body, and this is important in maintaining its temperature balance and climate.

- There is strong evidence from climate change modelling that anthropogenic effects are causing significant global warming.
- Economic and cultural factors prevent this evidence being accepted and acted on to reduce risks associated with climate change.
- The photoelectric effect is the ejection of electrons from a polished metal surface by incident light. The light must have a minimum frequency for this to occur.
- The minimum frequency corresponds to a minimum energy, $E = hf_0$.
- The minimum energy corresponds to the work function, W, of the metal.
- The maximum kinetic energy of the photoelectrons is $KE_{max} = hf - W$. This is a statement of conservation of energy.
- The photocurrent, which is proportional to the number of electrons, depends on the intensity of the light. The intensity is a measure of the number of incident photons.

CHAPTER GLOSSARY

albedo the ratio of light reflected by a surface to light incident on it; a surface with an albedo of 1 is perfectly reflective, and an albedo of 0 is perfectly absorbing

black body an object with a perfectly absorbing surface, which emits radiation with a spectrum that is characteristic of the temperature of the object

black body radiation the electromagnetic radiation emitted by a black body, with a spectrum characteristic of the temperature of the body

continuous able to take any value, sometimes within a fixed range, as distinct from discrete or quantised

continuous spectrum a spectrum containing radiation of all wavelengths; for example, a rainbow

cut-off frequency, f_0 the minimum frequency of light needed to eject an electron from a metal surface

discrete able to take only specific values, not continuous; for example, a line spectrum is a discrete spectrum

modes of vibration characteristic patterns of oscillation, usually with a discrete set of allowed frequencies

photocurrent the current formed by electrons ejected from a surface by incident photons

photoelectric effect the ejection of electrons from a surface by incident photons of sufficient energy

photoelectron an electron ejected from a metal surface following absorption of a photon of sufficient energy

photon a particle or quanta of light, having energy $E = hf$

Planck constant the constant of proportionality between energy and frequency for photons: $h = 6.626 \times 10^{-34}$ J s

quantised existing in discrete amounts, not able to be divided into arbitrarily small amounts

quantum a discrete unit or amount of some physical property, such as energy, charge, mass or angular momentum

spectrum the distributed components of light or another wave arranged by frequency (or wavelength)

stopping voltage the reverse bias voltage required to stop the flow of photoelectrons in a photoelectric effect experiment

uncertainty estimate of the range of values within which the 'true value' of a measurement or derived quantity lies; the extent to which the result of an experiment is unknown or unpredictable

work function the energy required to eject an electron from a metal surface; effectively, it is the ionisation energy for the bulk material

CHAPTER REVIEW QUESTIONS

Remembering

1 Name three physicists who contributed to the development of the quantum theory and briefly describe their contributions.

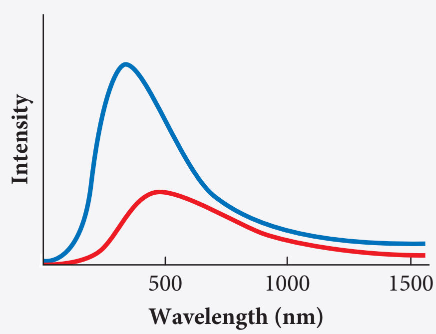
2 When was the photoelectric effect first observed, and by whom?

3 An incandescent light globe is connected to a variable power supply and the voltage gradually increased. Describe the sequence of colours produced by the filament of the globe.

Understanding

4 Explain why a wave model of light is needed to understand the interference pattern produced in the double-slit experiment. Give two other examples of wave-like behaviour of light.

5 Figure 7.21 shows the black body spectra for two stars A (red curve) and B (blue curve). Which star is hotter? Explain your answer.



▲ Figure 7.21

6 Why are fluorescent globes and LEDs so much more energy-efficient than incandescent globes?

7 Which metal in Table 7.2 would require the highest frequency light for photoelectrons to be emitted?

8 A polished lead surface is illuminated with light of wavelength 200 nm. What is the effect on the photocurrent if:
  a the wavelength is halved?
  b the intensity of the light is doubled?

9 Figure 7.22 shows a graph of maximum kinetic energy as a function of frequency for a particular metal. Copy the graph and, on the same set of axes, draw a line showing maximum kinetic energy for:
  a the same metal but with higher intensity light.
  b light of the same intensity but for a metal with a larger work function.

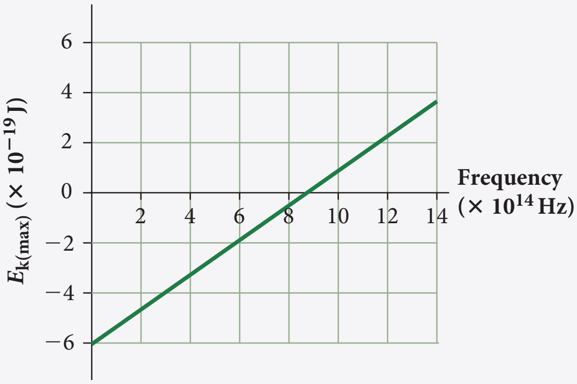


Figure 7.22

Applying

10 In a double-slit experiment, light with wavelength 589 nm is used to illuminate twin slits that are separated by 0.015 mm. The pattern produced is observed on a wall 2.00 m from the slits. Find the position of:
  a the first interference maximum (bright spot).
  b the first interference minimum (dark spot).
  c the second interference maximum (bright spot).

11 What is the peak wavelength emitted by a toaster element at 700°C? What colour would you expect it to be?

12 Cosmic background radiation (see Chapter 10) has a spectrum similar to that produced by a black body at 2.7 K. What is the peak wavelength of this radiation?

13 A photon has energy 5.5 eV. What is its:
  a energy in J?
  b wavelength?
  c frequency?

14 A polished sodium surface is illuminated with light.
  a Find the cut-off wavelength for sodium. What colour does this correspond to?
  b Find the maximum kinetic energy of ejected photoelectrons when light of wavelength 300 nm is used.

15 The graph shown in Figure 7.23 shows the results of a photoelectric experiment using an unknown metal. Calculate the work function for this metal. What is this metal?

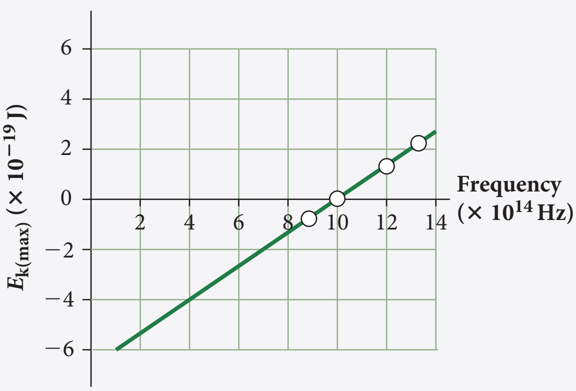


Figure 7.23

Analysing

16 In a double-slit experiment light with wavelength 589 nm is used to illuminate twin slits. The pattern produced is observed on a wall 2.00 m from the slits. The tenth interference minimum (dark spot) is found to be at a position 7.26 mm from the central bright spot. Calculate the slit separation.

17 A microwave oven produces electromagnetic radiation with a wavelength of 12.2 cm.

- **a** What energy does this correspond to? Give your answer in J and eV.
- **b** Is this enough energy to produce a photocurrent from a metal surface? Why do you think it might be a bad idea to put metal objects in a microwave? *Caution: Don't try it!*

18 The following data was collected by students doing an experiment on black body radiation.

- **a** Plot a graph showing power as a function of wavelength.
- **b** Calculate the temperature of the black body that produced this radiation. Explain your answer.

Wavelength (nm)	Power at detector (W)
100	0
200	0.01
300	0.18
400	0.30
500	0.34
600	0.32
700	0.27
800	0.21
900	0.17
1000	0.14
1100	0.10
1200	0.09

19 The table below shows data collected in a photoelectric experiment. Plot an appropriate graph and find:

- **a** Planck's constant.
- **b** the work function for this metal.

Wavelength (nm)	KE_{max} of photoelectrons (eV)
588	0.67
505	0.98
445	1.35
399	1.63

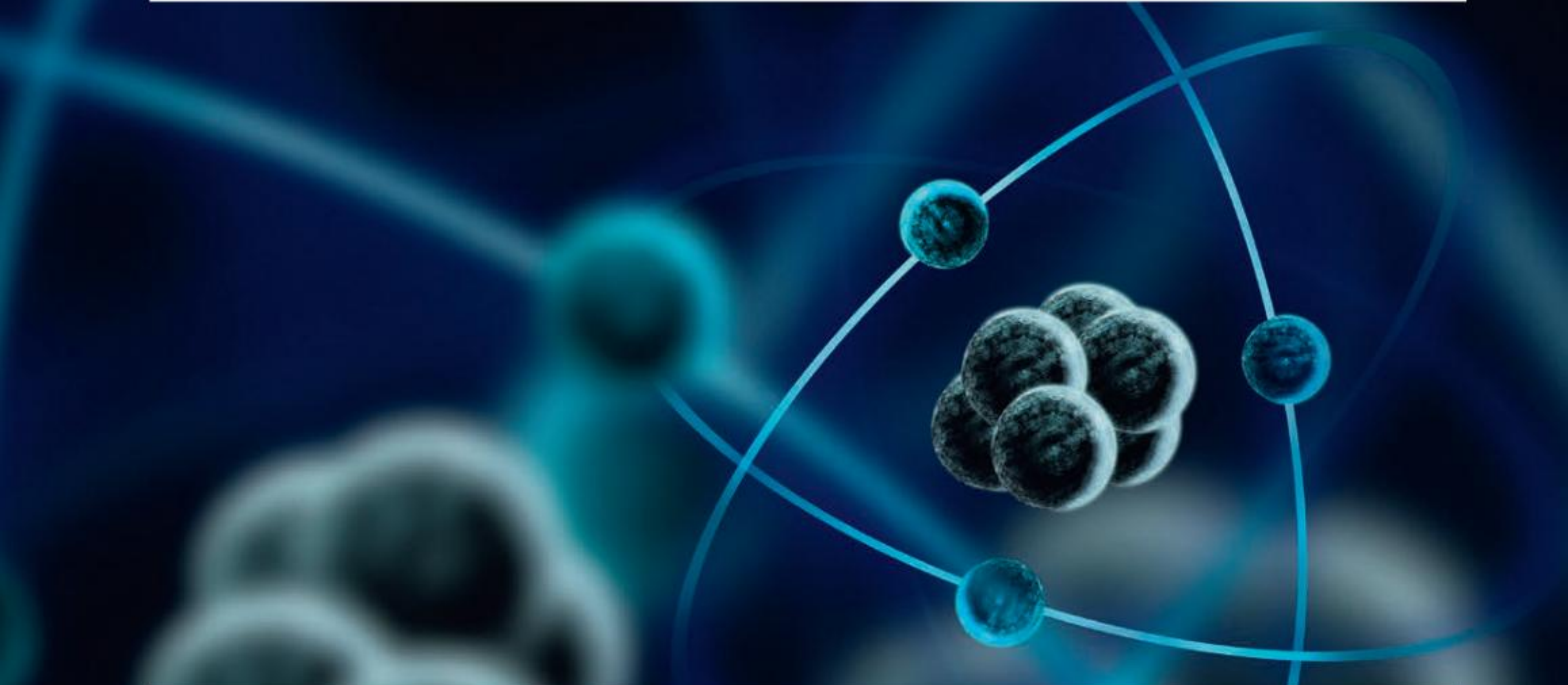
Reflecting

20 Research Robert Millikan. Why did Millikan initially object to Einstein's explanation of the photoelectric effect? Why did he eventually accept the photon explanation?

21 Why does the decrease in the Earth's albedo result in more extreme weather, such as storms, floods and droughts? You will need to think about what you learnt when you studied thermal physics, as well as what you have learnt in this chapter.

22 Research on the internet the views of climate sceptics and climate scientists. What evidence does each group use to support their opinions? Evaluate the arguments put forward by both sides and summarise and justify your own opinion.

CHAPTER 8
QUANTUM THEORY AND MATTER



Shutterstock.com/solarseven

By the end of this chapter you will have covered the following material.

Science Understanding

- Atomic phenomena and the interaction of light with matter indicate that states of matter and energy are quantised into discrete values (ACSPH135)
- Atoms of an element emit and absorb specific wavelengths of light that are unique to that element; this is the basis of spectral analysis (ACSPH138)
- The Bohr model of the hydrogen atom integrates light quanta and atomic energy states to explain the specific wavelengths in the hydrogen spectrum and in the spectra of other simple atoms; the Bohr model enables line spectra to be correlated with atomic energy-level diagrams (ACSPH139)
- On the atomic level, energy and matter exhibit the characteristics of both waves and particles (for example, Young's double-slit experiment is explained with a wave model but produces the same interference pattern when one photon at a time is passed through the slits) (ACSPH140)

Introduction

In the previous chapter we saw that observations of black body radiation and the photoelectric effect led to the development of the photon (particle) model of light. At the same time that Planck, Einstein and others were developing this new quantum model of light, other physicists were investigating the nature of atoms.

The idea that matter consists of atoms dates back to ancient Greece. Until the end of the 19th century it was believed that atoms were indivisible. Then in 1897, J.J. Thomson experimentally observed electrons being emitted from atoms. This established that an atom was not a single homogenous particle, but consisted of matter with both positive and negative charges. At that time it was not known how the positive and negative parts were arranged.

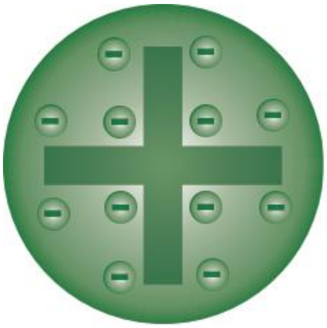


Figure 8.1 ▲ Thomson's plum pudding model of the atom

Measurements of the mass of electrons showed that almost all of the mass of an atom was associated with the positive part of the atom. Based on this evidence, in 1904 Thomson suggested a model in which the atom consisted of a sphere of positive charge with electrons embedded in it. This model was known as the 'plum-pudding' model of the atom, as the electrons were scattered through the positive matter, like raisins scattered through a plum-pudding (or chocolate chips in a chocolate chip muffin).

Then, in 1909 Hans Geiger and Ernest Marsden, who were students working with Ernest Rutherford, performed the alpha particle scattering experiment you may remember from Chapter 3 of *Nelson Physics Units 1 & 2 for the Australian Curriculum*. A thin gold foil was placed in a beam of alpha particles and the paths of the scattered alpha particles were measured. What they found was that, although most alpha particles were not deflected by much, about one in 20 000 was scattered backwards towards the source. This result implies the presence of a tiny and very dense positively charged nucleus. Rutherford later said 'It was as if you fired a 15-inch shell at a sheet of tissue paper and it came back to hit you.'

Rutherford suggested a planetary model of the atom to explain these observations. This model involved electrons orbiting the nucleus in a circular path under the influence of the electrostatic force. This was analogous to the way planets orbit the Sun under the influence of the gravitational force.

The electrostatic force was described in Chapter 3. The orbit of planets around the Sun due to the gravitational force was described in Chapter 2.

Although Rutherford's model fitted the experimental data from the gold-foil scattering experiment, there were still problems. In this planetary model, the electrons (charged particles) were undergoing centripetal acceleration. This is not a problem for uncharged objects like planets, but accelerating charged particles emit energy. The electrons should be acting like those in an antenna and emitting electromagnetic waves, as described in Chapter 5. Remember that energy is always conserved. If the electron was emitting energy as electromagnetic radiation, then the electron would lose kinetic and potential energy and spiral into the nucleus and the atom would collapse. This was a significant flaw with this model.

None of the classical models of the atom, such as those described above, were able to fully explain or predict the behaviour of atoms. New models and theories were needed.

Experimental data showing that energy in atoms is quantised would prove vital for developing the new quantum mechanical atomic theory. This data came from measurements of atomic line spectra.

Getty Images/De Agostini

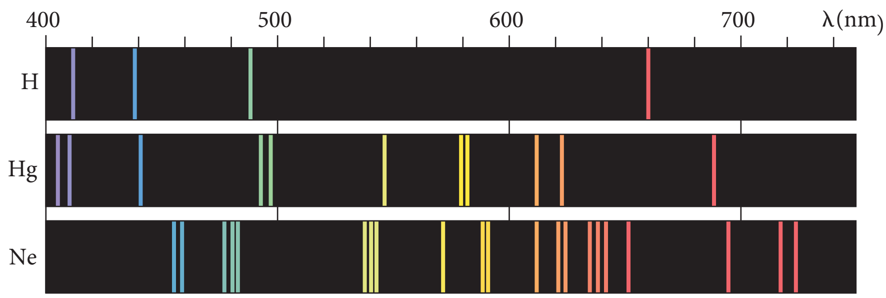
Figure 8.2 ▲ In a spectroscope a prism is used to disperse the light. A diffraction grating can be used in place of the prism to cause dispersion.

Atomic spectra

A **spectroscope** uses a prism or diffraction grating to separate, or disperse, light into its component colours. White light disperses into all the colours of the rainbow. We have already seen in the previous chapter that black bodies produce a continuous spectrum. In contrast, when a gas is heated it produces a spectrum consisting of discrete colours. This is observed through a spectroscope as a pattern of parallel coloured lines and hence is called a **line spectrum**. A heated gas produces an **emission spectrum**. When white light is passed through a cold gas an **absorption spectrum** is produced. It has the same characteristic lines, but they are dark on a continuous coloured background.

Gustav Kirchhoff and Robert Bunsen had recognised as early as the 1860s that line spectra can be used to identify elements. They discovered two new elements, caesium and rubidium in 1861 using spectroscopy.

▲Figure 8.3
Neon lights are an example of a gas discharge tube. The electrical energy is transformed into light.



▲Figure 8.4
Emission spectra of hydrogen, mercury and neon

Although Kirchhoff, Bunsen and others had observed and used spectra, there was no theory that explained why they existed. However, it was presumed that the characteristic spectra were related to the internal structure of the atom. To solve this puzzle, the simplest atom, hydrogen, was subject to intense theoretical and experimental investigation.

ACTIVITY 8.1

OBSERVING SPECTRA

Aim

To observe the spectra from different light sources using a diffraction grating

You will need

A diffraction grating or pair of diffraction grating glasses and two or more light sources – LEDs and fluorescent lights are good sources. *Do not use a laser.*

What are the risks in doing this experiment?	How can you manage these risks to stay safe?
Looking directly at the Sun through a diffraction grating can cause eye damage.	Never look directly at the Sun. Only observe sunlight indirectly.

BUILD YOUR OWN SPECTROSCOPE

Build a simple spectroscope from a CD and a cardboard box by following this guide.

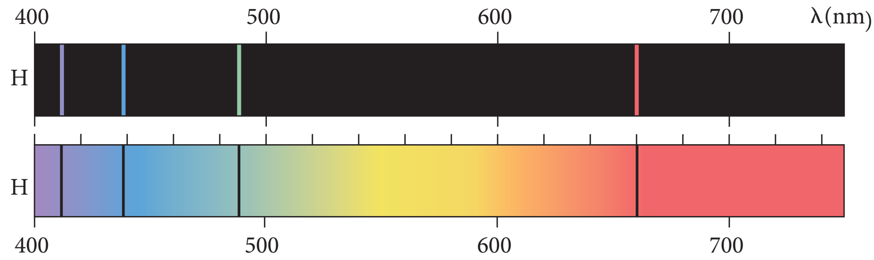
What to do

In an otherwise darkened room, look through the grating(s) at the light source. Note what you see. Do you see a complete spectrum or are there some colours that are missing or very bright? Compare the spectrum of a fluorescent light with that of an incandescent light globe. You should be able to see some bright green and purple lines from a fluorescent light.

The hydrogen spectrum

Hydrogen produces infrared, visible and ultraviolet emission spectra. The emission and absorption spectra of hydrogen are shown in Figure 8.5. Dark lines in the absorption spectrum of any element coincide with the bright lines in its emission spectrum.

Figure 8.5 ▶
The emission and absorption spectra of hydrogen



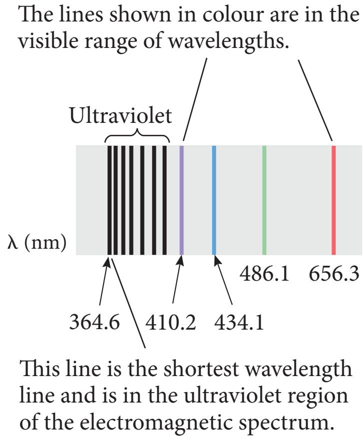


Figure 8.6 ▲
The Balmer series of spectral lines for atomic hydrogen

In 1855, Johann Balmer derived an empirical formula for the visible series that now bears his name.

Balmer showed that the observed wavelengths were proportional to $\frac{m^2}{(m^2 - n^2)}$ with $n = 2$ and m greater than n.

The other spectral series of hydrogen are named after their discoverers. The Lyman series is in the ultraviolet and the Paschen series is in the infrared part of the spectrum. Other series of even longer wavelength are the Brackett series and the Pfund series. These lines have a similar pattern of separation, but with different values for n and m.

Table 8.1 Lines in the spectral series of hydrogen

Series name	Part of spectrum	Shortest wavelength (nm)	Longest wavelength (nm)
Lyman	Ultraviolet	91.1	121.6
Balmer	Visible	364.5	656.3
Paschen	Infrared	820.1	1870

In the 1880s Johannes Rydberg was working on finding a mathematical description of the line spectra of alkali metals (lithium, sodium and so on). He read Balmer's work on hydrogen and realised that his own mathematical model and Balmer's were equivalent. Rydberg expressed the relationship as:

$$\frac{1}{\lambda} = R\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right)$$

where λ is the wavelength of the line, n_1 and n_2 are integers with $n_1 < n_2$. R is a constant known as the Rydberg constant, $R = 1.097 \times 10^7\ \text{m}^{-1}$.

Rydberg arrived at his formula empirically; that is, by fitting an equation to the observed data. At the time there was no theoretical model of the atom that could predict the relationship between positions of spectral lines, or even the existence of spectral lines.

His work was important in that it led to the development of the first quantum mechanical model of the atom – the Bohr model.

WORKED EXAMPLE 8.1

For the Brackett series in the far infrared, $n_1 = 4$.

a Find the longest wavelength in the series. (5 marks)

b Find the shortest wavelength in the series. (7 marks)

Answers	Logic	
a $\frac{1}{\lambda} = R\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right)$	Relate wavelength to other variables.	1 mark
$\frac{1}{\lambda} = 1.097 \times 10^7 \text{m}^{-1}\left(\frac{1}{4^2} - \frac{1}{5^2}\right)$	Insert numerical values, noting that for the longest wavelength the quantity in brackets must be the smallest, hence n_2 must have its minimum value.	2 marks
$\frac{1}{\lambda} = 2.47 \times 10^5 \text{ m}^{-1}$	Calculate a value for $\frac{1}{\lambda}$.	1 mark
$\lambda = 4050 \text{nm}$	Calculate λ.	1 mark
b $\frac{1}{\lambda} = R\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right)$	Relate wavelength to other variables.	1 mark
$\frac{1}{\lambda} = R\left(\frac{1}{n_1^2} - \frac{1}{\infty^2}\right)$	Recognise that for the shortest wavelength the quantity in brackets must be the smallest, hence n_2 must have its maximum value. As there is no upper limit to n_2 we take $n_2 = \infty$.	2 marks
$\frac{1}{\lambda} = R\left(\frac{1}{n_1^2}\right)$	Simplify the equation.	1 mark
$\lambda = \frac{n_1^2}{R} = \frac{4^2}{1.907 \times 10^{-7} \text{m}^{-1}}$	Rearrange the equation to make λ the subject and substitute values.	2 marks
$\lambda = 1460 \text{nm}$	Calculate λ.	1 mark

Part a is one of the very rare occasions where it is sensible to substitute numbers into an equation before rearranging for the desired quantity. In general, it is better to rearrange first. However, in this case we are finding the reciprocal of a lengthy expression, so it is acceptable to substitute first. But be very careful doing this! Don't forget to find λ, not $\frac{1}{\lambda}$.

The Bohr model

In 1913, Niels Bohr combined the concepts of Rutherford's planetary model of the atom and Einstein's photons to predict the observed spectra of hydrogen. To solve the problems of the planetary model, Bohr made several postulates.

Bohr's postulates

1. An electron in an atom moves in a circular orbit about the nucleus under the influence of the electrostatic attraction of the nucleus.
2. Only certain orbits are stable. Electrons in these orbits do not emit energy.
3. The greater the radius of the orbit, the greater is its energy. Atoms emit radiation when an electron goes from one orbit to another orbit with lower energy. The energy released is:

$$E = E_i - E_f = hf$$

4 **The orbits are characterised by quantised radii, given by**

$$r = \frac{nh}{2\pi m_e v}$$

where r is the radius in m, m_e is the mass of the electron in kg, v is its velocity, h is Planck's constant and n is an integer.

You may have come across these n numbers as shell numbers in chemistry. In physics, n is known as the principle quantum number.

These postulates were a mixture of classical physics (postulate 1), recently introduced quantum principles (postulate 3) and completely new ideas (postulates 2 and 4).

Bohr's first postulate is drawn directly from the earlier planetary model of the atom. His second postulate is simply stating what had been observed – that atoms do not collapse. It does not explain it; it just states it as a given. Bohr's third and fourth postulates are the ones that distinguish his model as the first quantum model of the atom.

Bohr's third postulate states that energies are quantised. They may take only a discrete set of values. The relationship between radius and energy is given by classical electromagnetism, as described in Chapter 5. The greater the separation between a positive charge (the nucleus) and a negative charge (the electron), the greater the potential energy of the system. The different possible energies are called **energy levels**. Using classical electromagnetism Bohr showed that the energy of any given level was proportional to $\frac{1}{n^2}$, where n is the integer in the equation for the radius in postulate 4. Hence:

$$E_n = \frac{k}{n^2}$$

where k is a constant.

Bohr's fourth postulate was also based on the quantisation of a physical property, in this case the **angular momentum, *L***, of the electron. Angular momentum is to circular motion (such as orbiting planets) what momentum is to linear motion. It is a conserved quantity and is given by:

$$L = mvr$$

where L is the angular momentum of the object, m is its mass, v is its velocity and r is the orbital radius.

You have met various conservation principles already, and you will meet more in the next chapter. Conservation of angular momentum is a particularly useful conservation principle in understanding the behaviour of planets and satellites, and the interaction of elementary particles.

By saying that L may only have discrete values or that:

$$L = mvr = \frac{nh}{2\pi}$$

Bohr was saying that only discrete values of r were allowed. This in turn means that only discrete values of energy are allowed, as stated in postulate 3.

We often refer to these energy levels as electron energies, but remember that this energy belongs to the electron–nucleus system because they are separated charged particles. All isolated atoms of one element will have the same set of energy levels, but different elements have different sets because they have different nuclei and different numbers of electrons. An atom can make a transition from one level to a lower level by emitting a photon of energy equal to the difference between the levels. This occurs when an electron moves from an orbit further from the nucleus to one closer to the nucleus. To move from a lower energy level to a higher level an electron must absorb energy. Figure 8.7 shows the Bohr model of the atom. The force directed towards the nucleus is the Coulomb force, as described in Chapter 3.

The Bohr model was able to explain the existence of discrete line spectra. Emission spectra can be explained by electrons moving from higher to lower energy orbits. The energy gap between the energy levels is equal to the energy of the photon emitted when the transition occurs:

$$E_{gap} = E_i - E_f = hf_{photon}$$

and recalling that $E_n = \frac{k}{n^2}$ we can write:

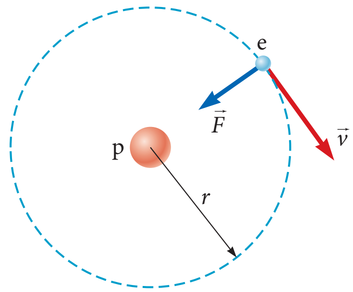
$$hf_{\text{photon}} = E_i - E_f = k\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right)$$

If we now use the dispersion relation for light, $f = \frac{c}{\lambda}$, we can see that:

$$\frac{1}{\lambda} = \frac{k}{h}\left(\frac{1}{n_f^2} - \frac{1}{n_i^2}\right)$$

You should recognise this equation as being of the same form as Rydberg's equation. In fact Bohr showed that the constant $\frac{k}{h}$ was equal to the Rydberg constant, R.

This was a very important success for the Bohr model. Not only did it predict the existence of line spectra, but it quantitatively predicted the positions of the lines for the hydrogen atom.



▲Figure 8.7
In Bohr's model of the hydrogen atom, the electron occupies discrete orbits.

The hydrogen spectrum revisited

Bohr's model explained the observed line spectra as resulting from transitions between energy levels in atoms. The lowest wavelength or highest energy line corresponds to the ionisation energy of electrons in the lowest possible energy level. This level is called the **ground level** of the atom and corresponds to the electron being in the orbit with the smallest radius. Ionisation is the removal of an electron from the atom to infinitely far away, or at least so far that the electrostatic attraction is negligible.

Energy levels for electrons in atoms were described in Chapter 3 of Nelson Physics Units 1 & 2 for the Australian Curriculum. The structure of the levels determines a material's conductivity when the atoms are bound in a solid.

Energy levels can be represented in two ways:

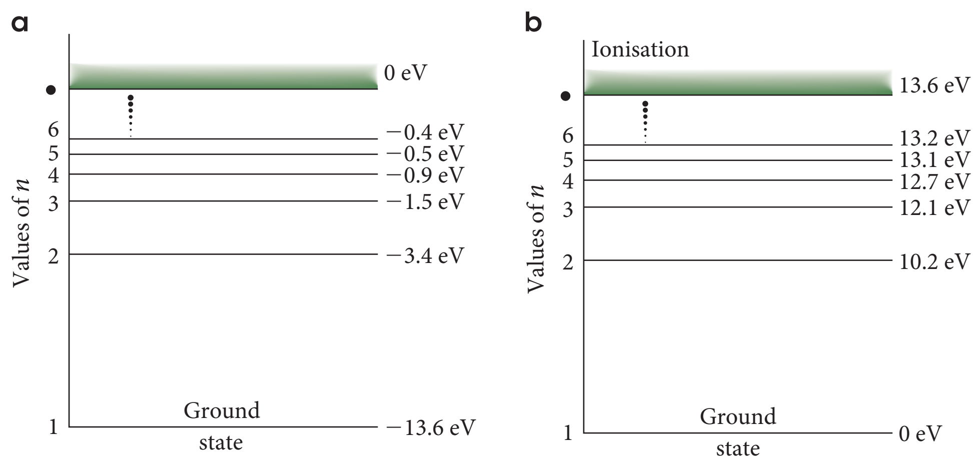
1 with the ionisation energy being taken as zero (all the energy states then have a negative potential energy)

or

2 with the ground state level being taken as zero.

In both representations, the energy difference between levels is the same. We shall generally use the first representation, as this corresponds to the usual convention for choosing the zero of potential energy as in electromagnetism.

Recall from Chapters 2 and 3 that the zero of potential energy for systems of objects that interact via a field is taken as infinite separation of all the objects in the system.



▲Figure 8.8
The energy levels of the hydrogen atom have a) the ionisation energy as zero and b) the ground state as zero.

The energy of each level can be deduced from the wavelengths, and hence energies, of the lines in the emission spectra. The highest energy lines for hydrogen are those in the Lyman series in the ultraviolet. These lines correspond to transitions to the ground state from higher energy levels, hence for these lines $n_f = 1$ and $n_i = 2, 3, 4 \ldots$ The Balmer series, in the visible region, corresponds to transitions to the $n = 2$ level, so $n_f = 2$ and $n_i = 3, 4, 5 \ldots$

Figure 8.9 shows the energy levels for hydrogen and the transitions corresponding to these two series of spectral lines.

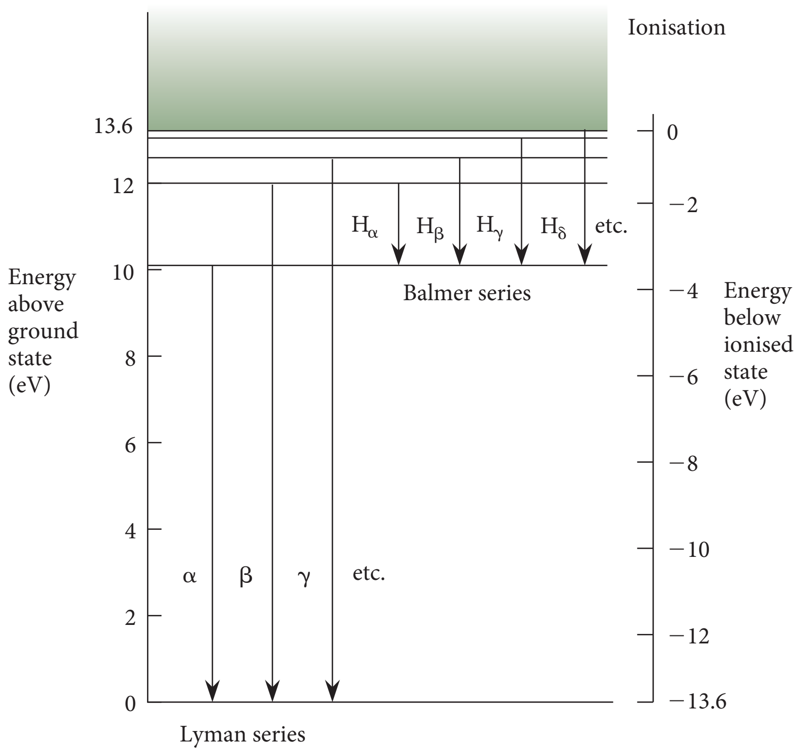


Figure 8.9 ▶ This energy level model for hydrogen shows the transitions corresponding to the Lyman and Balmer emission spectra series.

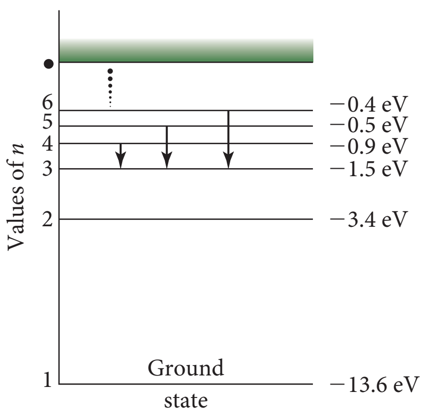
WORKED EXAMPLE 8.2

a Construct an energy level diagram like that shown in Figure 8.10 showing the transitions corresponding to the Paschen series, for which $n_f = 3$. (3 marks)

b Calculate the frequency of a photon released in a transition from $n_i = 6$ to $n_f = 3$. (4 marks)

Answers

a



◀**Figure 8.10**

Logic

Use the energies given in Figure 8.8 to draw an energy level diagram showing the energy levels up to $n = 6$. On this diagram the Paschen series corresponds to transitions to $n_f = 3$ from levels $n_i = 4$, 5 and 6.

3 marks

b $E_f - E_i = hf$ — Relate the change in energy of the electron to the photon frequency. — 1 mark

$f = \frac{E_i - E_f}{h}$ — Rearrange for frequency. — 1 mark

$f = \frac{-0.4\ \text{eV} - (-1.5\ \text{eV})}{4.14 \times 10^{-15}\ \text{eV s}}$ — Insert numbers with correct units, noting that we must use h in units of eV s or convert our energies to J. — 1 mark

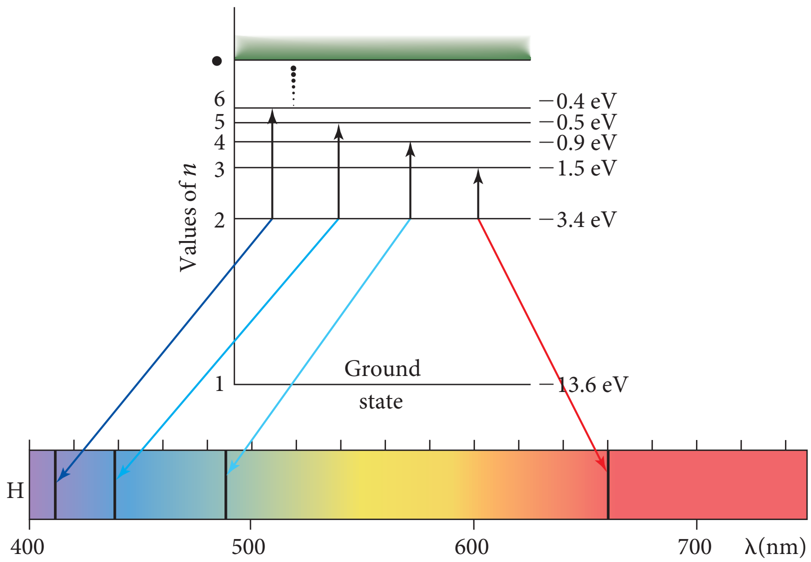
$f = 2.6 \times 10^{14}\ \text{Hz}$ — Calculate final value — 1 mark

Try these yourself

1 Calculate the frequency of the lowest energy line in the Lyman series. (4 marks)

2 Calculate the wavelength of the highest energy line in the Balmer series. (5 marks)

Absorption spectra occur when a photon of exactly the right energy is incident on an atom. A photon with too little or too much energy cannot be absorbed; only a photon with energy corresponding exactly to a gap between energy levels can be absorbed. In a gas of atoms most of the atoms will be in their ground state; that is, most will have their electrons in their lowest possible energy levels. Hence, the most likely transitions will be from the ground state, $n = 1$, to higher levels. However, as long as the temperature is above absolute zero, there will always be some atoms in an excited state, with electrons in levels $n = 2$, $n = 3$ and so on. That is why in an absorption spectrum we see lines corresponding to all the transitions that we see in the emission spectrum. The Balmer series in the absorption spectrum corresponds to electrons going *from* the $n = 2$ state to higher levels.



◀**Figure 8.11** The absorption spectrum of hydrogen and corresponding transitions are represented on this energy level diagram.

WORKED EXAMPLE 8.3

Consider photons of the following energies: 1.9 eV, 10.2 eV, 12.5 eV, 13.6 eV. Which of these could be absorbed by hydrogen gas? Explain your answer. (5 marks)

Answer	Logic	
$E_f - E_i$	For a photon to be absorbed its energy must be exactly equal to the energy gap between two energy levels.	1 mark
The 1.9 eV photon could be absorbed.	The 1.9 eV photon could be absorbed by an electron in the $n = 2$ state which then transitions to the $n = 3$ state: $E_3 - E_2 = -1.9$ eV	1 mark
The 10.2 eV photon could be absorbed.	The 10.2 eV photon could be absorbed by an electron in the $n = 1$ (ground) state, which then transitions to the $n = 2$ state: $E_3 - E_2 = -10.2$ eV	1 mark
The 12.5 eV photon could not be absorbed.	There is no energy gap corresponding to 12.5 eV, so this photon could not be absorbed, it would simply pass through the gas.	1 mark
The 13.6 eV photon could be absorbed.	The 13.6 eV photon could be absorbed by an electron in the ground state, and ionise the atom: $E_\infty - E_1 = -13.6$ eV	1 mark

The negative values mean that this much energy is absorbed by the system to undergo the transition.

Try this yourself

Which of these wavelengths could be absorbed? 365 nm, 420 nm, 656 nm (6 marks)

Spectra of other atomic species

Each type of atom produces a unique line spectrum. The energy levels are unique because each different type of atom has a different number of protons and hence a different nuclear charge. This means that the force exerted by the nucleus on the electrons is different, giving rise to different potential energies at different orbital radii, $r = \frac{nh}{2\pi m_e v}$. There is also a different number of electrons in different atoms. Electrons close to the nucleus 'shield' the outer electrons somewhat from the nuclear charge. This also acts to change the potential energy at the different allowed radii. These effects combine to give a unique fingerprint for each atom in the form of a line spectrum.

This shifting of energy levels is the origin of the conduction and valence bands in metals and semiconductors that you met in Chapter 5 of Nelson Physics Units 1 & 2 for the Australian Curriculum.

This is extremely useful, as it allows the presence of different types of atoms to be detected. For example, we know from the line spectrum of the Sun that there is a great deal of hydrogen and helium present as well as larger atoms. Other stars have different characteristic spectra, indicating the presence of other atomic species.

As we saw in the case study in Chapter 5, molecules also have characteristic spectra. The spectrum of a molecule is not simply the sum of the spectra of the atoms that make up the molecule. This is because when atoms bind together to form a molecule the energy levels change. Hence, it is possible to distinguish between ethanol and methanol by their spectra, even though both contain only carbon, hydrogen and oxygen.

EXPERIMENT 8.1

MEASURING PLANCK'S CONSTANT

Light-emitting diodes (LEDs) produce photons of a particular wavelength when electrons transition from a higher to a lower energy level within the semiconductor material of the LED. It is possible to measure Planck's constant using the wavelengths of LEDs. The minimum voltage that will cause an LED to emit light is called the threshold voltage, V_t:

$$V_t = \frac{hc}{e\lambda}$$

where h is Planck's constant, c is the speed of light, e is the charge of the electron and λ is the wavelength of the light produced by the LED.

Aim

To measure Planck's constant

Materials

- 5 different coloured LEDs
- DC power supply
- 1 kΩ resistor
- 10 kΩ variable resistor
- voltmeter or multimeter

What are the risks in doing this experiment?	How can you manage these risks to stay safe?
High voltages can damage equipment or cause electric shock.	Only use low voltages. Ask your teacher to check your circuit before you turn the power supply on.
Large currents will damage the LEDs.	Always have a resistor in the circuit with the LED and power supply.

In your write-up, add any more risks you can think of, as well as ways to manage them.

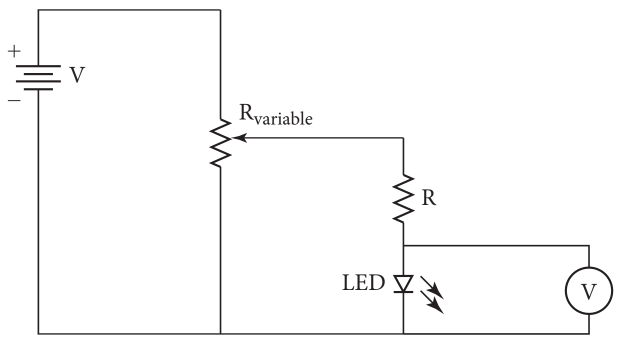
Procedure

1 Connect the circuit as shown in Figure 8.12.

When your circuit is set up and has been checked you will need to work in a dark room to make your measurements.

2 Slowly increase the voltage until the LED just starts to glow.

3 Record the voltage at which this occurs.



▲Figure 8.12
Circuit diagram for measuring Planck's constant

Results

1 Record your results in a table like that shown below.

Colour of LED	Wavelength, λ (m)	$\frac{1}{\lambda}$ (m^{-1})	Threshold voltage, V_t (V)

2 Draw a graph of V_t against $\frac{1}{\lambda}$.

Analysis of results

1 Find the gradient of your graph. Remember to include units.

2 From the gradient you can now calculate Planck's constant.

Discussion

1 Refer back to Chapter 3, in particular the definition of electric potential. Explain how the equation $V_t = \frac{hc}{e\lambda}$ arises.

2 Why do you plot V_t vs $\frac{1}{\lambda}$ rather than V_t vs λ? What would be the shape of a V_t vs λ graph?

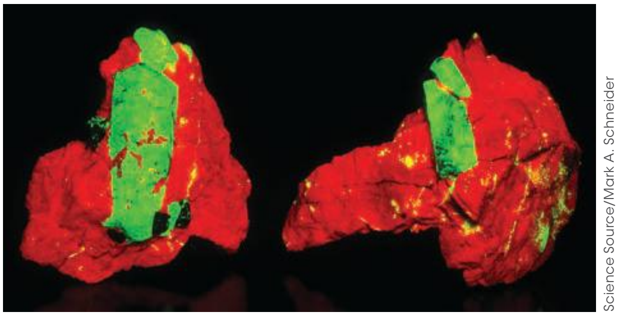
3 Does your value for Planck's constant agree with the accepted value? To be able to compare the values you will need to estimate the uncertainty in your value. How will you do this?

4 What could you do to minimise uncertainties in your experiment?

WOW

Fluorescence

Fluorescence is the glowing of substances when illuminated by ultraviolet light. An atom is excited to an energy state several levels above its ground state by absorption of a high-energy ultraviolet photon. The atom can decay (de-excite) in a number of ways. One is through a series of smaller steps back down to the ground state. Lower-energy photons are emitted, some of which may be in the visible range. Fluorescent dyes are used in paints and even laundry detergents to make colours brighter in sunlight by converting some of the ultraviolet into visible light – making your whites whiter!



Science Source/Mark A. Schneider

Figure 8.13 ▲
These rocks (willemite on calcite) fluoresce in the visible spectrum when illuminated by an ultraviolet light.

Limitations of the Bohr model

Bohr was able to explain qualitatively and quantitatively the existence and positions of the spectral lines of a hydrogen atom. His estimate of the size of the largest stable radius also agreed closely with the measured size of the hydrogen atom.

However, the Bohr model could not predict the spectra of multi-electron atoms, even one as simple as two-electron helium. It also could not explain the different intensities of lines or why some lines split into multiple, closely spaced lines – fine and hyperfine structure – or the magnetically induced Zeeman effect.

WOW

Zeeman effect

In 1896, Dutch physicist Pieter Zeeman was using laboratory equipment to observe the effect of strong magnetic fields on spectral lines. He observed that the individual lines were split into multiple closely spaced lines. Unfortunately, he was supposed to be working on something else, and had been explicitly told by his supervisor not to use the equipment for his own research. When his supervisor found out he was fired.

Zeeman and Hendrik Lorentz later showed that the splitting was due to the electron's intrinsic magnetic field (spin) and the magnetic field due to its orbital motion interacting with the applied magnetic fields. Spin is described further in the next chapter. For this they were awarded the 1902 Nobel Prize in Physics.

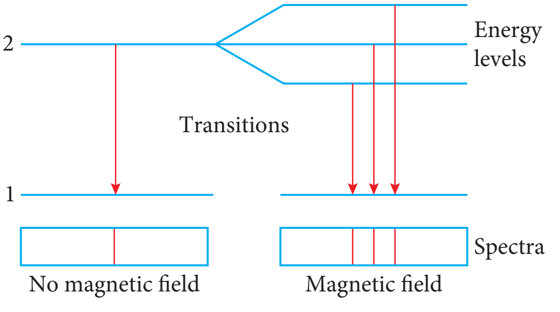


Figure 8.14 ▲
The Zeeman effect. Spectral lines are split by the presence of the magnetic field.

Finally, Bohr's model introduced the idea of quantised atomic energy levels, but it did not offer any explanation for why they should be quantised. Successful models have both predictive power and explanatory power. Bohr's model lacked explanatory power, and had limited predictive power. It was superseded by a more fully quantum mechanical model developed by Schrödinger and Heisenberg. This modern quantum model built on the ideas of de Broglie, described in the next section.

QUESTION SET 8.1

Remembering

1 Write down Bohr's postulates. What was quantised in Bohr's model of the atom?

2 What are the two purposes of a model?

Understanding

3 Explain the difference between an absorption spectrum and an emission spectrum. Why are the lines at the same frequencies?

4 Explain why Bohr's model of the atom is considered the first quantum mechanical model.

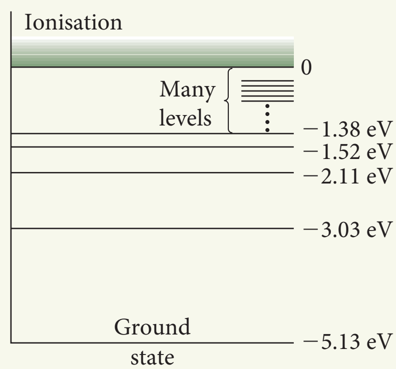
Applying

5 How many different emission spectral lines would there be as a result of transitions from the −1.52 eV energy level in a sodium gas? The sodium energy levels are shown in Figure 8.15.

6 What is the frequency and wavelength of the photons emitted when an electron in a hydrogen atom moves from the second energy level above the ground state to the first energy level above the ground state?

7 What is the shortest wavelength photon that can be emitted from an excited hydrogen atom when its electron returns to the ground state?

8 A hydrogen atom is in the $n = 4$ energy state. It then returns to the ground state. What are all the possible energies the emitted photon(s) could have?



▲**Figure 8.15**
Energy level diagram for sodium

Analysing

9 Calculate the velocity of an electron orbiting the nucleus at a radius of $r = 5.3 \times 10^{-11}$ m, with $n = 1$, according to the Bohr model.

Reflecting

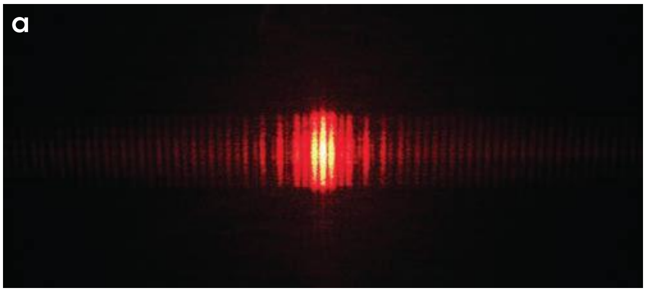
10 What have you learnt about the development of scientific explanations? Give specific examples from Rutherford and Bohr.

Waves and particles

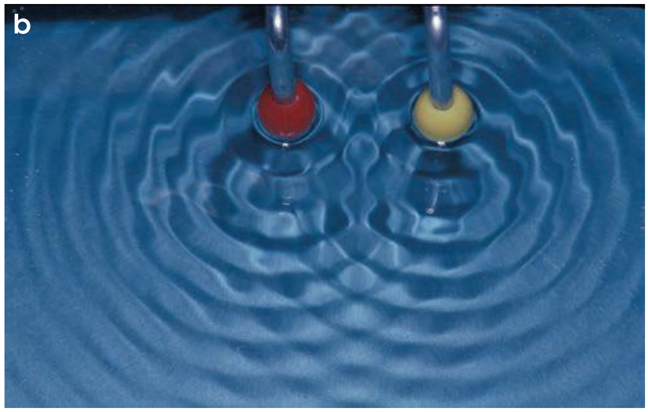
We have seen that the photoelectric effect demonstrates that light acts like a particle. However, the particle model of light, which is quantum mechanical in nature, does not explain the behaviour of light in some other experiments. We have seen that if light is shone through two slits, as in a Young's double-slit experiment, an interference pattern is seen. This is shown in Figure 8.16(a). In this experiment, light clearly acts like a wave, and produces a pattern of high and low intensity just like any other wave, such as that shown in Figure 8.16(b).

So is light a particle or a wave? The answer is that it acts like both, and the behaviour that you see depends on the experiment that you do. On its own, neither the wave nor the particle model completely explains the behaviour of light. They are *complementary* models; both are needed, and which one is used depends on the situation. We describe this need for both of the two distinct models as **wave–particle duality**.

You studied wave behaviour including interference of light in Unit 2 of Nelson Physics Units 1 & 2 for the Australian Curriculum.



Getty Images/GIPhotoStock



Alamy/Bruce Coleman Inc.

Figure 8.16 ▲
An interference pattern from a) a double-slit experiment using a laser and b) two wave sources in water

LOUIS DE BROGLIE

You can read more about Louis de Broglie at the Nobel Prize website.

The question then arises: What about things which we know to be particles? If light is a wave and a particle, what about an electron? Or a proton? Could they also have a dual wave–particle nature? And if so, what experiments can we do to observe this dual nature?

De Broglie wavelength for particles

In 1924 Louis de Broglie introduced the idea that any moving particle has an associated wavelength. He introduced this hypothesis in his doctoral thesis. The idea was revolutionary in physics, and the implications enormously important. However many physicists were sceptical.

De Broglie claimed that a particle of mass, m, moving at a velocity v, would have an associated wavelength:

$$\lambda = \frac{h}{mv}$$

or

$$\lambda = \frac{h}{p}$$

where h is Planck's constant and $p = mv$ is the momentum of the particle.

The wavelength λ is now known as the de Broglie wavelength.

WORKED EXAMPLE 8.4

a What is the de Broglie wavelength of an electron travelling at $25\,\text{m}\,\text{s}^{-1}$? (4 marks)

b How fast would a cricket ball with a mass of 160 g have to travel to have the same de Broglie wavelength as the electron in part a? (4 marks)

Answers	Logic	
a $\lambda = \frac{h}{mv}$	Relate wavelength to other values.	1 mark
$m = 9.1 \times 10^{-31}\,\text{kg}$	Find the data needed.	1 mark
$\lambda = \frac{6.6 \times 10^{-34}\,\text{J s}}{(9.1 \times 10^{-31}\,\text{kg})(25\,\text{m s}^{-1})}$	Substitute numbers including units.	1 mark
$\lambda = 2.9 \times 10^{-5}\,\text{m}$	Calculate the final value.	1 mark

b $\lambda = \frac{h}{mv}$	Relate wavelength to other values.	1 mark
$v = \frac{h}{m\lambda}$	Rearrange for v.	1 mark
$v = \frac{6.63 \times 10^{-34}\ \text{J s}}{(0.16\ \text{kg})(2.9 \times 10^{-5}\ \text{m})}$	Substitute numbers including units.	1 mark
$v = 1.4 \times 10^{-28}\ \text{m s}^{-1}$	Calculate the final value.	1 mark

Try this yourself

Calculate the de Broglie wavelength of a cricket ball moving at 25 m s^{-1}. (4 marks)

De Broglie's equation also allows us to calculate the momentum of a photon:

$$p = \frac{h}{\lambda}$$

In any interaction between objects, including collisions, both energy and momentum must be conserved. As we have seen, photons carry energy that can be transferred to an electron in the photoelectric effect. The ejected electron also gains momentum from the photon, and this momentum must be conserved in the interaction.

We shall see in the next chapter that photon momentum is important in reactions involving elementary particles.

You saw in Chapter 6 that the momentum of a photon is $p = \frac{E}{c}$. As $E = hf = \frac{hc}{\lambda}$, we can also write the momentum as $p = \frac{h}{\lambda}$.

De Broglie waves and the Bohr model

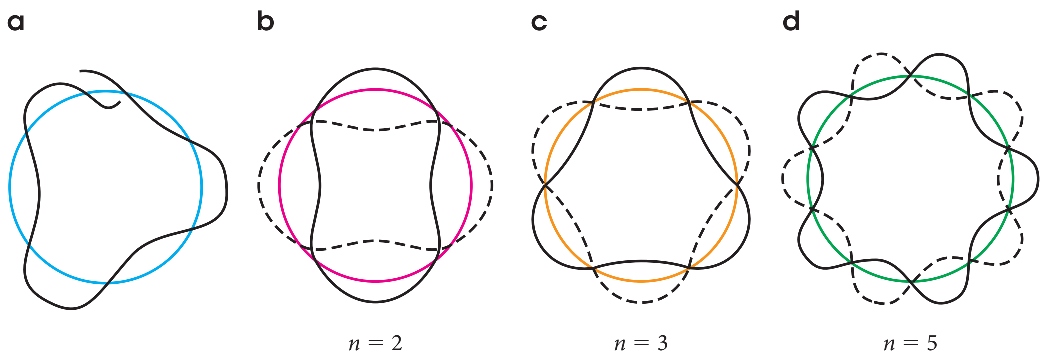
When the idea of the de Broglie wavelength was incorporated into the Bohr model it gave a justification for the quantisation of energies. The explanation treats the electrons as standing waves. If the electrons are standing waves then an integer number of wavelengths must fit into the orbit. The electrons act like standing waves on a string the length of the orbit.

The condition for a stable orbit is then:

$$n\lambda = 2\pi r$$

where λ is the de Broglie wavelength of the electron, n is an integer and r is the radius of the orbit.

You studied standing waves in Chapter 10 of Nelson Physics Units 1 & 2 for the Australian Curriculum.



◀Figure 8.17 Electrons as standing waves, showing how energy levels correspond to different standing wave modes. (a) is not a standing wave and hence not a stable orbit. (b), (c) and (d) are stable orbits.

WORKED EXAMPLE 8.5

Calculate the longest three wavelengths of an electron in an orbit of radius 5.3 nm. (7 marks)

Answer	Logic	
$n\lambda = 2\pi r$	Relate wavelength to other values.	1 mark
$\lambda = \frac{2\pi r}{n}$	Rearrange for wavelength.	1 mark
$\lambda_1 = \frac{2\pi r}{1}, \quad \lambda_2 = \frac{2\pi r}{2}, \quad \lambda_3 = \frac{2\pi r}{3}$	Write the expression for the first three wavelengths.	2 marks
$\lambda_1 = \frac{2\pi(5.3 \times 10^{-9}\text{ m})}{1}$	Substitute numbers and calculate value.	1 mark
$= 3.3 \times 10^{-8}\text{ m}$		
$\lambda_2 = \frac{2\pi(5.3 \times 10^{-9}\text{ m})}{2}$		
$= 1.7 \times 10^{-8}\text{ m}$	Calculate second value.	1 mark
$\lambda_3 = \frac{2\pi(5.3 \times 10^{-9}\text{ m})}{3}$		
$= 1.1 \times 10^{-8}\text{ m}$	Calculate third value.	1 mark

DE BROGLIE WAVELENGTH CALCULATOR

The Hyperphysics website has a de Broglie wavelength calculator, with links to much more about wave–particle duality.

Although the idea of electrons as waves gave some physical explanation for the quantisation of energy levels in the Bohr model, it did nothing to address the other failings of the model. As the wave nature of electrons became better understood, better models of the atom were developed. However, the Bohr model was of great historical importance, and can still be used to give a useful model of simple hydrogen-like systems.

The wave nature of matter

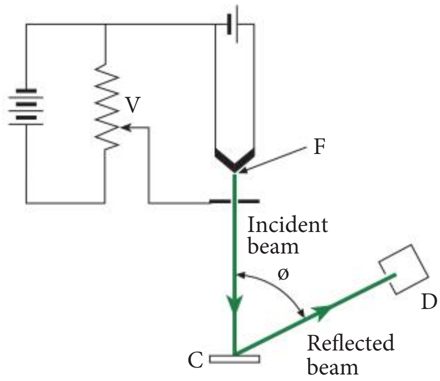
What does it mean for a particle or object to have a wavelength? It does *not* mean that it follows a wiggly path, undulating up and down as it travels. What it means is that in some sense the particle is de-localised in space, just as a wave is.

Particles do not undulate as they move because of their wave nature. Their wave nature means that they are delocalised or spread out in space as they travel.

Think about the way waves behave when they interact with things of a size similar to their wavelength. Waves passing through a hole will diffract and spread out on the other side. Waves passing through slits spaced on a similar scale to their wavelength will produce an interference pattern (Figure 8.16). De Broglie predicted that when particles interacted with objects or apertures, such as slits, on a scale similar to their de Broglie wavelength, they too would show wave behaviour.

De Broglie's theory was based on the work of Einstein and Planck. They had introduced the idea of a particle or quantum aspect for waves. Their model gained acceptance because it explained experimental results such as the photoelectric effect and the black body radiation spectrum. At the time that de Broglie proposed his idea, there was no experimental evidence for it, and no experiments waiting to be explained. Hence it did not initially gain a lot of support.

Electron interference

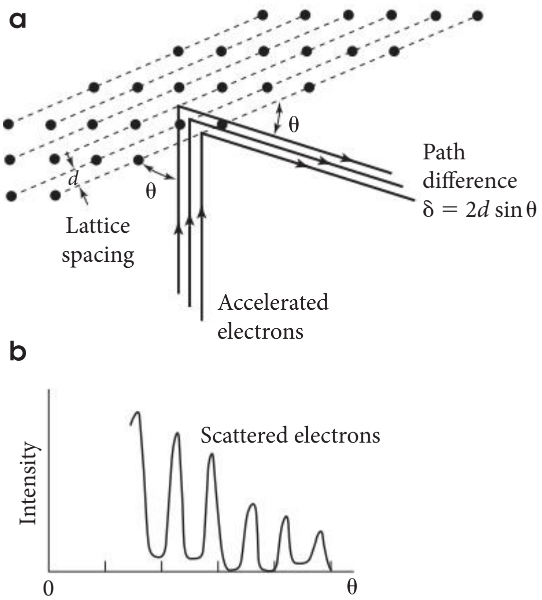


▲Figure 8.18
Davisson–Germer experiment. The filament, F, is heated and the electrons released are accelerated across a potential difference. The beam reflected from the nickel target showed a diffraction pattern of maxima and minima, which was detected at D as a function of scattering angle, ϕ.

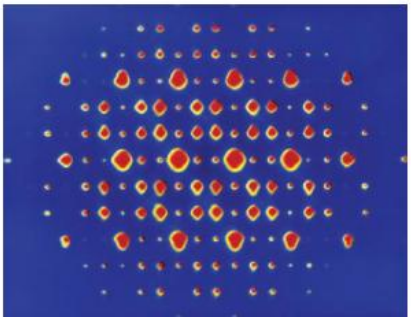
In 1927 Clinton Davisson and Lester Germer were doing an experiment to study the surface of a piece of nickel. Their sample of nickel was mounted inside a vacuum chamber. Electrons were fired at the surface and the reflected electrons were detected using a movable detector. During the experiment a leak developed in the vacuum system and air got into the chamber. The air reacted with the surface of the nickel, forming a layer of oxide. To remove this layer, they heated the piece of nickel in a furnace with hydrogen gas. They then put it back in the vacuum chamber and repeated the experiment. To their surprise, they no longer detected electrons at all positions, but only at specific, regularly spaced positions.

What they observed was **electron diffraction**. The electrons reflected from the metal surface had formed an interference pattern. They measured the spacing between maxima, and found that it corresponded to that predicted by the de Broglie wavelength. The layers of atoms in the nickel were acting as a diffraction grating for the 'electron waves'.

A diffraction grating acts like the slits in a double-slit experiment, except with thousands of slits. They produce an interference pattern with narrow lines, separated due to the path difference for light travelling through adjacent slits. However, the crystal has layers of atoms in different planes, so the resulting pattern is a two-dimensional pattern of bright spots, such as that shown in Figure 8.20.



▲Figure 8.19
a) Electrons reflecting from different layers of atoms in the nickel crystal interfere. b) This results in a diffraction pattern.



Science Photo Library/Dr David Wexler, Coloured by Dr Jeremy Burgess

▲Figure 8.20
An electron diffraction pattern

As shown in Figure 8.19(a), the path difference for electrons reflecting from different layers of atoms in the metal crystal lattice is given by

$$\delta = 2d\sin\theta$$

where d is the spacing between layers of atoms and θ is the angle between the incident beam of electrons and the parallel layers.

This looks a lot like the expression for the path difference in the double-slit experiment. There are two differences, however. First, the angle is measured to the surface rather than the normal. Second, there is a factor of two because one electron wave travels the extra distance between layers *twice* compared to the other electron wave.

Just as with the double-slit experiment, and indeed any interference experiment, maxima are observed where the path difference is equal to an integer number of wavelengths.

In electron diffraction:

$\delta = 2d\sin\theta = m\lambda$, $m = 1, 2, 3 \ldots$ for constructive interference.

$\delta = 2d\sin\theta = \left(m + \frac{1}{2}\right)\lambda$, $m = 1, 2, 3 \ldots$ for destructive interference.

This observation by Davisson and Germer was only possible because of the accident with the vacuum system. The heating of the sample to remove the layer of oxide had melted their sample, which originally was made of many microscopic crystals. When the sample cooled, it formed a single crystal with regularly spaced atoms. The spacing between atoms in nickel was already known at the time, from measurements using X-rays. Davisson and Germer were able to work out the momentum of their electrons from the potential difference used to accelerate them (see Chapter 3). The result was the first experimental evidence for wave–particle duality in particles with mass. De Broglie's hypothesis was supported, and only two years later, in 1929, he was awarded the Nobel Prize.

WORKED EXAMPLE 8.6

Davisson and Germer used electrons which had been accelerated through a potential difference of 54 V. They found an interference maximum at an angle of $\theta = 65°$. Assume $m = 1$ for this interference maximum.

a What was the de Broglie wavelength of the electrons? (6 marks)

b What was the spacing between layers of atoms that produced this peak in intensity? (4 marks)

Answers	Logic	
a $E = Ve$	Relate the energy to the accelerating potential.	1 mark
$E = \frac{1}{2}mv^2$, so:		
$v = \sqrt{\frac{2E}{m}} = \sqrt{\frac{2Ve}{m}}$	Relate velocity to potential.	1 mark
$p = mv = \sqrt{2mVe}$	Relate momentum to potential.	1 mark
$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2mVe}}$	Relate wavelength to momentum.	1 mark
$\lambda = \frac{6.626 \times 10^{-34}\ \text{Js}}{\sqrt{2 \times 9.1 \times 10^{-31}\ \text{kg} \times 54\ \text{V} \times 1.6 \times 10^{-19}\ \text{C}}}$	Substitute values with correct units.	1 mark
$\lambda = 1.67 \times 10^{-10}\ \text{m}$	Calculate the final value.	1 mark

Note that we have kept three significant figures here as we will use this value in part b.

$\lambda = 1.7 \times 10^{-10}\ \text{m}$ is the correctly rounded answer to part a.

b $2d\sin\theta = \lambda$ — Relate path difference to de Broglie wavelength for constructive interference, with $m = 1$. 1 mark

$$d = \frac{\lambda}{2\sin\theta}$$ Rearrange for d. 1 mark

$$d = \frac{1.67\times10^{-10}\text{ m}}{2\sin 65°}$$ Substitute values with correct units.

$$d = 9.2\times10^{-11}\text{ m}$$ Calculate the final value.

Try these yourself

1 Find the angle at which the m = 2 interference maximum occurs for this system. Is it possible to observe this maximum? (4 marks)

2 Find the angular positions of the first and second interference minima. (4 marks)

It is not only electrons that show interference patterns. Atoms including helium, sodium and even large molecules such as carbon-60 'buckyballs' have been used to produce interference patterns. However, the more mass an object has, the smaller its de Broglie wavelength, and the harder it is to observe its wave behaviour.

The wave–particle theory originally put forward by de Broglie is now the basis of several important technologies. Electron microscopes use the wave nature of electrons to create images of objects too small to be visualised using light. Electron microscopy has been enormously important in medicine and the biological sciences, allowing organisms such as bacteria and viruses to be 'seen' and studied. Neutron diffraction is carried out at the OPAL reactor in Sydney. The wave nature of the neutron is used to produce diffraction patterns. These are used to investigate the properties of materials and test for stress damage in machine parts, among other applications.

WAVE–PARTICLE DUALITY SEEN IN CARBON-60 MOLECULES

This experiment showed that even the large carbon molecules called 'buckyballs' can produce interference.

WOW

Electron microscopes

In a scanning electron microscope (SEM) the electrons interact with atoms at the surface of the sample. Some electrons are reflected and some electrons are ejected from atoms at the surface. These are used to create images such as in Figure 8.21(b). SEMs can be used to 'see' objects as small as 1 nm.

In a transmission electron microscope (TEM) the electrons are passed through a very thin sample and collected by a detector on the other side. The image can be formed by the sample simply absorbing some electrons and allowing others to pass through, like a shadow. The electrons can also produce a diffraction pattern on the other side. TEMs can be used to 'see' objects as small as 0.1 nm.

▲Figure 8.21
a) An electron microscope. b) A photo taken with an electron microscope of a mite that is less than 1 mm long.

Single particle interference experiments

We have still not answered the question: what is it that is 'waving'? *How* do electrons, which show particle behaviour, produce an interference pattern? What exactly is distributed over space such that interference can take place?

Experiments using electrons and double slits also showed interference patterns, with a pattern of maxima and minima predicted by the de Broglie wavelength. Initially, it seemed that perhaps electrons passing through each slit were interfering with each other. Lots of experiments were done to try to determine what exactly was passing through the slits, and lots of competing theories were developed.

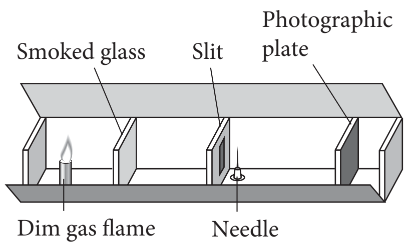


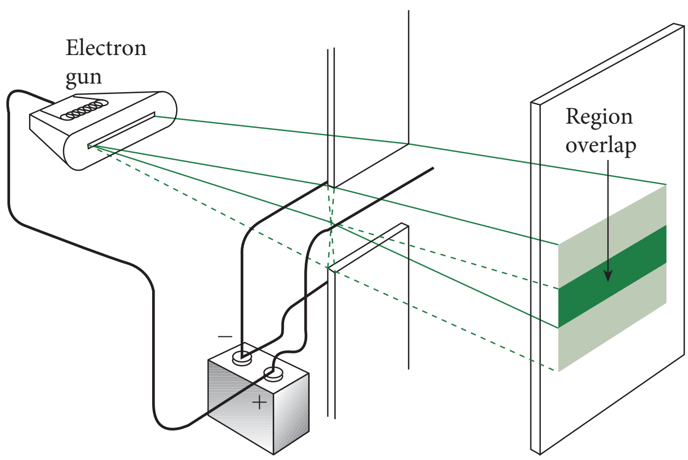
Figure 8.22 ▲ Taylor's experiment (1909). There was never more than one photon in the box at any one time, yet an interference pattern developed on the photographic plate.

One experimental result that was troubling to physicists was the single photon interference demonstrated by G.I. Taylor in 1909. The apparatus is shown in Figure 8.22.

Taylor demonstrated that an interference pattern occurred when the light was so dim that only a single photon could be present in the apparatus at a time. This wasn't a problem if light was a wave. Regardless of the intensity, which is related to the amplitude, the wave still spreads out in space and hence passes through both slits and recombines on the other side. But how could this result be explained by the particle theory of light?

Physicists wanted to see if the same thing happened with other particles. In 1974, Pier Giorgio Merli, Giulio Pozzi and GianFranco Missiroli performed a double-slit experiment with electrons. In their experiment, as in Taylor's, only a single particle was in the apparatus and incident on the slits at a time. Again, interference was observed. The most famous single-electron interference experiment was performed by Akira Tonomura and colleagues in Japan in 1989. Their experimental set-up is shown in Figure 8.23.

Figure 8.23 ▶ Tonomura's electron biprism experiment. Electrons diffract around the charged central wire and recombine to produce an interference pattern.



The electron double-slit experiment can be analysed in exactly the same way as Young's double-slit experiment for light, as described in Chapter 7.

Their results clearly showed the build-up of an interference pattern, one electron at a time, as shown in Figure 8.24.

This experiment can be analysed in the same way as the double-slit experiment for light in Chapter 7. The electron biprism acts as a double slit for the incident electrons. Interference maxima are observed at angles given by

$$d\sin\theta = m\lambda,\ m = 0, 1, 2, \ldots$$

where d is the effective slit spacing due to the biprism and λ is the de Broglie wavelength of the electron. The angle θ is the angle between the normal line joining the biprism and the detector screen and the line from the biprism to the point of interest on the screen (see Figure 7. 2).

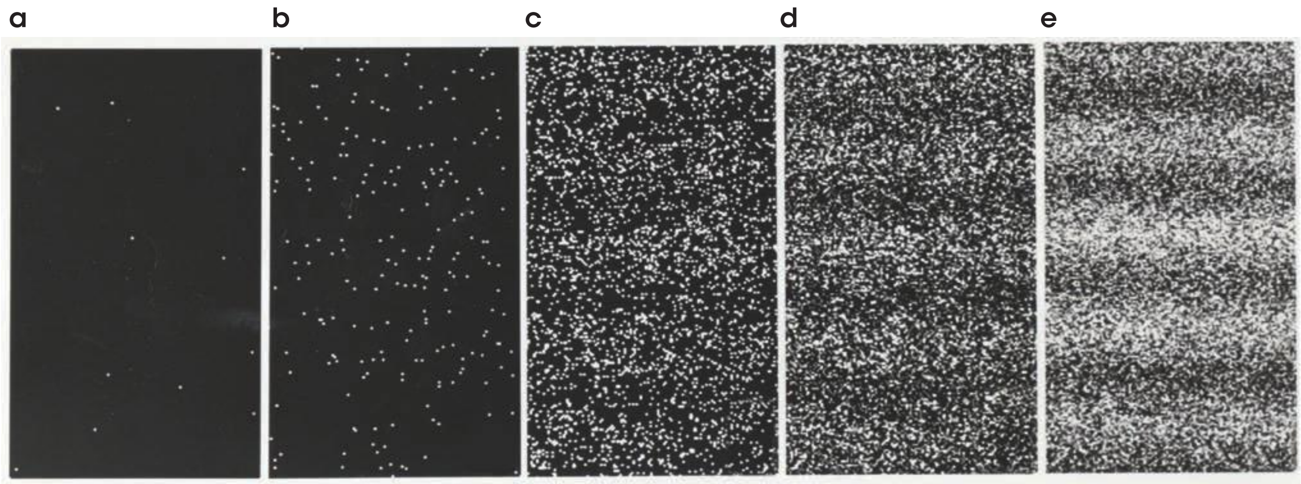
The results were now clear and indisputable. Single electrons can produce interference patterns. So it cannot be that the electrons passing through one slit interfere with electrons passing through the other.

Each electron in some way interacts with and passes through both slits simultaneously.

It is hard to imagine how a particle could do this. We need to use the wave model here. In the modern interpretation of quantum mechanics we think of the particle as being built up of waves that can pass through both slits at the same time. What exactly these waves represent has been a matter of dispute among physicists ever since. The most accepted theory currently is that the waves represent the *probability* of the electron passing through each slit. The theory states that the probability of the particle passing through each slit is represented by a wave. It is these probability waves that interfere, creating the pattern seen in Figure 8.24. The pattern appears even though the electrons themselves are detected as particles at single distinct points.

DOUBLE-SLIT EXPERIMENT

Read more about this experiment and watch a video of the pattern forming.



Wikipedia Creative Commons/ Provided with kind permission of Dr Tonomura

▲**Figure 8.24**
Interference pattern produced by Tonomura's single-electron interference experiment. a) After 10 electrons and b) after 100 electrons the pattern looks random, but c) after 3000 electrons have arrived, the interference pattern is starting to emerge. In d) and e) the interference pattern has become even clearer

So what happens if we know which slit the electron passes through? In this case the probability of the wave passing through one slit is zero, and through the other slit it is one. Surely if this interpretation of the wavelength is correct, the pattern should disappear. When the experiment was done, the pattern *did* disappear. There was argument over the validity of the experiment, over whether the detector prevented the interference, and so on. But many versions of the experiment have been performed, with the same results.

TONOMURA EXPERIMENT

The Tonomura experiment was voted by physicists to be the most beautiful experiment ever performed.

Modern quantum mechanics – probability and uncertainty

From the 1920s, what is now known as modern quantum mechanics was developed by physicists including Niels Bohr, Louis de Broglie, Werner Heisenberg and Erwin Schrödinger. The term 'modern' is used to distinguish it from earlier quantum mechanics, such as the Bohr model.

The new model that they came up with included the idea of uncertainty. In any experiment, as you know, there will be uncertainties due to various sources including equipment limitations. Heisenberg proposed that there is also an intrinsic uncertainty and that the behaviour of particles is **probabilistic**. In other words, it cannot be predicted with certainty no matter how much you know about the particles. In classical mechanics the behaviour of all objects including subatomic particles is **deterministic**. This means that it is completely predictable once you have enough information. This is a fundamental difference between quantum mechanics and classical mechanics.

QUANTUM TUNNELLING

Quantum tunnelling is another quantum phenomenon that is unexplainable by classical physics. Particles 'tunnel' through barriers they do not have enough energy to 'jump'.

TUNNELLING

Tunnelling is the first vital step in the thermonuclear reaction that powers the Sun.

Classical mechanics is deterministic. Given enough information, the outcome of any experiment is predictable. Modern quantum mechanics is probabilistic. It is not possible to predict the outcome of an experiment with complete certainty.

In the double-slit experiment, a probability wave associated with each particle passes through both slits. These probability waves interfere on the other side. When we measure which slit the wave passes through, we no longer have the probability wave passing through both slits, so we no longer have an interference pattern. The modern probabilistic model of quantum mechanics correctly predicts the results of these types of experiments.

Although it is now the orthodox view, there are still some physicists who do not accept this model. Einstein, for example, was vigorously opposed to this interpretation of quantum mechanics, and made his famous statement 'God does not play dice' in response. To the end of his days Einstein argued against the theory, but it has now become accepted by most physicists. You can read about the arguments between Einstein, Bohr and others in the Scientific literacy box 'God does not play dice' – Einstein's objections to quantum mechanics.

The probability of an electron being at a given position is described by a wave equation. This wave equation is the Schrödinger wave equation.

The modern quantum mechanical model of the atom

The modern quantum mechanical model of the atom treats the electrons as having a *probability* of being at any given position near the nucleus. Thus, electrons are not confined to circular orbits as in the Bohr model.

The principle quantum number, *n,* is still used, but other quantum numbers have been introduced. These other quantum numbers describe the quantisation of other physical properties such as the magnetic fields due to spin and orbital motion (see Chapters 9 and 10). In addition, the Pauli **exclusion principle** states that no two electrons in an atom may have identical sets of quantum numbers. Hence, each electron in an atom has a unique set of quantum numbers describing it.

In chemistry, n is called the shell number. The different possible combinations of n and the other quantum numbers determine the orbital structure of an atom and the filling of the shells by electrons. This structure is the physical basis of the periodic table.

The different possible combinations of quantum numbers determine the energy level structure of the atom. This model of the atom closely predicts the observed spectra of atoms.

The idea of orbits is used in this model to represent regions of high probability of finding the electrons. Figure 8.25 shows some examples of these 'orbits'. The shapes shown represent the region in which there is a high probability of finding the electron for a given value of *n*, the principal quantum number. Note that unlike the Bohr model, the 'orbits' are not generally circular or spherical.

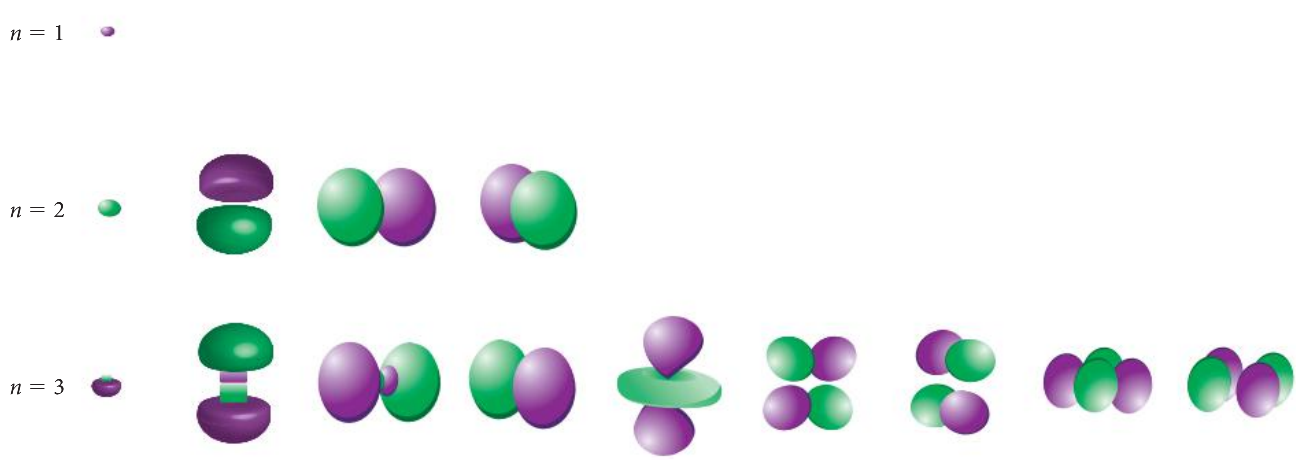


Figure 8.25 ▲ Electron orbits for $n = 1$, $n = 2$ and $n = 3$ shells

The development of the modern quantum mechanical model of the atom gave a physical and theoretical basis for many of the empirical rules in chemistry. We now have a better understanding of why atoms form bonds the way they do, and why materials such as conductors and semiconductors have the properties that they have.

Scientific literacy: 'God does not play dice' – Einstein's objections to quantum mechanics

Einstein disagreed with what has now become the orthodox interpretation of quantum mechanics. This interpretation says that nature is fundamentally probabilistic. Faced with two slits, an electron may go through one, or it may go through the other (or both). There is no way to predict which it will pass through, no matter how much you know about the system. This is in contrast to classical mechanics which is deterministic. In classical mechanics, if you have enough information about a particle and the system containing it, you can predict with certainty what will happen.

Einstein was not alone in objecting to the probabilistic or uncertain nature of quantum mechanics. Erwin Schrödinger developed the wave equation central to quantum mechanics. The modern quantum model of the atom is based on this equation. This model is used to understand and predict chemical bonding. In spite of its great success, Schrödinger said of quantum mechanics: 'I don't like it and I'm sorry I ever had anything to do with it.'

Einstein and colleagues proposed the theory of hidden variables in response to the probabilistic theory. In this theory, we cannot predict what path a particle takes because quantum theory is incomplete and there are variables yet to be discovered, hence the name, 'hidden variables'. Einstein believed that once these are discovered, we will be able to develop a deterministic quantum mechanical model.

There were great arguments between physicists, particularly between Einstein and Bohr. They disagreed vehemently, and neither ever conceded defeat. Bohr and his colleagues such as Heisenberg favoured the probabilistic interpretation. Uncertainty is a fundamental part of this interpretation, not just a result of the limits of measuring equipment. Einstein claimed that 'God does not play dice', and made many similar statements, to which Bohr eventually retorted 'Einstein, stop telling God what to do'.

Einstein and his colleagues, Podolsky and Rosen, proposed a thought experiment in which the waves of two particles are entangled or mixed up. The particles then move off in opposite directions. If you interact with one particle, it instantly has an effect on the other particle because they are connected by their probability waves. In theory, this should allow information to travel faster than the speed of light. This contradicts special relativity, and Einstein called it 'spooky action at a distance'. This thought experiment was designed to show that quantum mechanics could not be complete, because the predicted result was impossible.

In the end, the probabilistic interpretation favoured by Bohr is the one accepted by most (but not all) physicists. Einstein's 'spooky action at a distance' has been observed experimentally, and is known as 'entanglement'. There are many groups of physicists around the world now trying to develop quantum computing and quantum cryptography based on it.

Quantum computers will store and process information in a different way, allowing much faster processing and increased information storage. Quantum cryptography would revolutionise internet security, and there is much military interest in its development.

The development of quantum mechanics was not just a revolution in physics. Our way of life has changed due to the technology that has grown from it. The way we see the universe has also changed – it is no longer a deterministic 'clockwork universe', but a spooky, uncertain place. Of course, that depends on who you agree with!

Corbis/Bettmann

Alamy/GL Archive

▲**Figure 8.26**
a) Albert Einstein and b) Neils Bohr

Questions

1 Define 'deterministic'.
2 The thought experiment proposed by Einstein, Podolsky and Rosen is often referred to as the EPR paradox. Why do you think it is called a paradox?
3 Why do you think financial institutions such as banks might be interested in quantum cryptography?
4 Read some more about the hidden variables and probabilistic interpretations of quantum mechanics. Who do you agree with and why? Is your choice based on experimental evidence, or on some other basis?
5 Why do you think Einstein was so opposed to the probabilistic interpretation of quantum mechanics? Read more about Einstein's objections to this interpretation and critique his argument.

Quantum physics today

The field of quantum physics is still very active, and there is much to be done. There are still competing models and different interpretations of what the models mean. We shall see in the next chapter how new models such as the quark theory have changed the way we understand what particles such as protons and neutrons really are.

These models will be further developed, some existing models and interpretations will be abandoned, and new ones will arise. New experiments will help to distinguish between competing models and stimulate the development of new theories. The new theories will inspire further experiments to test them, and thus science will progress. Alongside this, new technologies will develop that will change the way we work and communicate, just as the development of quantum mechanics has already done.

Some of it may sound like science fiction, but quantum physics is the basis of many of the technologies you use every single day. Semiconductors used in diodes and transistors rely on electrons behaving like waves, which allows them to move in ways that a classical particle cannot. Without the development of quantum physics, you would not have a smart phone, an MP4 player, or even a computer.

Quantum physics has changed the way we think about the world, and the way in which we live more than any other advance in science. That is why it was such an important revolution in science.

QUESTION SET 8.2

Remembering

1 Summarise the probabilistic interpretation of quantum mechanics. Name one physicist who agreed with this interpretation, and one who disagreed.

2 Describe two experiments that demonstrate the wave nature of electrons.

Understanding

3 A photon, an electron and a neutron all have the same wavelength. Rank the velocities of the three particles.

4 Explain why the Davisson–Germer experiment was so important.

Applying

5 A bullet of mass 50 g travels at 1.2×10^3 m s^{-1}.

- a What is its de Broglie wavelength?
- b Why is the wave nature of the bullet not observed through diffraction effects?

6 A beam of electrons with de Broglie wavelength 1.5×10^{-10} m is incident on an electron biprism, such as that shown in Figure 8.23. The biprism can be modelled as a double slit with separation 4.5×10^{-9} m. If the detectors are arranged on a screen a distance 20 cm from the biprism, what is the location of:

- a the first interference maximum?
- b the first interference minimum?
- c the second interference maximum?

Analysing

7 A photon of energy 3.6×10^{-15} J collides with a stationary electron that is free to move.

- a What is the magnitude of the momentum of the photon before the collision?

After the collision, the photon returns along its original path and the electron moves forwards with a momentum of 2.1×10^{-23} kg m s^{-1}.

- b What is the de Broglie wavelength of the electron after the collision?

c Is the wavelength of the photon now:

i the same as

ii greater than

iii less than before?

Give reasons for your answer.

d Calculate the wavelength of the photon before and after the collision.

8 A beam of neutrons with momentum $1.55 \times 10^{-23}\,kg\,m\,s^{-1}$ is incident on a crystal and produces consecutive interference maxima at angles 18.9° and 40.4°. What is the lattice spacing, *d*, in this crystal?

Reflecting

9 Think of all the devices that you have used today that rely on semiconductors. How would your life be different if quantum physics had not been developed?

10 Read about the 'Many Worlds' interpretation of quantum mechanics. Write a brief summary of this interpretation. Reflect on how this interpretation fits with either the hidden variables or probabilistic model of quantum mechanics. Justify this interpretation based on the arguments of one of these two models.

CHAPTER SUMMARY

- The quantum theory was developed when classical models, such as the wave model of light, were unable to explain experimental results such as the black body spectrum, atomic line spectra and the photoelectric effect.
- The development of quantum theory was a revolution in science, and not a smooth transition between theories. It was surrounded with great argument and debate.
- Atoms emit spectra with discrete lines with characteristic frequencies when heated. They also absorb light with the same characteristic frequencies. These are called line spectra.
- The line spectrum of each atomic species is unique, and enables atoms to be identified by their spectra.
- Classical models of the atom such as Thomson's 'plum pudding' model and Rutherford's 'planetary' model could not explain the existence of line spectra.
- The Bohr model, which was a semi-classical model, was able to explain the existence of atomic spectra, and predict the positions of lines in the hydrogen spectrum.
- The Bohr model assumed that the orbital radius of an electron and hence the kinetic energy of the electron and the potential energy of the electron–nucleus system were quantised.
- This quantisation arose from the assumption of quantisation of the angular momentum of the electrons: $L = mvr = nh$
- The electrons in the Bohr model are treated as standing waves with wavelength $\lambda = \frac{2\pi r}{n}$. This was a later addition to the model as a result of de Broglie's work.
- Transitions of electrons between the quantised energy levels correspond to lines in the atomic spectra.
- Each transition is the result of a photon being absorbed or emitted. The photon must have an energy exactly corresponding to the gap between energy levels.
- The Bohr model was not able to predict the spectra of atoms other than hydrogen, or explain features such as line splitting.
- Particles such as electrons also have a wave nature, this is known as wave–particle duality.
- The wavelength of a particle with mass, such as an electron, is given by the de Broglie wavelength: $\lambda = \frac{h}{mv} = \frac{h}{p}$
- Particles such as electrons show interference effects, such as an interference pattern produced when they are passed through a pair of slits of separation comparable to their de Broglie wavelength.

- Single-particle interference has been observed with photons, electrons and other particles. This behaviour is explained in the modern quantum model as being due to probability waves associated with particles interacting with both slits.
- In the modern quantum theory, the behaviour of particles is probabilistic, rather than deterministic. The theory has extremely good explanatory and predictive power, hence is considered a good or effective theory.
- This probabilistic nature of the theory was unacceptable to some physicists, including Einstein. However, they were unable to offer a better alternative theory.
- The theory has led to the development of modern electronics, computing and communications technology. It has made an enormous impact on our lives.

CHAPTER GLOSSARY

absorption spectrum the wavelengths (or frequencies or energies) of radiation absorbed by a material

angular momentum, *L* momentum associated with rotational or orbital motion, $L = mvr$

deterministic predictable, able to be determined if enough information is available

electron diffraction the interference of electrons due to the interaction of their probability waves

emission spectrum the spectrum of radiation emitted by an object, for example black body radiation or atomic spectra from a discharge tube

energy levels the allowed energies of a nucleus–electron system; often referred to as electron energy levels, even though they are characteristic of the atom, not of individual electrons

exclusion principle the principle that no two electrons in an atom may have identical sets of quantum numbers; it was first stated by Wolfgang Pauli in 1925

ground level the lowest possible energy level of a nucleus–electron system

line spectrum an emission or absorption spectrum consisting of discrete lines, characteristic of the energy levels of a particular atom or molecule

probabilistic not deterministic, unable to be predicted regardless of how much information is known

spectroscope a device that disperses radiation by energy (or wavelength or frequency) so that a spectrum may be observed and measured

wave–particle duality the dual nature of matter and energy, requiring both wave and particle models to completely explain all observed behaviour of matter and energy

CHAPTER REVIEW QUESTIONS

Remembering

1 Name three physicists who contributed to the development of the model of the atom. Briefly describe their contributions.

2 What were the most important successes and shortcomings of the Bohr model?

3 What physical properties of the nucleus–electron system were quantised in the Bohr model?

Understanding

4 The spectral lines for transitions to the $n = 2$ state from higher levels for a particular atom are in the infrared region of the spectrum. Would you expect this atom to have any series of spectral lines in the visible range? If so, to what transitions would they correspond?

5 In the Bohr model of the hydrogen atom, the electron is described as a standing wave. How many nodes does this wave have in each of the first four modes of vibration?

6 An electron has the same de Broglie wavelength as a photon. In what ways does the electron differ from the photon? In what ways is it the same?

7 Consider the Davisson–Germer experiment in which electron diffraction was observed for electrons reflecting off layers of atoms in a nickel crystal. How would the pattern they observed have been different if they had used a larger potential difference to accelerate the electrons? Explain your answer.

8 Explain the difference between deterministic and probabilistic. Why do you think this was an important difference between the classical (Newtonian) and quantum theories?

9 The ground-state energy level of the electron in a hydrogen atom is negative. What is the significance of the negative sign? What is the zero energy level in this model?

Applying

10 Figure 8.15 shows an energy level diagram for sodium.

a How many possible emission spectral lines are there for transitions from the −2.11 eV level?

b What are the wavelengths of the photons emitted in these transitions? If any are visible, identify the colour.

11 What is the energy of the shortest wavelength photons that can come from hydrogen atoms as they return from excited states to the ground state?

12 Two energy levels within a particular atom are 10 eV and x eV. When an atom of the element concerned returns from the higher level to the lower energy level, radiation of wavelength 450 nm is emitted. What are the possible values of x?

13 a What is the wavelength of X-ray photons of energy 3.6×10^4 eV?

b At what energy do electrons have a de Broglie wavelength equal to that of 3.6×10^4 eV X-rays?

14 An atom is excited to its third energy level above the ground state.

a How many different spectral lines can it emit?

b Which of the energy level transitions will produce the photon of greatest energy?

c Which of the energy level transitions will produce a photon of the longest wavelength?

15 A beam of electrons with de Broglie wavelength 1.5×10^{-8} m is incident on a pair of slits spaced 150 μm apart. Calculate the positions of the first three maxima in the interference pattern detected at a detector bank placed 25 cm from the slits.

16 In question 15 above, what de Broglie wavelength would the electrons need for the separation of interference maxima to be doubled?

Analysing

17 Hydrogen atoms have only one electron, yet the spectrum of hydrogen contains a large number of lines. Explain how this is possible.

18 A polished lead surface is illuminated with light of wavelength 200 nm. What are the maximum and minimum de Broglie wavelengths of the emitted photoelectrons?

19 Figure 8.7 shows the Bohr model of the atom.

a Write an expression for the force exerted on the electron by the nucleus (refer to Chapter 3).

b Set this force equal to the centripetal force, and find an expression for the kinetic energy of the electron.

c Find an expression for the total energy of the atom, assuming the nucleus is stable and given that the potential energy of the nucleus–electron system is $U = -\frac{k_e q_1 q_2}{r}$.

d Given that the ground level for hydrogen is −13.6 eV, find the value of r, the orbital radius of the electron. This value is known as the Bohr radius.

20 Electrons travelling at a speed of 150 m s^{-1} are incident on a pair of slits spaced 0.015 mm apart. An array of detectors is placed 2.00 m behind the slits.

a What wavelength photons would give an interference pattern with maxima at the same positions?

b Find the positions of the first two interference maxima.

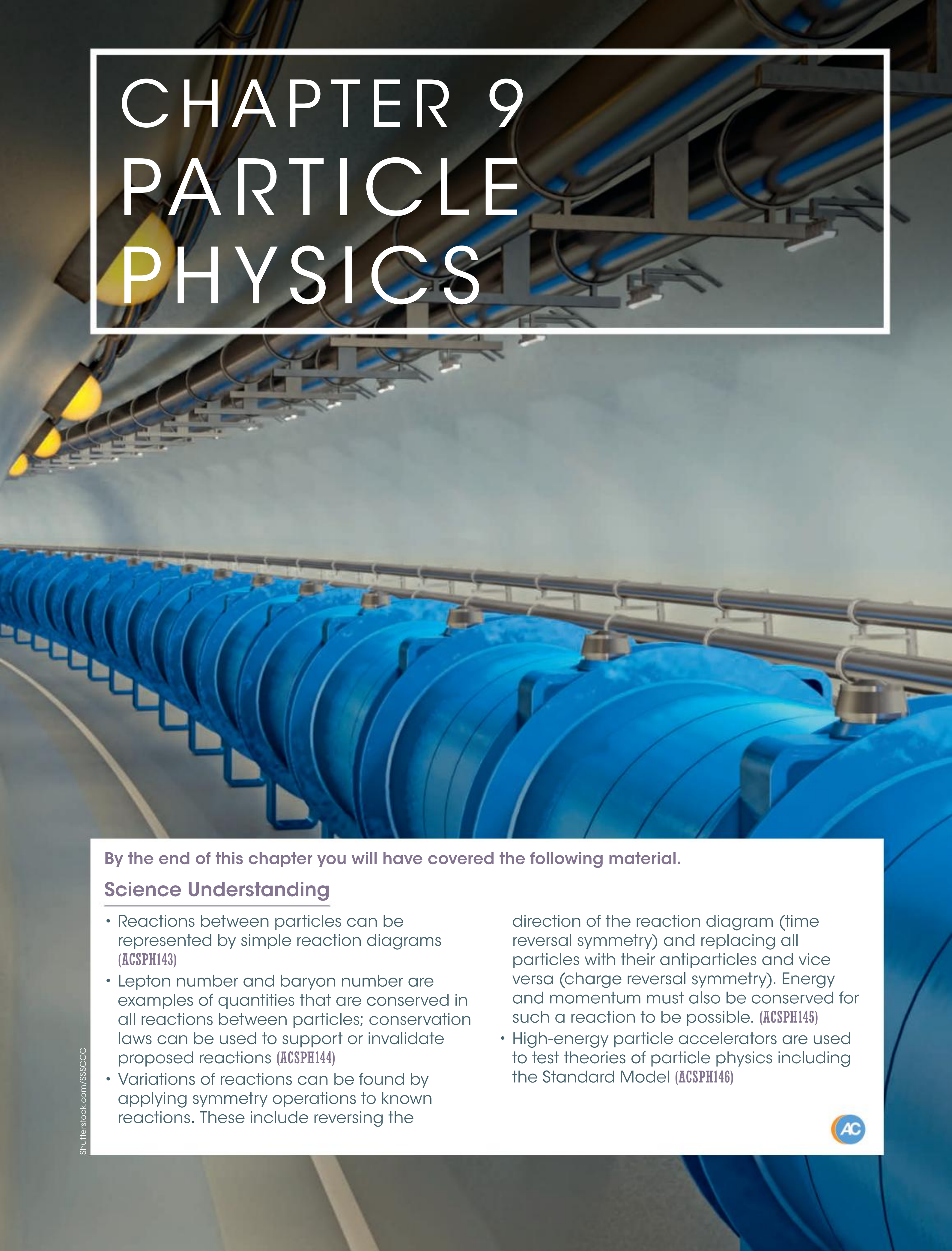
21 A crystal has a lattice spacing of 0.52 nm. In an electron diffraction experiment using this crystal, the first interference maximum occurs at $\theta = 34°$. What potential difference was used to accelerate the electrons for this experiment?

Reflecting

22 In what ways does the study of quantum theory help you to understand the interplay between theory and experiment? Give examples.

23 Summarise the main ideas and mathematical formulations in this chapter in a concept map. Link into material learnt in previous chapters.

24 Write a diary entry for a normal school day. Include how you got to school, what lessons you did and what you used, and what you did after school, including how you did homework or communicated with friends. Now write an imaginary diary entry for the same day, but if quantum physics and all the technology that relies on it had not been invented. Include the same activities and how they would be done instead, if possible at all.



Shutterstock.com/SSSCCC

CHAPTER 9 PARTICLE PHYSICS

By the end of this chapter you will have covered the following material.

Science Understanding

- Reactions between particles can be represented by simple reaction diagrams (ACSPH143)
- Lepton number and baryon number are examples of quantities that are conserved in all reactions between particles; conservation laws can be used to support or invalidate proposed reactions (ACSPH144)
- Variations of reactions can be found by applying symmetry operations to known reactions. These include reversing the direction of the reaction diagram (time reversal symmetry) and replacing all particles with their antiparticles and vice versa (charge reversal symmetry). Energy and momentum must also be conserved for such a reaction to be possible. (ACSPH145)
- High-energy particle accelerators are used to test theories of particle physics including the Standard Model (ACSPH146)

Introduction

One of the key characteristics of physics as a science is the belief that complex systems can be understood by examining and understanding the motion and interactions of simpler, component parts. Underlying this belief is another: that, as we break the system into smaller and smaller parts, we will eventually come to a point where there are no internal components. Then we will be dealing with the basic building blocks out of which everything else is constructed.

The idea that there are fundamental building blocks of matter has persisted through the ages, although it has undergone many changes as new discoveries and observations have been made. The ancient Greeks thought that the universe was made out of four basic elements: earth, air, fire and water. Then Democritus introduced the concept of the atom. Later it was found that atoms were made of even smaller components.

We now believe that the fundamental building blocks are **elementary particles**.

Over the past century, scientists and engineers working in particle physics have made a series of significant discoveries. These discoveries have led to new theories that describe and predict these elementary particles and their interactions. In this chapter, we describe some of the sequence of observations and discoveries in particle physics. These highlight the important connections between experimental and theoretical progress in physics. In the next chapter we introduce the model that describes our current understanding of particle physics: the Standard Model of Particle Physics (the Standard Model for short).

Splitting the atom

The history of the development of atomic models from Thomson's plum pudding model through to the modern quantum mechanical model is given in Chapter 8.

For a long time atoms were thought to be indivisible, in fact that is what the name 'atom' means: *a* – not, *tom* – divisible. Then, the electron was discovered by J.J. Thompson in 1897, leading to the understanding that atoms are *not* elementary particles.

It was known from Rutherford's scattering experiments (see Chapter 8) that most of the mass of the atom was localised in a massive, positively charged nucleus. The nucleus of the hydrogen atom was called a proton by Rutherford. The mass and charge of other atomic nuclei were measured, and it was found that each nucleus had a charge equal to some integer number of protons, but that the mass was generally equal to approximately twice this number of proton masses or more.

In 1932 this observation was explained when James Chadwick discovered the neutron. Chadwick also used scattering experiments in which particles were collided with nuclei.

As we shall see, physicists have learnt a lot about the nature of matter by smashing things into other things and seeing what happens!

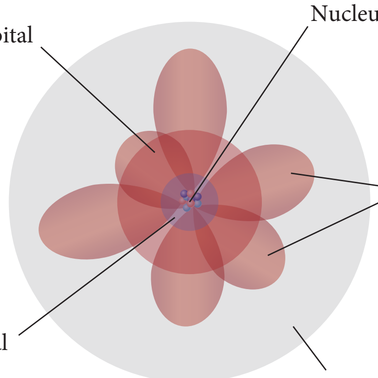


Figure 9.1 ▲ In the modern quantum mechanical model of the atom, the protons and neutrons are contained in the nucleus, surrounded by the electron clouds.

Protons and neutrons are not elementary particles

The model of the atom that you have learnt about in Chapter 8 was largely developed as of 1932. There is a tiny, massive nucleus that contains positively charged protons and neutral neutrons. Surrounding this is a cloud of negatively charged electrons.

A fourth particle, the photon, was also already known by the 1930s. As we saw in Chapter 7, the idea of a particle of light was introduced to explain the photoelectric effect. Recall that the photon has no mass and no charge. It does have quantised energy and momentum, which depend on its frequency. The electron and the photon both appear to be elementary particles.

Recall that when we are considering the gravitational force between the Sun and Earth, we can treat Earth as if all its mass were concentrated at its centre and use $F = GM_E\frac{m}{r^2}$.

However, Earth is not really a point mass because it has *internal structure*. When we are close to Earth's surface, the gravitational field no longer appears to be that due to a point mass; the internal structure becomes important. Similarly, a charged object can be treated as a point charge when it interacts with another charged object a long way away, but close up we must treat it differently.

The electric field of an electron is exactly what we expect for a point charge, suggesting that it has no internal structure and strongly supporting the idea that it is an elementary particle.

On the other hand, experimental evidence suggests that the neutron is *not* an elementary particle. Although the neutron has no charge, it does have a magnetic moment and hence its own intrinsic magnetic field. As we have seen in Chapter 4, a magnetic field is the result of moving charged particles or a changing electric field. This indicates that the neutron has some internal structure, and contains charges that add to zero total charge. This suggested to physicists that the neutron is *not* an elementary particle.

As the proton has a mass very close to that of a neutron and is otherwise a similar particle in its behaviour, the same question was also raised of the proton. Evidence that the proton is not an elementary particle came from studies of radioactivity. You have seen (*Nelson Physics Units 1 & 2 for the Australian Curriculum*, Chapter 3) that when a nucleus decays, different types of particles may be emitted. These include β^- and β^+ particles. It was observed that these β particles had the same mass as an electron, and either a positive or negative charge equal to the electron charge.

It turned out that the β^- particle was an electron. When a nucleus emits a β^-, its proton number increases by one and its neutron number decreases by one. Similarly, when a β^+ particle is emitted, a proton is converted to a neutron. Hence, it appeared that neutrons could be converted into protons and vice versa by the emission of these negatively or positively charged electrons, suggesting that neither the neutron nor the proton is an elementary particle.

The positively charged electrons, called **positrons**, that are involved in β^+ decay were discovered in cosmic rays in 1932 by Carl Anderson. Positrons were the first **antimatter** particle to be observed.

By 1932 there were five particles known: three 'normal' matter particles, the electron, proton and neutron; one antimatter particle, the positron; and the photon. Of these, it was already believed that the proton and neutron were not themselves elementary particles, and so the hunt for their component parts began.

Antimatter

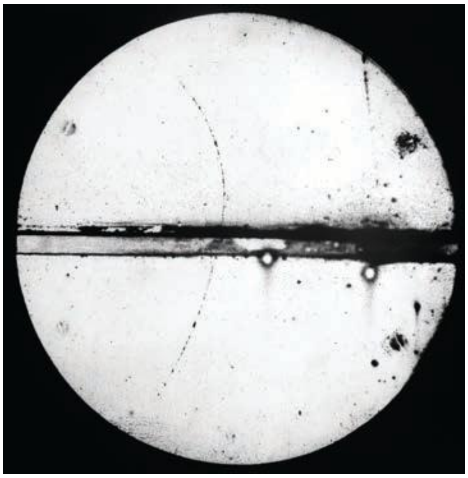
In the 1920s, Paul Dirac developed a relativistic version of the Schrödinger equation. Recall from Chapter 8 that the Schrödinger equation is the wave equation that describes the behaviour of particles such as electrons. This new relativistic equation explained the origin of the electron's spin and magnetic moment. This was an important theoretical development in quantum mechanics.

However, there appeared to be a difficulty with the theory: the equation that Dirac had developed to describe the electron had two solutions. One of these solutions gave the wave functions for electrons and correctly described their mass, charge and spin. The other solution described a particle with the same mass but the opposite charge and magnetic moment. It also predicted that if these two particles should meet, they would both be destroyed, producing a burst of energy. This is called annihilation.

Spin is a quantum number that can take half-integer or integer values. Recall from Chapter 4 that magnetic moment is a vector, it has direction and can have a positive or negative value. The magnetic moment of a particle depends on the spin, charge and mass of the particle.

Hence this second particle was the 'anti-electron'. It was this particle, the **antiparticle** to the electron, that Anderson observed in 1932.

Anderson was using a **cloud chamber** to study cosmic rays. A cloud chamber is a particle detector that detects ionising radiation (see Chapter 3 of *Nelson Physics Units 1 & 2 for the Australian Curriculum*). It uses a supersaturated vapour, like a layer of fog, in a container or chamber. When a charged particle enters the chamber it ionises the vapour. The resulting electrically charged vapour particles act as sites for condensation of the vapour. This produces a visible trail of condensation along the path of the particle.



Science Photo Library/Emilio Segre Visual Archives

Figure 9.2 ▲ This cloud chamber photograph shows the track of the first identified positron.

Anderson placed his cloud chamber in a magnetic field. The magnetic field caused the moving charged particles to follow curved paths, just like the ions in a mass spectrometer, as we saw in Chapter 4. This allowed him to distinguish between positive and negative charges. Recall that the direction in which a particle's path curves depends on its charge, and that the sharpness of the curve (radius of curvature) depends on the mass.

Anderson noted that some of the tracks in his cloud chamber had electron-like curvature but were deflected in a direction corresponding to a positively charged particle.

Anderson had discovered the anti-electron, or positron. Anderson was awarded a Nobel Prize for Physics in 1936 for this discovery.

Dirac's theory suggested that *an antiparticle exists for every particle*. It has subsequently been verified that almost every known elementary particle has a distinct antiparticle.

The antiparticle for a charged particle has the same mass as the particle but opposite charge and hence opposite sign for magnetic moment.

Antiparticles for uncharged particles, such as the neutron, are more difficult to describe. For neutral particles with non-zero magnetic moments such as the neutron, the antiparticle can be defined by the sign of its magnetic moment. A few particles, such as photons, are their own antiparticle.

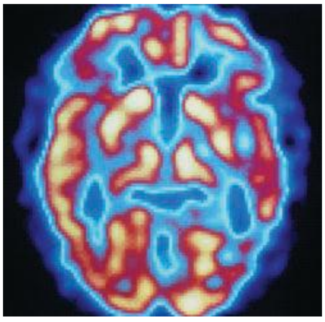
When we write the symbol for a particle, such as n for neutron or e^- for electron, we represent the antiparticle either by placing a bar over the symbol, as in $\overline{n}$, or by showing that the sign is reversed, as in e^+.

WOW

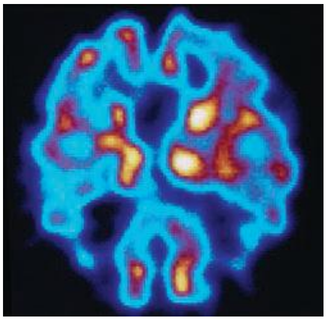
PET – how annihilation is saving lives

Electron–positron annihilation is used in the medical diagnostic technique called positron-emission tomography (PET). The patient is injected with a glucose solution containing a radioactive substance that decays by positron emission, and the glucose is carried throughout the body by the blood. A positron emitted during a decay event in one of the radioactive nuclei in the glucose annihilates with an electron in the surrounding tissue, resulting in two gamma-ray photons emitted in opposite directions. A gamma detector surrounding the patient pinpoints the source of the photons. A computer then displays an image of the sites at which the glucose has accumulated. This is useful for locating cancers because glucose metabolises rapidly in cancerous tumours and accumulates at those sites, providing a strong signal for a PET detector system.

The images produced in a PET scan can also indicate a wide variety of disorders in the brain, including Alzheimer's disease.



Science Photo Library/Tim Beddow



Science Photo Library/Tim Beddow

▲ **Figure 9.3** PET scans of the brain of a healthy older person (left) and that of a person with Alzheimer's disease (right). Lighter regions contain higher concentrations of radioactive glucose, indicating increased brain activity.

EXPERIMENT 9.1

BUILD YOUR OWN CLOUD CHAMBER AND DETECT COSMIC RAYS

There are very few particle physics experiments that you can do without expensive, and usually high voltage, equipment. However, with some fairly simple equipment you *can* build your own cloud chamber that will detect cosmic rays. You will need some chemicals and dry ice so you need to be very careful and follow all safety instructions. Typically, it takes about 10 to 20 minutes to detect a high-energy cosmic ray, depending on solar flare activity, so you may also need to be a bit patient after you have built your cloud chamber.

Aim

To build a cloud chamber and detect cosmic rays

Materials

- a clear glass or clear plastic tank, such as a small fish tank, about 15 cm tall and 15 cm wide by 30 cm long
- a strong source of light such as an overhead or slide projector
- a sheet of metal for a lid for the tank (same size as tank)
- a sheet of cardboard cut to fit the metal lid
- 3 sheets of felt, 30 cm by 30 cm
- a whole roll of black electrical tape
- foam padding, about 5 cm thick, same dimensions as the tank
- cardboard box just slightly bigger than the clear tank
- glue that is not soluble in alcohol such as silicon sealant
- isopropyl alcohol (isopropanol) (pure, about 500 mL)
- dry ice, about 500 g
- safety equipment: tongs, gloves, lab coats, safety glasses

What are the risks in doing this experiment?	How can you manage these risks to stay safe?
Isopropyl alcohol is toxic and can cause skin irritation. Refer to materials safety data sheet (MSDS) for more information.	Wear a lab coat, safety glasses and gloves when using isopropyl alcohol. Use in a well-ventilated space or fume cupboard. Dispose of gloves and wash hands thoroughly at the end of the experiment.
Isopropyl alcohol is highly inflammable.	Store isopropyl alcohol away from any sources of heat and ensure bottles are correctly labelled.
The isopropyl alcohol vapour used in the cloud chamber has a low flash point.	Keep chamber well away from all heat and flame sources.
Dry ice is very cold and can cause cold-burns.	Wear thick gloves and use tongs to handle dry ice.

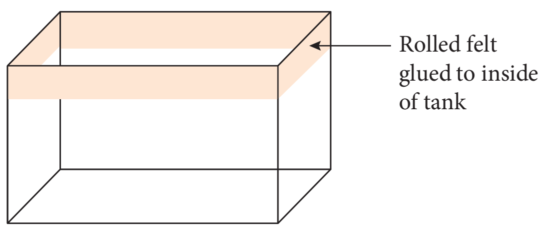
In your write-up, add any more risks you can think of, as well as ways to manage them.

Procedure

Building the cloud chamber will take some time. You will probably need a whole lab class to build your cloud chamber and then a second class to test it. Stop at step 6 unless you have ample time to proceed and test your cloud chamber.

1 Fold or roll the felt into strips about 5 cm wide and 30 cm long. Attach these to the inside of the clear container with the glue or sealant to form a ring around the bottom, as shown in Figure 9.4. Give the glue or sealant time to dry. (You may want to do this the day before.)

2 Cover one side of your cardboard sheet with strips of black electrical tape, as neatly as you can. This is to make particle tracks easier to view.

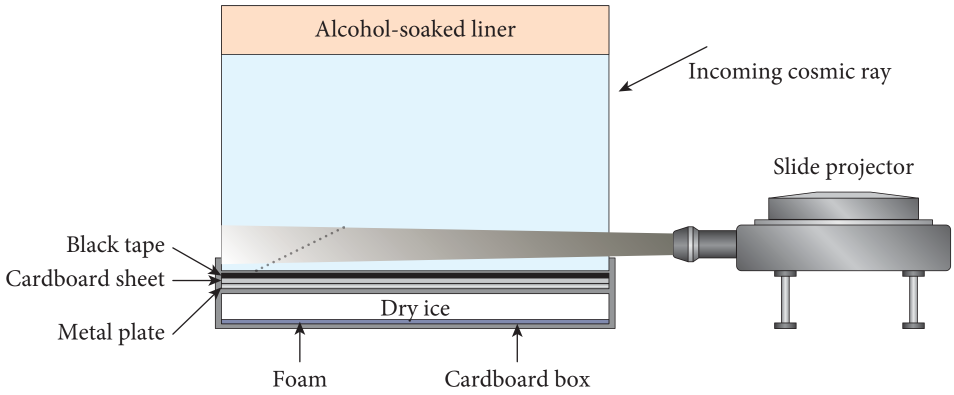


▲ **Figure 9.4** Felt folded and attached to tank as a 'soak zone'

3 Attach the cardboard, tape side out, to the metal lid so that when the lid is in place the black tape faces into the tank.

4 Cut the cardboard box down to about 7 cm tall and place the piece of foam in the bottom of it.

5 Check that your cloud chamber all fits together: the metal lid sits on top of the foam with the black tape facing upwards. The tank then sits upside-down on top of this so the ring of felt padding is at the top.



◀ **Figure 9.5** Cloud chamber and light source

It needs to all fit tightly and be well sealed when you are performing your observations. If it is not, air currents will make the particle tracks hard to see.

If it all fits together neatly, then you are ready to add the dry ice and isopropyl alcohol. If not, you will need to make some modifications until it fits neatly together.

6 Remove the tank and metal lid and place a layer of dry ice on top of the foam using tongs. Your teacher may do this step for you.

7 Soak the felt strips with isopropyl alcohol. Your teacher may do this step for you.

8 Seal the metal lid to the tank with tape (more electrical tape, or duct tape).

9 Place the tank upside-down on top of the dry ice so the metal lid is in contact with the dry ice.

10 Arrange the light source so it shines horizontally through the side of the tank. You need a bright light source to clearly illuminate the particle tracks.

11 Turn on the light, and observe your chamber for at least 10–20 minutes.

You should see a mist-like fog form inside your chamber. This will be most obvious near the bottom; don't worry if you can only see a thin layer, that will be enough.

Results

1 During the 10 to 20 minutes you observe your chamber, you should see fine tracks forming in your chamber. The tracks, which are the results of the random passage of cosmic rays through the vapour, can form at any time and will last only a very brief time before they disappear, so you will need to watch carefully for the entire period. Draw these tracks and/or photograph or video them.

2 Once your chamber is working, you can try putting a strong magnet on one side and observe what happens to the new particle tracks. Note that you will need a *very strong* magnet to obtain any significant curvature of the particle paths.

Analysis of results

Record as many tracks as you can. Look for sudden changes of direction in tracks that are straight lines. These could be muon decays, with the incoming track the muon and the outgoing track the electron into which it decays. (You may remember muons from Chapter 6). You may also see tracks that branch like a Y. These are usually due to collisions, for example a muon colliding with an electron and transferring some kinetic energy to it. The stem of the Y is the incoming muon and the two branches are the muon and electron moving off after the collision.

Discussion

Can you identify what is happening in the various tracks that you observed?

You might only be able to hypothesise what the incoming particles were.

MORE INFORMATION

This is a step-by-step guide to building a cloud chamber.

TROUBLE-SHOOTING GUIDE

This describes a few common problems you might have with your cloud chamber and how to solve them.

Particle physics: continuing the search for elementary particles

Following on from the discoveries described above, many 'new' particles were discovered. Some of these were discovered in cosmic-ray experiments like that carried out by Anderson. Other particles were discovered in nuclear decays.

The mass–energy relationship, $E = mc^2$, was derived in Chapter 6.

Cosmic rays pass through Earth's atmosphere at random, and their energies vary widely. Hence it was impossible to design well-controlled experiments that used cosmic rays to create new particles.

The energies of the particles emitted in some nuclear decays are well defined, but they are generally fairly low. The equation $E = mc^2$ suggested that if higher energies could be used, more massive particles might be created.

Thus, physicists looked for ways to produce beams in which there were more particles with high enough energies to produce, and so enable them to study, new particles. The breakthrough was achieved with the invention of **particle accelerators**.

From the 1950s onwards, many more particles were discovered in experiments involving high-energy collisions between known particles using these accelerators. The more energy available in the collisions, the higher the number of the new particles produced and the greater their mass.

These new particles are characteristically very unstable and have very short half-lives that range between 10^{-6} s and 10^{-23} s. Their decays produce lighter particles, some of which are also unstable. So each collision between just two initial particles may result in many outgoing particles, which need to be detected and identified simultaneously. To enable this, physicists and engineers worked together to develop huge, complex apparatus to use at the new particle accelerators. This is an example of the role that technology plays in allowing new experiments, which in turn lead to the development and refinement of theory.

Particle accelerators

In order to detect a particle with a very short half-life, it is necessary to first create it. This is done by causing reactions between stable particles such as protons, electrons and neutrons. However, protons and neutrons have small masses compared to most other particles.

To create a particle of large mass, m, there needs to be at least an energy of $E = mc^2$ available to convert into mass. If the total mass of the reacting particles is less than m, some of this energy has to come from the kinetic energy of the reacting particles. Particle accelerators provide this extra energy.

A particle accelerator uses an electric field to accelerate charged particles to very high speeds and hence very high energies. A magnetic field may also be used to contain the charged particles within the accelerators.

A **linear accelerator** or **linac** accelerates the particles in a straight line. In a linear accelerator, large electric fields are used to accelerate the changed particles, such as electrons or protons. Recall from Chapter 3 that an electric field, E, exerts a force $F = Eq$ on a particle with charge q. In a linear accelerator this force accelerates the particle to very high speeds. We can also use the effective change in potential, ΔV, that the particle passes through in the linear accelerator to find the change in energy and hence the final speed of the particle.

The change in kinetic energy is $W = \Delta U = q\Delta V = \frac{1}{2}mv^2$. The final speed is hence

$$v = \sqrt{\frac{2q\Delta V}{m}}$$

This is the classical expression, derived using classical electromagnetism and mechanics. As the accelerating potentials may be very large, and the particles reach speeds approaching the speed of light, we need to use the relativistic correction as described in Chapter 6:

$$\text{KE} = q\Delta V = \frac{1}{2} m_0 c^2 (\gamma - 1)$$

where m_0 is the rest mass of the particle and

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Relativistic expressions for energy and momentum are given in Chapter 6.

This gives us a value for the speed, which is less than that given by the classical expression, and which asymptotically approaches the value of c as the accelerating potential is increased.

At many facilities the linear accelerator is used to feed high-speed particles into a synchrotron ring. The synchrotron ring uses large magnetic fields to contain the charged particles in a circular path. Recall from Chapter 4 that the magnetic force supplies the required centripetal force:

$$F = \frac{mv^2}{r} = qvB$$

Recalling from Chapter 9 of *Nelson Physics Units 1 & 2 for the Australian Curriculum* that momentum is given by $p = mv$, we can see from the above equation that the momentum of a particle circulating in a magnetic field is given by

$$p = mv = qBr$$

Synchrotrons are described in more detail in Chapter 4. The Australian Synchrotron, located in Melbourne, is described in the Scientific literacy box in Chapter 5.

As the particles are likely to be at very high speeds in a synchrotron, we need to use the relativistic expression for momentum from Chapter 6:

$$p = \gamma m_0 v = qBr$$

In the synchrotron ring, the particles may be held for storage at constant speed, or they may be further accelerated. To increase the speed of the particles, a varying magnetic field must be used. Recall from Chapter 4 that a constant magnetic field cannot do work on a charged particle and hence change its kinetic energy. However, a time-varying magnetic field creates an electric field (Chapter 5), which *can* do work on a charged particle and accelerate it. By carefully varying the field in a periodic way, a synchrotron can be used to further increase the speed of the particles.

Once the charged particles have been accelerated to the desired speeds they are aimed, again using magnetic and electric fields to steer them, at a target. They are made to smash into the target at very high speed, close to the speed of light.

WOW

'Parasitic' radiation

In the 1950s and 60s synchrotron radiation was called 'parasitic' radiation, because it took energy from the accelerating particles intended for use in high energy and nuclear physics experiments. At early particle physics facilities this radiation was shielded against and not used. Later, it was realised that this high-energy electromagnetic radiation is useful in its own right, and in 1961, at the National Institute of Standards and Technology (NIST) in the USA, the first beamline into a synchrotron was opened up to allow use of this radiation. Beamlines were subsequently opened at other synchrotrons. The experiments and beamlines (and even the scientists who used them) were also referred to as parasitic. However, the science that they did with their parasitic experiments turned out to be very useful. Now there are large synchrotrons, such as the Australian Synchrotron in Melbourne, purpose built to produce and use this synchrotron radiation.

Detectors placed around the target are then used to measure the momentum, mass and charge of the particles that result from the collision. These particles might be a mixture of known and new types of particles. In some particle accelerators the target is held fixed in place and all the energy needed for the reaction has to come from the kinetic energy of the accelerated charged particle.

To reach even higher energies and so have the chance of creating particles with even larger mass, scientists and engineers developed colliding-beam accelerators called **colliders**.

In the simplest type of collider, particles that have equal masses but opposite charges, such as protons and antiprotons, travel in opposite directions in an accelerator ring. Because they have equal mass and are accelerated by the same electric field, they have equal kinetic energy. The particles collide head-on, and many new particles may be produced in the high-energy collision. Because the total momentum of the interacting particles is zero, all their kinetic energy is available for the reaction. Recall from your study of collisions in Chapter 9 of *Nelson Physics Units 1 & 2 for the Australian Curriculum* that momentum must be conserved in a collision, but kinetic energy need not be.

You studied collisions in Chapter 9 of Nelson Physics Units 1 & 2 for the Australian Curriculum. Recall that momentum is conserved for all collisions. Kinetic energy is conserved in elastic collisions.

Particle accelerators are massive pieces of equipment. They are extremely expensive to build and maintain, and so require the collaboration of governments, universities and international organisations and committees such as the International Science Advisory Committee and the International Machine Advisory Committee.

Particle accelerator facilities are used for a range of applications and research purposes, including testing the predictions of particle physics and of models such as the Standard Model (see Chapter 10).

The Large Hadron Collider (LHC)

THE LHC GAME

Play the LHC game at CERN.

The Large Hadron Collider at CERN is the world's largest and most powerful particle accelerator. It took 10 years to build, cost around 3 billion euros, and began operating in September 2008. Thousands of scientists and engineers collaborated to design and build the LHC and the associated detectors and other instruments.

Linear accelerators are used to give beams of protons energies of up to 4 TeV (10^{12} eV) before they enter the LHC synchrotron ring. After upgrades made in 2013–14, the proton beams have energies up to 7 TeV, which means the protons travel at 0.999 999 991 times the speed of light. Lead nuclei are also used in some experiments, and these can be accelerated to energies of nearly 600 TeV per nucleus.

The synchrotron ring uses superconducting electromagnets to steer and contain particle beams in a 27 km long ring.

When beams of protons collide head-on, a huge variety of particles are created. Seven detector stations are used at beamlines around the ring. These stations are where the particles are detected and their properties such as mass, charge and spin are measured.

Spin was discussed in Chapter 4, in which we saw that electrons have an intrinsic magnetic moment. We describe this magnetic moment as being due to a property that we call 'spin', even though we don't believe the electrons are actually spinning.

Data is collected at a very high rate. The LHC is connected to the world's largest computer grid for data analysis. The grid consists of 140 computing centres in 35 countries.

The most famous discovery made at the LHC to date is that of the Higgs boson. This is discussed in detail in the scientific literacy box in the next chapter.

Getty Images/AFP

◀ **Figure 9.6** Inside the tunnel containing the Large Hadron Collider at CERN

The particle zoo

So far, several hundred particles have been identified in particle accelerator experiments. The wide variety of properties and behaviours of these new particles led to the term 'particle zoo'. These newly discovered particles included the antiproton, discovered by Emilio Segré and Owen Chamberlain in 1955, and the antineutron, discovered in 1956 by Bruce Cork.

Some of these particles are listed, along with their properties, in Table 9.1. The masses in Table 9.1 are given in units of MeV/c^2. Recall from Chapter 6 that mass and energy are related by Einstein's mass equivalence relationship: $E = mc^2$. In particle physics it is common to give masses in terms of their energy equivalent, rather than in kilograms. $1\,\text{MeV}/c^2 = 1.78 \times 10^{-30}\,\text{kg}$.

Spin is another important property of particles listed in Table 9.1.

Like other properties of particles, spin is quantised; that is, it can take only specific discrete values. Some particles have spins with integer values, others have half-integer values of spin. These values relate to the magnetic moment and hence magnetic fields of the particles. Particles with non-zero spin have a magnetic moment and hence their own magnetic field. How the spins of the various particles in a material interact is important in magnetism, as we saw in Chapter 4. Spin is also a conserved quantity.

Particles can be classified according to their spin. Particles with half-integer spin, $s = \frac{1}{2}, \frac{3}{2}$, and so on, are called **fermions**. **Leptons** and **baryons** are fermions. Fermions obey the Pauli exclusion principle (Chapter 8). The Pauli exclusion principle states that any two fermions in the same quantum system cannot have identical sets of quantum numbers. Electrons in an atom are an example of this – no two electrons in a given atom can have identical quantum numbers. Particles with integer spin, $s = 0, 1, 2, \ldots$ are called **bosons**. **Mesons** and photons are bosons. Bosons do not obey the exclusion principle. Fermions and bosons play different roles in the Standard Model, as we shall see in Chapter 10.

DISCOVERY OF THE ANTIPROTON

Read about the discovery of the antiproton.

ELEMENTARY PARTICLES

Read about the search for elementary particles here.

Table 9.1 Some particles and their properties

Category	Particle name	Symbol	Anti-particle	Mass (MeV/c²)	B	L_e	L_μ	L_τ	Lifetime (s)	Spin
Leptons	Electron	e^-	e^+	0.511	0	+1	0	0	Stable	$\frac{1}{2}$
	Electron-neutrino	ν_e	$\bar{\nu}_e$	<7 eV/c^2	0	+1	0	0	Stable	$\frac{1}{2}$
	Muon	μ^-	μ^+	105.7	0	0	+1	0	2.20×10^{-6}	$\frac{1}{2}$
	Muon-neutrino	ν_μ	$\bar{\nu}_\mu$	<0.3	0	0	+1	0	Stable	$\frac{1}{2}$
	Tau	τ	τ^+	1 784	0	0	0	+1	$<4 \times 10^{-13}$	$\frac{1}{2}$
	Tau-neutrino	ν_τ	$\bar{\nu}_\tau$	<30	0	0	0	+1	Stable	$\frac{1}{2}$
Hadrons										
Mesons	Pion	π^+	π^-	139.6	0	0	0	0	2.60×10^{-8}	0
		π^0	Self	135.0	0	0	0	0	0.83×10^{-16}	0
	Kaon	K^+	K^-	493.7	0	0	0	0	1.24×10^{-8}	0
		K^0_S	$\bar{K}^0_S$	497.7	0	0	0	0	0.89×10^{-10}	0
		K^0_L	$\bar{K}^0_L$	497.7	0	0		0	5.2×10^{-8}	0
	Eta	η	Self	548.8	0	0	0	0	$<10^{-8}$	0
		η'	Self	958	0	0	0	0	2.2×10^{-10}	0

Category	Particle name	Symbol	Anti-particle	Mass (MeV/c²)	B	L_e	L_μ	L_τ	Lifetime (s)	Spin
Baryons	Proton	p	$\bar{p}$	938.3	+1	0	0	0	Stable	$\frac{1}{2}$
	Neutron	n	$\bar{n}$	939.6	+1	0	0	0	614	$\frac{1}{2}$
	Lambda	Λ^0	$\bar{\Lambda}^0$	1115.6	+1	0	0	0	2.6×10^{-10}	$\frac{1}{2}$
	Sigma	Σ^+	$\bar{\Sigma}^-$	1189.4	+1	0	0	0	0.80×10^{-10}	$\frac{1}{2}$
		Σ^0	$\bar{\Sigma}^0$	1192.5	+1	0	0	0	6×10^{-20}	$\frac{1}{2}$
		Σ^-	$\bar{\Sigma}^+$	1197.3	+1	0	0	0	1.5×10^{-10}	$\frac{1}{2}$
	Delta	Δ^{++}	$\bar{\Delta}^{--}$	1230	+1	0	0	0	6×10^{-24}	$\frac{3}{2}$
		Δ^+	$\bar{\Delta}^-$	1231	+1	0	0	0	6×10^{-24}	$\frac{3}{2}$
		Δ^0	$\bar{\Delta}^0$	1232	+1	0	0	0	63×10^{-24}	$\frac{3}{2}$
		Δ^-	$\bar{\Delta}^+$	1234	+1	0	0	0	6×10^{-24}	$\frac{3}{2}$
	Xi	Ξ^0	$\bar{\Xi}^0$	1315	+1	0	0	0	2.9×10^{-10}	$\frac{1}{2}$
		Ξ^-	$\bar{\Xi}^+$	1321	+1	0	0	0	1.64×10^{-10}	$\frac{1}{2}$
	Omega	Ω^-	Ω^+	1672	+1	0	0	0	0.82×10^{-10}	$\frac{3}{2}$

The lifetimes given in Table 9.1 are the mean lifetimes, t_{mean}. Recall from your studies of radioactivity that unstable nuclei have a half-life that determines how likely they are to decay. In a large population of nuclei with a half-life $t_{\frac{1}{2}}$, one half of the nuclei will decay in a time $t_{\frac{1}{2}}$.
The mean lifetime is related to the half-life by: $t_{mean} = \frac{t_{\frac{1}{2}}}{\ln 2} \approx 1.44 t_{\frac{1}{2}}$.

We are using the symbol t_{mean} rather than τ to denote the mean lifetime, to avoid confusion with the τ (tau) particle. You may see lifetimes written as τ in other sources.

The other particle properties, **baryon number (B)** and **lepton number (L)**, are discussed in detail later.

QUESTION SET 9.1

Remembering

1 Why are electrons considered to be elementary but protons are not?

2 State the charge and mass of:

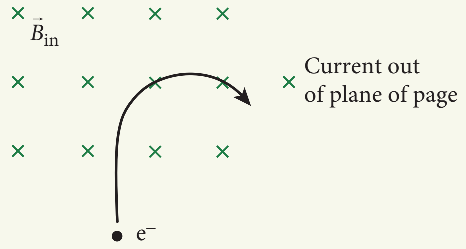
- a a positron.
- b an antiproton.
- c an antineutron.

3 What is the difference between a collider particle accelerator and one with a fixed target? Why is a collider type able to produce more massive particles as a result of a collision?

Understanding

4 A synchrotron with a constant magnetic field can hold a particle in orbit, but cannot change its speed. Explain why this is the case. Draw a diagram to help explain your answer.

5 An electron enters a region of uniform magnetic field and curves to the right as shown in Figure 9.7. A second charged particle enters the field following the same initial trajectory as the electron, but curves to the left. How could we tell if this second particle is a positron or a proton?



◀ **Figure 9.7** An electron's trajectory curves to the right when it enters the magnetic field.

Applying

6 At the Stanford Linear Accelerator Center (SLAC) electrons and positrons are accelerated in a tube 3 km long to energies of 50 GeV. What is the potential difference through which the particles are accelerated?

7 Consider a proton with energy of 4 TeV in the Large Hadron Collider.

a What is the energy of this particle in joules?

b How fast would a mosquito, with a mass of about 3 mg, need to fly to have this same kinetic energy? Comment on your answer.

8 The mass of an electron is 9.11×10^{-31} kg. Convert this mass into units of MeV/c^2.

9 A proton in a synchrotron ring has speed 0.999 *c*. What magnetic field is required to keep the proton in orbit if the orbital radius is 4.29 km? (Note that you will need to use a relativistic correction.)

Reflecting

10 What is your intuitive reaction to the idea that there may be hundreds of elementary particles? Critique the idea that for a theory to be correct it should be simple and elegant. What arguments can you find for and against this idea?

PARTICLE ADVENTURE

Work through the Particle Adventure and try the quiz.

The search for order

If you look back at Table 9.1 you will note that the particles listed are classified into two different types: leptons and **hadrons**.

Leptons are mostly particles with very small mass, such as the electron. We have already noted that electrons appear to be elementary particles because they show no signs of having internal structure. Because the heavier leptons and their antiparticles appear to be identical to electrons and positrons in all respects except for mass, it appears that they too are elementary particles. This leads physicists to believe that the whole family of leptons, including neutrinos, are elementary particles. Neutrinos are discussed further later in this chapter.

Hadrons mostly have larger mass, and are further divided into mesons and baryons. Families of particles with similar masses but different electric charges are seen in hadrons and not leptons.

You may have noted that the photon is not listed in Table 9.1. The photon is known as a **field particle**. This is a third variety of particle. Recall from Chapter 5 that light is an electromagnetic wave, made up of coupled electric and magnetic fields. Hence, the photon is the quantum of the electromagnetic field, as we saw in the previous chapter. This idea of field particles was immensely important in the search for underlying order in particle physics, as we shall see in the next section.

Particles can be divided into leptons, hadrons and field particles. Leptons, such as electrons, generally have small mass. Hadrons, such as protons and neutrons, generally have large mass. Field particles, such as photons, are involved in interactions between particles and may be massless.

Hadrons

We now believe that hadrons are not elementary particles. Evidence for this comes from the:

- existence of groups of hadrons with similar masses but different charges
- ways in which hadrons decay, often to other, lighter hadrons
- ways particular hadrons are produced in particular reactions, and
- magnetic moment of uncharged particles such as the neutron.

Hadrons interact via the strong nuclear, electromagnetic, gravitational and **weak forces**.

They are further divided into two classes, mesons and baryons, based on their masses and spins.

The gravitational, electromagnetic and weak forces are believed to be fundamental forces. As we shall in Chapter 10, the strong nuclear force is a consequence of another force, the strong force, which is believed to be fundamental.

Mesons

The name 'meson' means middle-sized. Several mesons have masses in the range between the masses of the electron and the proton, although mesons having masses greater than that of the proton have been found. Mesons all have zero or integer spin (0, 1, …) and hence are bosons.

No stable meson has ever been observed. All mesons decay. Some decays produce lighter (but themselves unstable) mesons, but all meson decay chains ultimately produce electrons, positrons, neutrinos and photons.

Baryons

The name 'baryon' means 'heavy' in Greek. Baryons have masses equal to or greater than the proton mass. Their spin is always a half-integer value ($\frac{1}{2}$, $\frac{3}{2}$, ...) and so they are fermions. Protons and neutrons are baryons. Protons are the only stable baryon. All others decay in such a way that the end products of the decay chain include a proton.

Leptons

Leptons (from the Greek *leptos*, meaning 'small' or 'light') are particles that do not interact by means of the **strong nuclear force**, but do interact via the gravitational, electromagnetic and weak forces. All leptons have spin $\frac{1}{2}$. Unlike hadrons, which have size and structure, leptons appear to be truly elementary, meaning that they have no structure and are point-like.

Also unlike hadrons, the number of known leptons is small. Currently, scientists believe that only six leptons (plus their antiparticles) exist: the electron, the muon and the tau, plus a neutrino associated with each. The antiparticles are the antielectron or positron, the antimuon and the antitau, and each of these has an antineutrino.

NEUTRINOS

You can read more about neutrinos here.

The neutrino was predicted in 1930 by Dirac. At this time radioactive beta decay had been experimentally observed, but momentum and energy did not appear to be conserved in these decays. Hence, Dirac proposed the existence of a very light, uncharged particle that could carry the unaccounted-for energy and momentum.

Current experiments indicate that neutrinos have a small, but non-zero, mass. Direct experimental evidence for the neutrino associated with the tau was announced by the Fermi National Accelerator Laboratory (Fermilab) in 2000.

Field particles

There are four forces believed to be responsible for all interactions. You have met these forces before in your study of physics. They are the gravitational force, the electromagnetic force, the strong nuclear force and the weak force. Each of these forces is mediated by a field particle.

In Chapters 3 and 4 we saw that electrons and other charged particles interact via the electric and magnetic fields. We also know from Chapter 8 that the electromagnetic field consists of particles called photons. The electromagnetic force can therefore be pictured in terms of the exchange of photons between electrically charged particles. The electromagnetic force is said to be *mediated* by photons.

In this particle exchange model of interactions, the basic process is that of one particle emitting a field particle that is subsequently absorbed by another elementary particle.

As we shall see in Chapter 10, the pion is now understood to be a composite particle. It is not an elementary particle, and is not in fact a true field particle like the photon. Instead, we now understand the nuclear force to be the most visible result of a more elementary interaction, the strong force, between the constituent particles that make up the nucleons themselves.

The idea that electromagnetic interactions could be described as being mediated by photon exchange led physicists to question whether other types of interaction might be modelled in the same way.

The discovery that atomic nuclei are composed of protons and neutrons led to the introduction of a new type of force. The protons in any nucleus should strongly repel one another due to their charges of the same sign. So there must be some other force acting to stop the nucleus flying apart. We call the force that holds the nucleus together the strong nuclear force, or the nuclear force, because it must be strong to overcome the proton–proton repulsion and it acts between nucleons.

The first theory attempting to explain the nature of the attractive force between nucleons was proposed in 1935 by Japanese physicist Hideki Yukawa. Yukawa was awarded the Nobel Prize in Physics in 1949 for his work.

Yukawa used the idea of **exchange particles** to explain the strong nuclear force. He proposed the existence of a new type of particle whose exchange between nucleons in the nucleus causes the strong nuclear force.

Yukawa predicted neutral, positively and negatively charged versions of his particle, all of which would be involved in interactions between protons and neutrons.

In 1947 a set of particles that matched Yukawa's predictions was discovered. These particles were named the pi mesons (π), or simply pions. The new pions came in electrically negative, neutral and positive varieties, and had zero spin.

Figure 9.8 shows the exchange of a pion between a neutron and a proton in a nucleus.

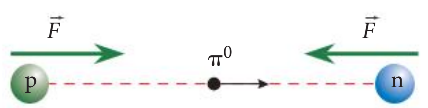


Figure 9.8 ▶ A proton and a neutron interact via the exchange of a neutral pion.

Following Yukawa's success, exchange particles for the weak nuclear force (also called the weak force) and gravity were also proposed.

The exchange particles for the weak nuclear force, W and Z bosons, have since been detected. The weak nuclear force is involved in nuclear decay. Like photons, W and Z bosons are their own antiparticles.

The proposed field particle for the gravitational field, the graviton, has not yet been detected. However, there are several major international research facilities, including the Australian International Gravitational Research Centre in Western Australia, trying to detect gravity waves or gravitons.

The different forces and their exchange particles are listed in Table 9.2.

AUSTRALIAN INTERNATIONAL GRAVITATIONAL RESEARCH CENTRE

Read about Australia's role in searching for the graviton.

Table 9.2 Forces and their exchange particles

Force	Exchange particle
Electromagnetic	Photon, γ
Gravitational	Graviton (not yet detected)
Weak	W and Z bosons, W^+, W^-, Z^0
Strong nuclear (not a fundamental force)	Pions, π^+, π^-, π^0

QUESTION SET 9.2

Remembering

1 List the six types of leptons. Give the name and symbol for each.
2 Make a table listing the properties of the two different types of hadrons.
3 What is an exchange particle? List the exchange particles for each force.

Understanding

4 Why are leptons considered to be elementary particles, but not mesons?
5 What is a baryon? Describe the features that distinguish it from a lepton or a meson.

Analysing

6 Follow the weblink to the Australian International Gravitational Research Centre. What evidence are they hoping to find to support the existence of gravitons? What are the main difficulties in finding this evidence?
7 Research the properties of each exchange particle shown in Table 9.2. Make a table showing the properties including mass, charge and spin of each.

Reflecting

8 Compare and contrast the field model for forces and the particle exchange model.

Particle reactions

The Standard Model, combined with a set of conservation laws, allows us to understand all the interactions and reactions that occur between particles. These interactions include those already described as mediated by the exchange particles listed in Table 9.2. They also include reactions such as the decay reactions you saw when you studied radioactivity and nuclear decay in Chapter 3 of *Nelson Physics Units 1 & 2 for the Australian Curriculum*.

First, we begin by introducing a useful tool to represent the reactions between particles. We use **reaction diagrams** to give us a simple visual representation of reactions.

Reaction diagrams

A reaction diagram provides a pictorial representation of a process.

The reaction diagrams that we will use here show the particles that interact during the reaction and the particles that result from the reaction. Those that exist before the reaction are shown to the left-hand side of the diagram and those that result from the reaction are shown to the right-hand side. Hence, the horizontal axis represents time, with the positive direction pointing to the right.

Particle reaction diagrams are similar to the reaction diagrams and reaction equations that you might have seen in chemistry.

Figure 9.9 shows a reaction diagram for the decay of a neutron into a proton, electron and electron antineutrino.

Each particle is represented by a line with an arrow. For particles, the direction of the arrow matches the direction of time. For antiparticles, the direction of the arrow is reversed. This allows us to distinguish between particles and antiparticles. Straight lines represent matter particles and wiggly lines represent massless exchange particles such as photons.

We are representing time on the horizontal axis. Other resources may represent time on the vertical axis – be sure to check which convention is being used!

The reaction itself is represented by the large dot or circle in the middle. The entire reaction is represented in the diagram as occurring at a single point in time, whereas in reality the reaction may include intermediate stages and the exchange of various field particles, which takes a finite time.

The arrows shown in Figure 9.9 do not represent actual paths through space. They do not represent the trajectories of the particles before and after the reaction.

The reaction shown in Figure 9.9 is one that you may recognise from *Nelson Physics Units 1 & 2 for the Australian Curriculum*, Chapter 3, in which you studied radioactive decay. This is the process of β^- decay, which occurs in nuclei.

As we saw in the Wow box about PET (page 258), matter particles annihilate when they meet their antimatter counterparts, resulting in the release of energy. In the case of PET, that energy takes the form of two photons. This is shown in Figure 9.10.

The reaction is represented by the circle and the particles are represented by lines going into or out of the reaction. Note that antiparticles, such as the electron antineutrino, have their arrows reversed.

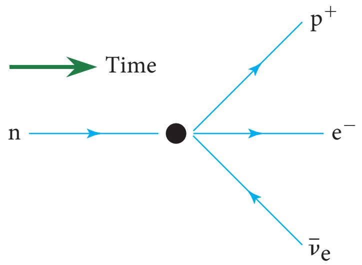


Figure 9.9 ▲ A simple reaction diagram showing a neutron decaying to a proton, an electron and an electron antineutrino.

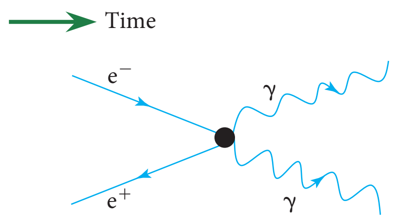
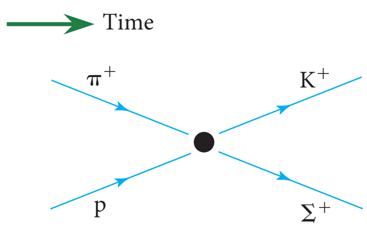


Figure 9.10 ▲ Positron–electron annihilation

WORKED EXAMPLE 9.1

Draw a reaction diagram for the process $\pi^+ + p \rightarrow K^+ + \Sigma^+$. (5 marks)

Answer



▲ Figure 9.11

Logic

We are given the reaction here: $\pi^+ + p \rightarrow K^+ + \Sigma^+$. 5 marks

All the particles are matter (not antimatter), so the arrows for all particles point to the right. The π^+ and p are on the left, because they are reacting, and the K^+ and Σ^+ are on the right, because they are the products of the reaction.

Try this yourself

Draw a reaction diagram showing the decay of a negatively charged tau to a tau neutrino, an electron and an anti-electron neutrino. (5 marks)

Reactions and conservation laws

Conservation principles are a unifying concept in physics. All systems and phenomena yet discovered obey these conservation principles. Energy, momentum, angular momentum, charge and spin are all conserved quantities.

The laws of conservation of energy, linear momentum, angular momentum, spin and electric charge provide us with a set of rules that all processes must obey.

You have already met the law of conservation of energy many times. It is central to understanding processes such as the photoelectric effect (Chapter 7), as well as the transmission of waves and mechanical collisions, which were discussed in Chapters 10 and 9 respectively of *Nelson Physics for the Australian Curriculum Units 1 & 2*. In Chapter 6 you saw that mass and energy are related via the equation $E = mc^2$.

You applied conservation of linear momentum when you studied collisions between objects in Chapter 9 of *Nelson Physics for the Australian Curriculum Units 1 & 2*.

In Chapter 8 we saw that angular momentum is a property of rotating or orbiting systems. It is conserved for all systems and processes, and is quantised in atomic and subatomic systems. Spin, which results in magnetic moments, is another quantity that is both conserved and quantised.

When you studied circuits in Chapter 5 of *Nelson Physics Units 1 & 2 for the Australian Curriculum*, you applied conservation of charge every time you used Kirchhoff's current law (junction rule), and conservation of energy when you applied the energy law (loop rule).

All of these conservation laws must be obeyed in any interaction between particles.

Particle physicists have discovered two further conservation laws. These are conservation of baryon number and conservation of lepton number.

Conservation of baryon number

We assign every particle a baryon number, as follows: $B = +1$ for all baryons, $B = -1$ for all antibaryons, and $B = 0$ for all other particles (see Table 9.1).

Experimental results show that whenever a baryon is created in a decay or nuclear reaction, an antibaryon is also created. Baryons and antibaryons can annihilate just as electrons and positrons do. However, in any other reaction in which a baryon or antibaryon is destroyed, another is created. From this we deduce that baryon number is a conserved quantity.

The law of conservation of baryon number states that whenever a reaction or decay occurs, the sum of the baryon numbers before the process must equal the sum of the baryon numbers after the process.

Baryon numbers for some particles are given in Table 9.1.

For example, consider the decay of a neutron resulting in β^- radiation:

$$n \rightarrow p + e^- + \bar{\nu}_e$$

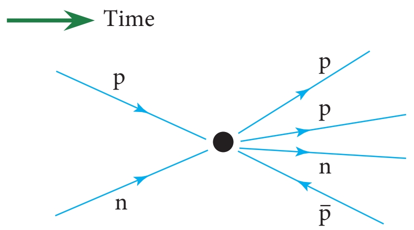
On the left-hand side of the equation, the neutron has baryon number $B = +1$, and on the right-hand side the proton has baryon number $B = +1$. The electron and antineutrino both have $B = 0$. Because baryon number is conserved, this is an allowed process.

In contrast, a decay of the proton to a positron and a neutral pion could satisfy conservation of energy, momentum and electric charge, but it would not satisfy the law of conservation of baryon number. Such a decay has never been observed.

WORKED EXAMPLE 9.2

Use the law of conservation of baryon number to determine whether each of the following reactions can occur.

a



(4 marks)

◀ **Figure 9.12**

b $p + n \rightarrow p + p + \bar{p}$ (4 marks)

Answers

a The reaction shown in Figure 9.12 is $p + n \rightarrow p + p + n + \bar{p}$

Before reaction		After reaction	
Particle	B	Particle	B
p	1	p	1
n	1	p	1
		n	1
		$\bar{p}$	−1

This reaction is allowed.

Logic

Before the reaction:
$B = (+1) + (+1) = +2$ 2 marks

After the reaction:
$B = (+1) + (+1) + (+1) + (-1) = +2$ 2 marks

As the total baryon number before and after the reaction is the same, this reaction is allowed by the law of conservation of baryon number.

b The reaction shown in Figure 9.12 is $p + n \rightarrow p + p + \bar{p}$

Before reaction		**After reaction**	
Particle	*B*	Particle	*B*
p	1	p	1
n	1	p	1
		$\bar{p}$	−1

This reaction is not allowed.

Before the reaction:
$B = (+1) + (+1) = +2$ 2 marks

After the reaction:
$B = (+1) + (+1) + (-1) = +1$ 1 mark

The total baryon number is reduced. Because the baryon number is not conserved, the reaction cannot occur. 1 mark

Try this yourself

Is the reaction shown in Figure 9.13 allowed? (4 marks)

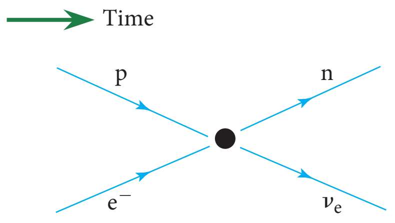


Figure 9.13

This reaction should be familiar from your study of nuclear processes in Unit 1 of Nelson Physics Units 1 & 2 for the Australian Curriculum.

In part a of Worked example 9.2, the mass on the right is larger than the mass on the left. Therefore, one might be tempted to claim that the reaction violates energy conservation. However, the reaction can indeed occur if the initial particles have sufficient kinetic energy to allow for the increase in the rest mass energy of the system. Remember that it is *total energy* that must always be conserved. Converting kinetic energy to rest mass energy does not violate conservation of energy.

Conservation of lepton number

There are three varieties of leptons: electrons (e), muons (μ) and tauons (τ), each with an associated neutrino (ν). There are three quantum numbers called lepton numbers, L_e, L_μ and L_τ associated with these particles. The electron and the electron neutrino have electron lepton number $L_e = +1$, and the antileptons e^+ and $\bar{\nu}_e$ have $L_e = -1$. All other particles have $L_e = 0$. Table 9.1 gives the lepton numbers of each type of lepton.

The law of conservation of electron lepton number states that in any reaction the sum of the electron lepton numbers before the process must equal the sum of the electron lepton numbers after the process. Similar conservation laws apply to muon and tau lepton numbers.

Lepton numbers for each type of lepton are given in Table 9.1.

Each of the lepton numbers L_e, L_μ, and L_τ is a conserved quantity.

Consider the decay of the neutron again:

$$n \rightarrow p + e^- + \bar{\nu}_e$$

Before the decay, the electron lepton number is $L_e = 0$; after the decay, it is $0 + 1 + (-1) = 0$. Therefore, electron lepton number is conserved.

Similarly, when a decay involves muons, the muon lepton number L_μ is conserved. The μ^- and the ν_μ are assigned a muon lepton number $L_\mu = +1$, and the antimuons μ^+ and $\bar{\nu}_\mu$ are assigned a muon lepton number $L_\mu = -1$. All other particles have $L_\mu = 0$. Tau lepton number, L_τ, is also conserved, with similar assignments made for the tau lepton, its neutrino, and their two antiparticles.

It has been found that lepton number is conserved in all reactions between particles; however, neutrinos have been observed to change from one type to another. This is called 'neutrino oscillation'.

WORKED EXAMPLE 9.3

Use the law of conservation of lepton numbers to determine whether this decay can occur:

$\mu^- \rightarrow e^- + \overline{\nu}_e + \nu_\mu$ (5 marks)

Answer

Before reaction			After reaction		
Particle	L_e	L_μ	Particle	L_e	L_μ
μ^-		1	e−	1	
			$\overline{\nu}_e$	−1	
			ν_μ		1

Both electron and tau lepton numbers are conserved and on this basis the decay is possible. — 1 mark

Logic

Before the decay:
$L_\mu = +1$, $L_e = 0$ — 2 marks

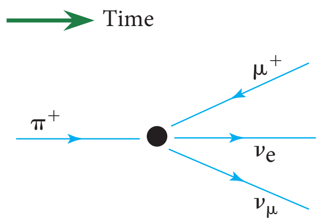
After the decay: — 2 marks

$L_\mu = (0) + (0) + (+1) = +1$

$L_e = (+1) + (-1) + 0 = 0$

Try this yourself

Is the decay shown in Figure 9.14 allowed? (5 marks)



◀ Figure 9.14

QUESTION SET 9.3

Remembering

1 List four quantities that must be conserved in all interactions.

2 Define 'baryon number'. List the baryon numbers for the following particles: n, p, $\overline{p}$, e^-.

Understanding

3 Draw a reaction diagram for each of the following reactions.

a $\Omega^- \rightarrow \Lambda^0 + K^-$

b $\pi^- \rightarrow \mu^- + \overline{\nu}_\mu$

Applying

4 Show that baryon number is conserved in the following reactions.

a $\pi^+ + p \rightarrow K^+ + \Sigma^+$

b $p + p \rightarrow p + p + \pi^0$

5 Show that charge is conserved in the following reactions.

a $\Omega^- \rightarrow \Lambda^0 + K^-$

b $p + e^- \rightarrow n + \nu_e$

6 Show that lepton number is conserved in the following reaction: $\mu^- \rightarrow \nu_\mu + e^- + \overline{\nu}_e$

Analysing

7 The following reaction is forbidden. Determine which conservation law is violated.

$p + \bar{p} \rightarrow \mu^+ + e^-$

8 a Draw a reaction diagram for this reaction: $n \rightarrow p + e^- + \bar{\nu}_e$

b Is the reaction allowed according to the conservation laws described in this section? If not, which law is violated?

9 Consider the following reaction: $\pi^- + \_ \rightarrow K^0 + \Lambda^0$. What properties must the missing particle have? Give an example of such a particle from Table 9.1.

Reflecting

10 In your earlier studies of radioactivity you saw the role of the neutrino as an energy carrier. Compare and contrast this description with your understanding of the role of neutrinos in particle reactions now.

The symmetry between the electric and magnetic fields allowed Maxwell to recognise them as aspects of a single phenomenon, as described in Chapter 5. The idea that all physical laws apply in any non-accelerating reference frame was central to the development of relativity, as described in Chapter 6.

Predicting new reactions using symmetry operations

The idea of **symmetry** in physics is a very important one. Physicists believe that the laws of physics will correctly describe physical phenomena under various transformations – such as reversal of direction in space or time, and rotations. These are called symmetries. We shall look at three symmetries in particular, because they give us a means of predicting possible particle reactions. These are **time-reversal symmetry**, **charge-reversal symmetry** and **crossing symmetry**.

If we apply any of these symmetries to an allowed reaction, then the resulting reaction is also allowed under the conservation laws we have discussed. This does not mean, however, that the reaction is *likely* to take place. In general, the probability of a new reaction occurring will be very different from the probability of the reaction from which it was derived. In some cases the new reaction does not in fact occur. The reason for this **symmetry breaking** is a matter of ongoing theoretical and experimental research.

Time-reversal symmetry

Recall that the horizontal direction on our reaction diagram represents time. If we flip a diagram around, swapping right to left, we reverse the direction of time. This means that the reaction occurs in the reverse order. If all the conservation laws previously described were obeyed by the original reaction, then this new process also obeys all the conservation laws. This process is called time reversal.

Figure 9.15 shows time reversal applied to electron capture. The first process shown (Figure 9.15(a)), in which an electron is captured by a proton, resulting in a neutron and an electron neutrino, occurs naturally in some large nuclei. If we apply time reversal to this process we get the reaction $n + \nu_e \rightarrow p + e$. This process also occurs. This is, in fact, the reaction by which the presence of neutrinos is detected at facilities such as Mount Kamioka where the Super-Kamiokande experiment runs (see Chapter 10).

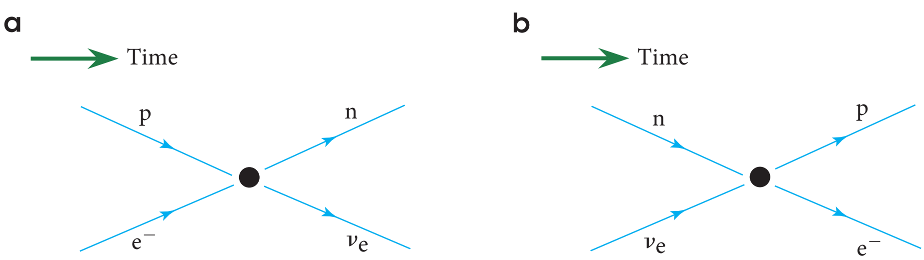
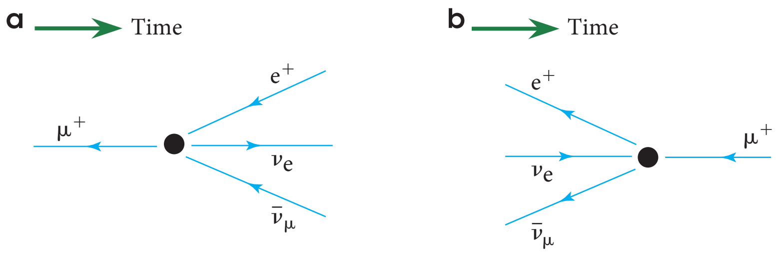


Figure 9.15 ▶ Time reversal applied to electron capture. a) An electron is captured by a proton. b) A neutron reacts with a neutrino to form a proton and an electron.

It may seem that you can apply time reversal to any reaction to generate a new possible reaction. However, energy and momentum conservation must be obeyed.

Consider a reaction in which the mass of the products is less than the mass of the reactants. If you apply time reversal, then the products have more mass than the reactants. The only way that this is possible is if the reacting particles have enough kinetic energy to convert into mass to create the new particles, or if there is some other source of energy available. This may be a particle that gives up energy but does not otherwise change in the reaction. An extreme example is electron–positron pair annihilation. You start with two massive particles, which annihilate to give photons with no mass. Applying time reversal to this process requires photons with enough energy and the correct momentum to interact such that an electron–positron pair is produced. The process of producing an electron–positron pair from high-energy photons is not simply the reverse of electron–positron annihilation. It involves a complex interaction with a nucleus to allow for conservation of momentum and energy.

A second aspect to consider is the *probability* of a reaction occurring. Consider a decay reaction such as that shown in Figure 9.16(a). This is the experimentally observed decay of a positive muon into a positron, an antimuon neutrino and an electron neutrino. If we apply time reversal to this process, as in Figure 9.16(b), we have a reaction between a positron, an antimuon neutrino and an electron neutrino. Although this is theoretically possible, in practice the probability of finding all three particles close enough together to react like this is negligible. This is particularly the case with short-lived exotic particles, such as muons, and neutrinos that interact only very weakly with matter. It is practically impossible to simultaneously collide three or more particles.



◀ **Figure 9.16** a) A positive muon decays into a positron, an antimuon neutrino and an electron neutrino. b) Time reversal applied to the process in part a. This reaction is not observed.

Applying time-reversal symmetry to a known allowed reaction generates a new reaction that is theoretically possible.
However this new reaction may not be experimentally realisable because of the requirement for conservation of energy and momentum, or because of the extremely low probability of simultaneously combining all the reacting particles.

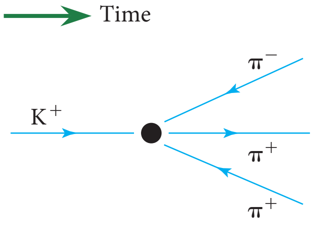
WORKED EXAMPLE 9.4

What is the result of applying time reversal to the β decay of a neutron? Do you think the time-reversed reaction is likely to occur? (4 marks)

Answer	Logic	
$n \rightarrow p + \beta^- + \bar{\nu}_e$	Apply conservation of baryon and lepton number to arrive at the correct reaction. Apply time reversal to obtain this equation.	3 marks
$p + \beta^- + \bar{\nu}_e \rightarrow n$ This reaction is unlikely.	The reaction requires both an electron and an antineutrino to interact simultaneously with a proton.	1 mark

Try this yourself

Apply time reversal to the reaction shown in Figure 9.17. Do you think the new reaction is likely? (2 marks)



◀**Figure 9.17** Reaction diagram for kaon decay

Charge-reversal symmetry

Charge-reversal symmetry says that if we reverse the charges on all particles in a reaction, this new reaction is also possible in that it doesn't violate any conservation laws. In fact, strictly speaking, charge reversal is used to refer to swapping all particles for their antiparticles, even those that are electrically neutral, such as neutrons and neutrinos.

Consider, for example, the decay of a muon:

$$\mu^- \rightarrow e^- + \overline{\nu}_e + \nu_\mu$$

Under charge reversal this becomes:

$$\mu^+ \rightarrow e^+ + \nu_e + \overline{\nu}_\mu$$

This decay also occurs.

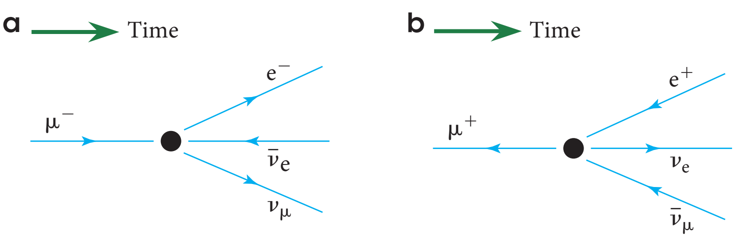
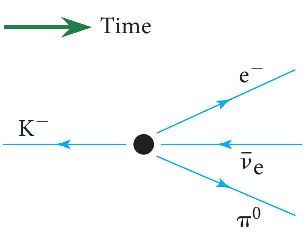


Figure 9.18 ▶ a) Decay of a negatively charged muon. b) Applying charge reversal gives us the reaction for the decay of a positively charged muon. Both these reactions are observed.

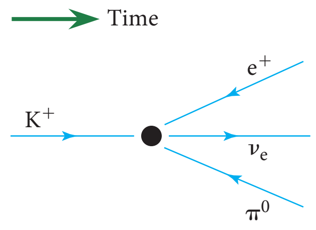
WORKED EXAMPLE 9.5

Apply charge-reversal symmetry to the decay of a negative kaon, shown in Figure 9.19, to draw a reaction diagram for the decay of a positive kaon. (3 marks)



◀**Figure 9.19** Decay of a negative kaon

Answer



▲ Figure 9.20

Logic

Reverse the charge on each particle, replacing it with its antiparticle. To show that the neutral pion is an antipion, reverse the direction of its arrow on the reaction diagram. 3 marks

Try this yourself

Apply charge reversal to the reaction for electron capture shown in Figure 9.15(a). Do you think this reaction is likely to be observed? (4 marks)

As with time reversal, applying charge reversal produces a reaction that does not violate conservation principles but may be so unlikely as to rarely or never occur naturally.

One reason for this is the extremely tiny amount of antimatter in the universe compared to the amount of normal matter.

Crossing symmetry

A third sort of symmetry that can be applied to predict new reactions is crossing symmetry.

In crossing symmetry we take one particle and cross it to the other side of the reaction, while converting it into its antiparticle.

WORKED EXAMPLE 9.6

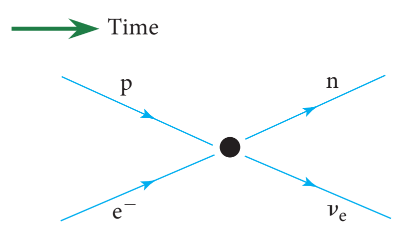
A proton captures an electron, to form a neutron and an electron neutrino.

a Draw the reaction diagram for this process. (2 marks)

b Apply crossing symmetry to the electron and draw the resulting reaction diagram. (3 marks)

Answers

a The reaction is $p + e^- \rightarrow n + \nu_e$.

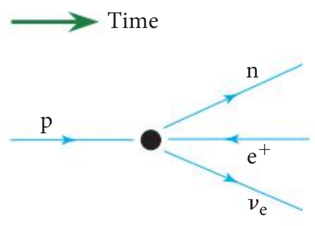


▲ Figure 9.21

Logic

Draw the reaction diagram (Figure 9.21), noting that there are no antiparticles, so all arrows point to the right. 2 marks

b The new reaction is $p \rightarrow n + \nu_e + e^+$



▲ Figure 9.22

We need to move the electron to the right-hand side and replace it with its antiparticle, the positron.

Note that we now have an antiparticle on the right-hand side, so its arrow must point left.

3 marks

Try this yourself

Write two other reactions by applying crossing symmetry to the reaction for electron capture. Do you think these reactions are likely to occur? (6 marks)

In Worked example 9.6 we can see that applying crossing symmetry to electron capture predicts the β^+ decay reaction. Both reactions are observed. However, other reactions predicted by applying crossing symmetry to electron capture are not observed.

Crossing symmetry, as with time-reversal and charge-reversal symmetry, predicts reactions that do not violate conservation principles. However, this does not mean that the reactions predicted actually occur.

Whether a reaction can occur, or how likely it is to occur, depends on the energy available, the mass differences of the particles, and conservation of other properties such as angular momentum. As with the other symmetries, reactions may also be theoretically possible, but have a negligible probability of ever occurring.

ACTIVITY 9.1

PIPE-CLEANER REACTION DIAGRAMS

Aim

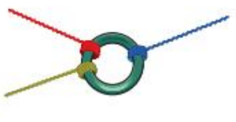
Your aim is to make some reaction diagrams using pipe cleaners. You can then apply the three different symmetry operations to generate new reactions.

You will need

Pipe cleaners in different colours, some thread and a large bead or ring.

What to do

1. The pipe cleaners represent particles. The bead or ring represents a reaction. Decide in advance what particles the different colours of pipe cleaners will represent. You can use the same colour but with a bit of thread wrapped around the pipe cleaner to represent the antiparticle.
2. Now apply the rules for drawing reaction diagrams to make a pipe-cleaner reaction diagram. Twist one end of each pipe cleaner into the bead or ring. Now bend the pipe cleaner so it is either coming into or going out of the reaction.
3. Sketch the diagram and write down what it represents.



▲ Figure 9.23 A pipe-cleaner reaction diagram

4 Now apply one of the symmetry operations – time reversal, charge reversal or crossing – to your pipe cleaners. You do this by bending the 'particles' to come into the reaction from different directions. Remember that when you do this, a particle is coming in from the opposite direction in time, so you need to convert it into its antiparticle.

See how many different reactions you can make.

5 Sketch and write down these new reactions.

QUESTION SET 9.4

Remembering

1 List the three types of symmetries that can be applied to generate new reactions.

2 When we apply crossing symmetry and move a particle from one side of a reaction to the other, what else must we do?

Understanding

3 Explain briefly the process of time reversal. For what reasons might a reaction predicted by applying time reversal not actually occur?

4 Why are many of the reactions predicted by charge reversal not observed, even though they are possible?

Applying

5 A negative sigma, Σ^-, commonly decays by the reaction $\Sigma^- \rightarrow n + \pi^-$. Apply charge-reversal symmetry to predict the decay process for a positive sigma.

6 The following decay reaction has been observed experimentally: $\tau^- \rightarrow e^- + \nu_\tau + \overline{\nu}_e$

Apply charge reversal to this reaction. Do you think this reaction is likely to occur? Explain your answer.

7 a Apply crossing symmetry to the electron in the reaction shown in question 6. Draw a reaction diagram for the resulting reaction. Do you think this reaction is likely to occur? Explain your answer.

b Apply time-reversal symmetry to the reaction you found in part a. Why is this reaction unlikely to occur? Give at least two reasons.

8 Apply time reversal to the reaction shown in Figure 9.10. Write the reaction and draw the reaction diagram. Explain why this reaction does not occur without other particles being involved.

Analysing

9 Consider the reaction $n \rightarrow p + e^- + \overline{\nu}_e$. By applying crossing symmetry, time reversal and then a second symmetry operation to this reaction, we can arrive at the process of β^+ decay. Draw the series of reaction diagrams for these symmetry operations, starting with the neutron decay reaction as given, and ending with β^+ decay.

Reflecting

10 Explain what the term 'symmetry' means in physics. Compare and contrast this with other meanings of symmetry you have come across.

CHAPTER SUMMARY

- Elementary particles are those that have no internal structure and cannot be divided into smaller components.
- The electron is one of a family of elementary particles called leptons.
- The photon is an elementary particle, with zero mass.
- Protons and neutrons are not elementary particles. They can decay into other particles, and the neutron has a magnetic field even though it is uncharged.
- For every particle there is a corresponding antiparticle with opposite charge but the same mass. When a particle meets its antiparticle they annihilate, producing energy in the form of photons. Positron–electron annihilation is an example of this, and is used in medical imaging.
- High-energy particles can be detected because they cause ionisation of materials they pass through. This occurs in cloud chambers and bubble chambers.
- Particle accelerators are used to create new particles by colliding high-energy charged particles with a target, such as another particle or a nucleus.
- More than 300 distinct types of particles have been observed. These can be classified as leptons and hadrons. Hadrons can be further classified as mesons and baryons.
- Particle masses are usually given in units of MeV/c^2. $1\,\text{MeV}/c^2 = 1.78 \times 10^{-30}\,\text{kg}$.
- Interactions between particles can be represented using reaction diagrams. These show the reacting particles as lines, moving in the direction of increasing time (horizontal axis) and interacting at an instant in time. Antiparticles are shown with their arrow pointing in the opposite direction.
- Reactions between particles obey a set of conservation laws, including conservation of energy, momentum, angular momentum, charge, baryon number and lepton number.
- Baryons have baryon number $B = +1$; antibaryons have $B = -1$. All other particles have $B = 0$.
- There are three lepton numbers, one for each type of lepton and its neutrino; L_e, L_μ and L_τ. The leptons have $L_x = 1$, the antileptons have $L_x = -1$, where $x = \text{e}$, μ or τ depending on the type of lepton.
- If a reaction is allowed under all conservation laws, then new reactions that are generated by applying time-reversal symmetry, charge-reversal symmetry or crossing symmetry are also allowed under the conservation laws.
- Time-reversal symmetry is applied by reversing the direction of time for a reaction.
- Charge-reversal symmetry is applied by reversing the charge on all particles involved in the reaction and replacing them with their antiparticles.
- Crossing symmetry is applied by moving one particle to the opposite side of a reaction and replacing it with its antiparticle.
- Reactions generated in this way are not necessarily likely to occur, and may not be possible, even though they are not forbidden by conservation laws.

CHAPTER GLOSSARY

antimatter matter composed of antiparticles, such as positrons, antiprotons and antineutrons

antiparticle each particle has an antiparticle that has the same mass but opposite charge and magnetic moment to the particle

baryon a heavy particle, with baryon number $B = +1$

baryon number, *B* quantum number associated with baryons; baryons have $B = +1$, antibaryons have $B = -1$

boson particle with integer spin, $s = 0, 1$. These particles do not obey the exclusion principle. Examples include the exchange particles

charge-reversal symmetry if all particles in an allowed reaction are replaced with their antiparticles (which have opposite charge), the new reaction is also allowed under known conservation laws

cloud chamber a chamber containing a supersaturated vapour through which high-energy particles pass, causing condensation of the vapour, and thus allowing their tracks to be visualised

collider a particle accelerator in which two particles are accelerated in opposite directions and collide

crossing symmetry if a particle in an allowed reaction is crossed to the other side of the reaction and replaced with its antiparticle, the new reaction is also allowed under known conservation principles provided enough energy is available

elementary particle a particle that does not have internal structure and cannot be broken into constituent particles. Leptons are considered to be elementary particles; mesons and baryons are not

exchange particles or field particles particles that mediate interactions, such as photons, which are the field particle of the electromagnetic field; also called gauge bosons

fermion particle with half-integer spin, $s = \frac{1}{2}, \frac{3}{2}, \ldots$, that obeys the exclusion principle; fermions include protons, neutrons and electrons

field particles *see* exchange particles

hadrons particles with large mass; the two types of hadrons are mesons and baryons

leptons a family of elementary particles: electrons, taus and muons and their neutrinos and all their antiparticles

lepton number, *L* the quantum number associated with leptons, there is one for each lepton type: $L_e = +1$ for electrons and electron neutrinos, $L_\mu = +1$ for muons and muon neutrinos and $L_\tau = +1$ for the tau and tau neutrino; the corresponding antiparticles have lepton number −1

linear accelerator or linac a device in which electric and magnetic fields are used to accelerate charged particles to high speeds in a straight line

meson a particle, generally (but not always) with mass between that of a lepton and a baryon

particle accelerator a device in which electric and magnetic fields are used to accelerate beams of particles to high speeds

positron the antiparticle of the electron, with charge $+e$ and mass m_e

reaction diagram diagram showing the interaction of particles. Particles are represented as lines with arrows, and interactions are represented as a circle; time is shown on the horizontal axis

strong nuclear force or nuclear force the force that acts between nucleons (protons and neutrons) to hold the nucleus together; it is mediated by pions

symmetry the invariance of physical laws under transformations such as translation, reflection and rotation in time and space

symmetry breaking a change in the behaviour of a physical system or the laws of physics that govern its behaviour when a symmetry operation such as a translation, reflection or rotation in time or space takes place

time-reversal symmetry when an allowed reaction is written such that it runs in the opposite direction in time; the new reaction is also allowed in that it does not break any of the known conservation laws

weak force the force necessary for β decay of nuclei. It is mediated by W and Z bosons. At very high energies it is unified with the electromagnetic force as part of the electroweak force

CHAPTER REVIEW QUESTIONS

Remembering

1 Name the three types of particle described in this chapter. Which are elementary particles?

2 Which force is mediated by the W and Z bosons? How are these exchange particles different from those that mediate the electromagnetic force?

3 State the baryon number, *B*, of:
   - **a** a photon. **b** an electron.
   - **c** a pion. **d** a neutron.

4 State the electron lepton number, L_e of:
   - **a** a photon. **b** an electron.
   - **c** a muon. **d** an electron neutrino.

5 When we apply time reversal to a reaction, which axis of the reaction diagram is reversed?

Understanding

6 Suppose a claim is made that the decay of the neutron is given by $n \rightarrow p + e$. What conservation laws are violated by this decay?

7 List five conservation principles which must be obeyed in particle reactions. Are these same conservation principles obeyed in a chemical reaction?

8 List the particles that normally exist in your body. Include exchange particles.

9 Show that the unit MeV c^{-2} is a unit of mass by writing and simplifying its dimensions.

10 Why do physicists believe that protons and neutrons are not elementary particles? Describe the evidence that suggests this.

Applying

11 a Show that 1 MeV $c^{-2} = 1.78 \times 10^{-30}$ kg.

b What is the mass in kg of the following particles?

i τ

ii Ω

12 An electron is accelerated to 60% the speed of light in a linear accelerator before being fed into a synchrotron ring of radius 550 m.

a Using classical electromagnetism, calculate:

i the accelerating potential used in the linear accelerator.

ii the magnetic field required in the synchrotron.

b Would your answers to parts a and b be larger or smaller if you had used the relativistic corrections?

13 An electron is circulating at 98% of the speed of light in a synchrotron with radius 1.5 km. What magnetic field is required to keep it in orbit? Use the relativistic correction.

14 a Show that baryon number is conserved in the following reaction: $\Sigma^0 \rightarrow \Lambda^0 + \gamma$.

b Why is a photon also produced in this reaction?

15 Is lepton number conserved in the following reactions?

a $n \rightarrow p + e^-$

b $\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$

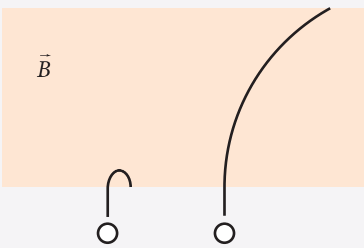
c $\pi^+ \rightarrow \mu^+ + \bar{\nu}_\mu$

16 Consider the following decay: $\pi^0 \rightarrow \mu^- + e^+ + \nu_\mu$. What conservation laws are violated by this decay?

17 A positron and a proton both enter a uniform magnetic field, as shown in Figure 9.24.

a Which of the two particles is the proton? Explain your answer.

b An electron enters the same field from the same direction and with the same initial velocity. Copy the diagram and draw the path of the electron.



▲ **Figure 9.24**

18 Draw a reaction diagram for the decay of a muon to a muon neutrino, an electron and an anti-electron neutrino.

19 Apply time reversal to the following decay reaction: $\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$. Is this reaction likely to occur? Why or why not?

20 Apply charge reversal to the reaction given in Question **19**. How likely do you think this reaction is to occur compared to the answer to Question **19**?

Analysing

21 An electron is accelerated to 0.999*c* in a linear accelerator. It is then fed into a synchrotron storage ring of radius 4.29 km. Calculate:

a the accelerating potential in the linear accelerator.

b the magnetic field in the synchrotron.

22 An electron circulates at a speed of 0.9999*c* in the 4.29 km radius ring at the large hadron collider. What frequency is the synchrotron radiation emitted by this electron?

23 Positron–electron pair annihilation most commonly produces a single virtual field particle that then decays to a new particle–antiparticle pair. For example, a positron and electron annihilate to create a high-energy photon that in turn decays to a muon–antimuon pair. Draw the two reaction diagrams that describe this process. Show that charge and lepton and baryon numbers are conserved.

Reflecting

24 Draw a spider diagram summarising what you have learnt in this chapter. Put in links to what you already knew from your study of radioactivity and nuclear physics. How has the material in this chapter extended your understanding of nuclear processes?

CHAPTER 10
THE STANDARD MODEL

Shutterstock.com/Bruce Rolff

By the end of this chapter you will have covered the following material.

Science Understanding

- The Standard Model is based on the premise that all matter in the universe is made up from elementary matter particles called quarks and leptons; quarks experience the strong force, leptons do not (ACSPH141)
- The Standard Model explains three of the four fundamental forces (strong, weak and electromagnetic forces) in terms of an exchange of force-carrying particles called gauge bosons; each force is mediated by a different type of gauge boson (ACSPH142)
- High-energy particle accelerators are used to test theories of particle physics including the Standard Model (ACSPH146)
- The Standard Model is used to describe the evolution of forces and the creation of matter in the Big Bang theory (ACSPH147)

Introduction

In the previous chapter we described the discovery of several new particles. We also looked at the evidence that indicates that familiar particles such as protons and neutrons are *not* elementary. This evidence includes the magnetic moment of uncharged particles such as neutrons and some other baryons.

The large number of new particles, more than 300, also suggested to physicists that not all of them could be elementary particles. This was partly because many of the particles seemed to come in 'families' or groups that had similar characteristics. For example, we have already seen that the proton and neutron have similar masses and the same spin. The group of three particles known as pions have almost identical masses and the same spin, but have different electric charges ($-1e$, 0 and $+1e$). As more and more particles were discovered, more patterns such as these began to emerge.

Other puzzles in particle physics included the nature of the field particles – photons, W and Z bosons, pions and the proposed gravitons. All but the last of these particles have now been detected. Photons are massless, but W and Z bosons and pions are not. What is the reason for the different masses of the field particles?

Physicists believed that there must be an underlying structure giving rise to the patterns observed in the properties and behaviour of all these particles. Just as the theory of evolution gave an underlying structure to the relationships between species, and the early theories of atoms gave an explanation for the organisation of the elements in the periodic table, physicists were searching for the underlying pattern to explain the 'particle zoo'. They wanted to know which of these particles were really elementary, and what the others were made of.

The answers to some of these questions came with the development of the **Standard Model** of particle physics. The first major development leading to the Standard Model was the **quark** theory.

QUARKS

Read about how the quark got its name.

At the same time that particle physicists were searching for the underlying structure of particles, astronomers and cosmologists were trying to understand how the large-scale structure of the universe came about. Astronomical evidence such as the red shift in the spectra of stars led to the development of the **Big Bang theory** – the theory that the universe began as a point-like **singularity** with enormous energy density. It may seem that the structure of the universe and the structure of a subatomic particle may not be closely related. However, according to the Big Bang theory, the universe began with the creation of elementary particles that condensed into the many particles we know of today, then into atoms, molecules and, finally, into the planets, stars and galaxies that give the universe its large-scale structure. Astronomical observations can only give us limited information about the history of the universe. To model what happened in the first stages of the evolution of the universe we need to use high-energy particle physics experiments.

Hence, an understanding of the underlying structure of the many particles in the particle zoo will also tell us about how the universe formed and then evolved to its current state. Once we understand the past evolution of our universe, we may be able to predict its future evolution.

The Standard Model of particle physics explains the underlying structure of particles and the forces via which they interact. It also describes the early stages of the evolution of the universe when particles first formed.

The evolution of the universe from a point-like singularity with enormous energy to the configuration of galaxies we know today is described by the Big Bang theory. The Big Bang theory relies on the Standard Model to describe and explain the first stages of the formation of the universe.

Matter and the Standard Model

As we have seen, in the 1950s there were a huge number of particles known. Broadly, these could be classified as particles with mass – leptons and hadrons – and those without mass, which include exchange particles such as photons.

The hadrons did not appear to be fundamental particles. So the question was, what were they made from?

Gell-Mann invented the term 'quark' for these elementary particles. The spelling was borrowed from a rhyme in James Joyce's Finnegans Wake*: 'Three quarks for Muster Mark'. Quark is also a type of cottage cheese.*

Quarks

In 1963, Murray Gell-Mann and George Zweig independently proposed a model for the substructure of hadrons. According to their model, all hadrons are composed of two or three elementary constituents called quarks.

This early quark model had three types of quarks, designated by the symbols u, d and s, which are given the arbitrary names up, down and strange. The various types of quarks are called **flavours**. Figure 10.1 is a pictorial representation of the quark compositions of several hadrons.

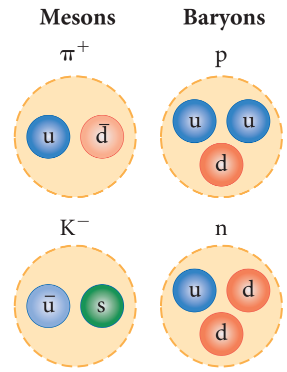


Figure 10.1 ▲ The quark composition of two mesons and two baryons

The compositions of all hadrons known when Gell-Mann and Zweig presented their model can be completely specified by three simple rules:

- **A meson consists of one quark and one antiquark.**
- **A baryon consists of three quarks.**
- **An antibaryon consists of three antiquarks.**

WORKED EXAMPLE 10.1

Which of these quark combinations is not a possible particle? (4 marks)

a $d\bar{d}$ **b** dd

c $\bar{u}\bar{d}d$ **d** $\bar{u}\bar{d}\bar{d}$

Answers	Logic	
a Possible	A meson consists of a quark–antiquark pair. A baryon consists of three quarks, and antibaryon of three antiquarks.	1 mark
b Not possible		1 mark
c Not possible	Hence, **a** is a meson, **b** is impossible, **c** is impossible and **d** is an antibaryon.	1 mark
d Possible		1 mark

Try this yourself

Which of these quark combinations is not a possible particle? (4 marks)

a uuu

b uu

c $u\bar{u}$

d $u\bar{u}\bar{u}$

Table 10.1 gives the properties of the quarks. The charmed c, top t, and bottom b quarks are described below.

Table 10.1 Properties of quarks

Name	Symbol	Spin	Charge	Baryon number	Strangeness	Charm	Bottomness	Topness
Up	u	$\frac{1}{2}$	$+\frac{2}{3}e$	$\frac{1}{3}$	0	0	0	0
Down	d	$\frac{1}{2}$	$-\frac{1}{3}e$	$\frac{1}{3}$	0	0	0	0
Strange	s	$\frac{1}{2}$	$-\frac{1}{3}e$	$\frac{1}{3}$	–1	0	0	0
Charmed	c	$\frac{1}{2}$	$+\frac{2}{3}e$	$\frac{1}{3}$	0	+1	0	0
Bottom	b	$\frac{1}{2}$	$\frac{1}{3}e$	$\frac{1}{3}$	0	0	+1	0
Top	t	$\frac{1}{2}$	$+\frac{2}{3}e$	$\frac{1}{3}$	0	0	0	+1

Each quark has an associated antiquark of opposite charge, baryon number, strangeness, charm, topness and bottomness.

Unlike any particles we have seen before, quarks carry a fractional electric charge. The u, d and s quarks have charges of $\frac{2e}{3}$, $-\frac{e}{3}$ and $-\frac{e}{3}$ respectively, where e is the electron charge ($e = 1.60 \times 10^{-19}$ C). Quarks have spin $\frac{1}{2}$, which means that they are classified as fermions (see Chapter 9), and have their own intrinsic magnetic moment and, hence, magnetic field. When quarks combine to form a particle, the charge of the particle is the arithmetic sum of the charges of its quarks. The spin of the particle is the sum of the spins of its quarks. However, spin is a vector, so we need to be more careful in how we add the spins. Both the magnitude and the direction of spin are quantised.

Recall from Chapter 4 that magnetic moment is a vector. It points in the direction of the north pole of an equivalent current loop. The magnitude of the magnetic moment of a particle depends on the spin.

In Table 10.1 you can see that quarks have baryon number $\frac{1}{3}$. Antiquarks have baryon number $-\frac{1}{3}$. Hence, a meson composed of one quark and one antiquark has baryon number $B = \frac{1}{3} + \left(-\frac{1}{3}\right) = 0$. A baryon, which is composed of three quarks, has $B = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1$, and an antibaryon composed of three antiquarks has $B = \left(-\frac{1}{3}\right) + \left(-\frac{1}{3}\right) + \left(-\frac{1}{3}\right) = -1$.

The law of conservation of baryon number, which we used in Chapter 9, is based on experimental observation. It is consistent with the model of quarks as elementary particles and this assignment of baryon numbers to quarks. Hence, the quark theory is able to explain the experimental observation of baryon number conservation.

Although no isolated quark has ever been experimentally observed, the quark model does describe the properties of mesons and baryons. Table 10.2 shows how mesons can be constructed from combinations of a quark and an antiquark.

Table 10.2 Quark composition of some mesons

		Antiquarks									
		$\bar{b}$		$\bar{c}$		$\bar{s}$		$\bar{d}$		$\bar{u}$	
Quarks	b	Υ	$(\bar{b}b)$	B_{c-}	$(\bar{c}b)$	$\bar{B}_s$	$(\bar{s}b)$	$\bar{B}_{d^0}$	$(\bar{d}b)$	B^-	$(\bar{u}b)$
	c	B_c^+	$(\bar{b}c)$	J/Ψ	$(\bar{c}c)$	D_s^+	$(\bar{s}c)$	D^+	$(\bar{d}c)$	D^0	$(\bar{u}c)$
	s	B_s^0	$(\bar{b}s)$	D_s-	$(\bar{c}s)$	ϕ	$(\bar{s}s)$	$\bar{K}_0$	$(\bar{d}s)$	K^-	$(\bar{u}s)$
	d	B_d^0	$(\bar{b}d)$	D^-	$(\bar{c}d)$	K^0	$(\bar{s}d)$	π^0	$(\bar{d}d)$	π^-	$(\bar{u}d)$
	u	B^+	$(\bar{b}u)$	$\bar{D}_0$	$(\bar{c}u)$	K^+	$(\bar{s}u)$	π^+	$(\bar{d}u)$	π^0	$(\bar{u}u)$

Consider the π^- meson as an example. We see in Table 10.2 that this meson has the quark composition $\bar{u}d$. The $\bar{u}$ quark has baryon number $-\frac{1}{3}$ and the d quark has baryon number $\frac{1}{3}$, giving a total of $B = 0$. The $\bar{u}$ quark has charge $-\frac{2e}{3}$ and the d quark has charge $-\frac{e}{3}$. This gives the π^- meson a total charge of $-1e$. Both the $\bar{u}$ and the d quarks have spin $\frac{1}{2}$. Spins are a little more complicated to add up because they represent magnetic moments, which are vectors. The magnetic moments can only line up in two ways: parallel or antiparallel (see Figure 10.2). When magnetic moments are parallel the total spin is $\frac{1}{2} + \frac{1}{2} = 1$; when they are antiparallel the total spin is $\frac{1}{2} - \frac{1}{2} = 0$. Looking back at Table 9.1 on page 264, we can see that a π^- meson has the properties $B = 0$, charge $= -1e$ and spin $= 0$. This is exactly what we expect from the combination of an anti-up quark and a down quark with their spins antiparallel. The alternative spin arrangement with spin $= 1$ corresponds to an excited (higher energy) state of the π^- meson.

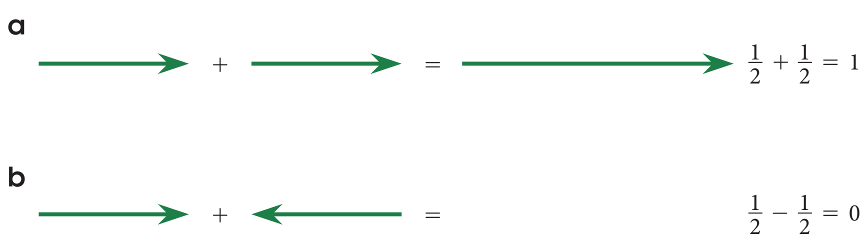


Figure 10.2
a) Parallel spins add to give spin = 1.
b) Antiparallel spins add to give spin = 0.

Baryons are composed of three quarks, and antibaryons of three antiquarks. Table 10.3 gives the quark composition of some baryons.

Note: Some baryons have the same quark composition, such as the p and Δ^+ and the n and Δ^0 baryons. In these cases, the Δ particles are considered to be excited states of the proton and neutron.

Table 10.3 Quark composition of several baryons

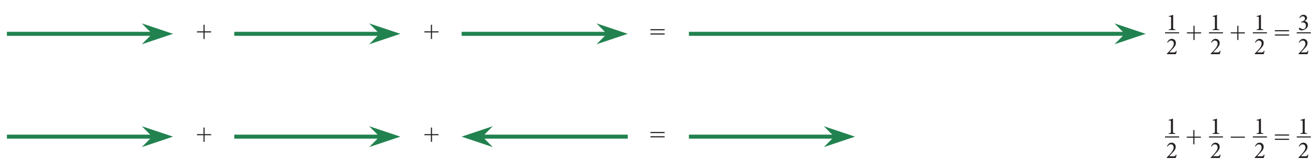
Particle	Quark composition
p	uud
n	udd
Λ^o	uds
Σ^+	uus
Σ^o	uds
Σ^-	dds
Δ^{++}	uuu
Δ^+	uud
Δ^o	udd
Δ^-	ddd
Ξ^o	uss
Ξ^-	dss
Ω^-	sss

Recall from Chapter 9 that particles with integer spin are bosons and particles with half-integer spin are fermions. The two types of particles behave differently when they interact and bind together. Fermions obey the exclusion principle; bosons do not.

WORKED EXAMPLE 10.2

Consider a baryon with quark composition uud. Show that it has the properties of a proton. (5 marks)

Answer	Logic	Marks
The uud combination has total charge $\frac{2}{3}e + \frac{2}{3}e - \frac{1}{3}e = 1e.$	An up quark has charge $\frac{2}{3}$ and a down quark has charge $-\frac{1}{3}$.	2 marks
Its baryon number is $B = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = 1$	Each quark has $B = \frac{1}{3}$.	1 mark
Spin $= \frac{3}{2}$ or $\frac{1}{2}$	The possible spin states are $\frac{3}{2}$ (all parallel) and $\frac{1}{2}$ (one antiparallel, two parallel), as shown in Figure 10.3.	2 marks



▲ **Figure 10.3** Possible spin arrangements for three quarks

Looking back at Table 9.1 in Chapter 9, we can see that the proton has $B = 1$, charge $= +1e$ and spin $= \frac{1}{2}$, in agreement with this quark composition with the lower energy arrangement of spins.

Try this yourself

Repeat the example above for a neutron with quark composition udd. (5 marks)

In the example above, the spin $\frac{3}{2}$ arrangement corresponds to a Δ^+ particle. As with the mesons, the parallel spin states are considered to be excited particle states. Hence, the Δ^+ particle is considered to be an excited state of the proton.

Although the original quark model was highly successful in classifying particles into families, some discrepancies occurred between its predictions and certain experimental decay rates. Consequently, several physicists proposed a fourth quark flavour in 1967. They argued that if four types of leptons existed (as was thought at the time), there should also be four flavours of quarks because of an underlying symmetry in nature. The fourth quark, designated c, was assigned a property called charm. A *charmed* quark has charge $+\frac{2e}{3}$, just as the up quark does, but its charm distinguishes it from the other three quarks. This introduces a new quantum number C, representing charm. The new quark has charm $C = +1$, its antiquark has charm of $C = -1$; all other quarks have $C = 0$.

At the time of the development of the quark theory only four types of lepton were known. We now know that there are six types, as described in Chapter 9.

Evidence that the charmed quark exists began to accumulate in the 1970s, when a series of heavy mesons with long lifetimes were discovered. The existence of these new mesons provided firm evidence for the fourth quark flavour.

Then in 1975, a fifth type of lepton, called the tau (t) lepton, was discovered. This led physicists to propose that more flavours of quarks might exist, on the basis of symmetry arguments similar to those leading to the proposal of the charmed quark. These proposals led to more elaborate quark models and the prediction of two new quarks, top (t) and bottom (b). To distinguish these quarks from the others, quantum numbers called *topness* and *bottomness* (with allowed values +1, 0, −1) were assigned to all quarks and antiquarks (see Table 10.1).

ACTIVITY 10.1

MAKING BARYONS

Aim

Your aim is to use 'quarks' to make some 'baryons' and identify their properties.

You will need

At least four beads or tokens in each of six different colours. Each different colour represents a different type of quark. You will also need a bag to hold the quarks, coloured pencils and paper.

What to do

1. Make a table listing the types of quarks and their properties, including charge, spin, strangeness, charm, topness and bottomness. Use Table 10.1 as a guide and add a column listing the colour you are using to represent each quark.
2. Place all the beads or tokens in the bag and randomly take 3 quarks from the bag.
3. These three quarks make up your new baryon. Using the table you have already drawn, work out the properties of your new baryon. List charge, spin, strangeness, charm, topness and bottomness.

   Note that spins can be in the same direction or opposite direction. Hence, there is not a single unique value for the spin of your new baryon.
4. Use your coloured pencils to record your new baryon and write down its properties.
5. Now see whether you can identify it in a list of known baryons, such as those given in Table 10.3 and Table 9.1 (page 264).
6. Put the quarks back in their bag and repeat the process.

What did you discover?

1. How many different baryons do you think you can make?
2. Did you make the same baryon more than once? Do you think all quarks are as common in nature as they are in your bag? Why or why not?

Matter particles in the Standard Model

The Standard Model of particle physics says that all matter is made up of two types of elementary particle. These are quarks and leptons.

As we saw in Chapter 9, there are six types of leptons, each with its own antiparticle. These are the electron, muon, tau, electron neutrino, muon neutrino and tau neutrino, and the antiparticle of each. Leptons do not combine to form other particles.

There are also six types or flavours of quarks, as we have just seen: up, down, top, bottom, strange and charmed. Each type of quark also has an antiquark. Quarks bind together to form other particles. They bind in groups of three to form baryons and in quark–antiquark pairs to form mesons.

All quarks and leptons have spin of $\frac{1}{2}$, and have an intrinsic magnetic moment. They are all fermions. This means that when they bind together into other particles, including baryons or large composite particles such as atoms, they obey the exclusion principle.

The second part of the Standard Model that we shall explore is the interaction of these particles. What force binds quarks together to form baryons and mesons? How is this force related to the force that holds nucleons together in the nucleus? We shall answer these questions in the next section.

Figure 10.4 ▶ The elementary matter particles described in the Standard Model

Quarks		Leptons	
u	$\bar{u}$	e^-	e^+
d	$\bar{d}$	ν_e	$\bar{\nu}_e$
s	$\bar{s}$	μ^-	μ^+
c	$\bar{c}$	ν_μ	$\bar{\nu}_\mu$
t	$\bar{t}$	τ^-	τ^+
b	$\bar{b}$	ν_τ	$\bar{\nu}_\tau$

QUESTION SET 10.1

Remembering

1 Name the three types of elementary particles.

2 a How many quarks do mesons consist of?

b How many quarks do baryons consist of?

Understanding

3 Explain why particle physics experiments are important in the development of cosmological theories such as the Big Bang theory.

4 Which of these quark combinations is a possible particle? Explain your answers.

a ud

b udd

c $\bar{u}$dd

d $\bar{u}\bar{d}\bar{d}$

e u$\bar{d}$

Applying

5 For each of these quark combinations, give the resulting particle's charge, strangeness and baryon number and all possible spins.

a $\bar{d}\bar{d}\bar{s}$

b bbt

c uus

d $\bar{u}\bar{u}\bar{s}$

6 Show that the properties of the quark combination $\bar{s}$u are consistent with the K^+ meson.

7 Show that the properties of the quark combination uds are consistent with the Λ^0. What other baryon could be composed of these quarks? Explain your answer.

▶

Analysing

8 A proton (uud) and an antiproton ($\bar{u}\bar{u}\bar{d}$) collide and a u and a $\bar{u}$ annihilate to create a total of eight quarks: ($u\,\bar{u}\,d\,\bar{d}\,d\,\bar{d}\,s\,\bar{s}$). A $\bar{\Sigma}^0$, a Λ^0 and what other particle are formed?

9 Analyse each of the following reactions in terms of constituent quarks and show that each type of quark is conserved.

  a $\pi^+ + p \rightarrow K^+ + \Sigma^+$

  b $K^- + p \rightarrow K^+ + K^o + \Omega^-$

10 Compare and contrast the aims of particle physics and cosmology. How has your understanding of the interaction between theories and researchers in different areas of physics changed?

Forces and the Standard Model

So far we have looked at four forces: the electromagnetic, gravitational, strong nuclear and weak nuclear forces. We said in the previous chapter that three of these, the electromagnetic, gravitational and weak nuclear, are fundamental forces. The fourth, the strong nuclear force, is not a fundamental force. Instead it is a consequence of another force, the strong force, which we shall introduce in this section.

The quark theory, combined with a theory called gauge theory, allows us to understand why the strong nuclear force is not a fundamental force. Rather, it arises from the interactions of quarks, as the nucleons (protons and neutrons) are composed of quarks. Gauge theory is highly mathematical and well beyond the scope of this course, so we shall look at some of the predictions of the theory only.

Gauge theory predicts the existence of particles called gauge bosons, which are massless particles via which massive particles interact.

We have already met one of these gauge bosons several times before: the photon.

The fundamental forces and field particles

In the Standard Model, particle interactions are described in terms of field particles or exchange particles. The emission of a field particle by one particle and its absorption by another particle manifests as a force between the two interacting particles.

Physicists now believe there are four fundamental forces which explain all interactions between particles.

In order of decreasing strength, they are the strong force, the electromagnetic force, the weak force and the gravitational force.

The strong force is responsible for the attractive force between nucleons. It has a very short range and is negligible for separation distances between nucleons greater than approximately 10^{-15} m (about the size of the nucleus). It acts between quarks, and the exchange or field particle is called the gluon. The name 'gluon' comes from the way that gluons 'glue' the nucleus together.

The electromagnetic force, which binds atoms and molecules together to form ordinary matter, has a strength of approximately 10^{-2} times that of the strong force. As we have seen, this long-range force decreases in magnitude as the inverse square of the separation between interacting particles. The field particles for the electromagnetic force are photons.

Do not confuse the strong force with the strong nuclear force. They are different forces, acting between different particles and mediated by different exchange particles.

The weak force is a short-range force that is responsible for beta decay and other decay processes. Its strength is only about 10^{-5} times that of the strong force. The weak force is mediated by field particles called *W* and *Z* bosons.

Finally, the gravitational force is a long-range force that has a strength of only about 10^{-39} times that of the strong force. Although this is the force that holds the planets, stars and galaxies together, its effect on elementary particles is negligible. The gravitational force is thought to be mediated by field particles called gravitons. The gravitational force is not part of the Standard Model, and the graviton has not yet been detected.

These interactions, their ranges, and their relative strengths are summarised in Table 10.4.

You can see that these forces differ in several respects, including their relative strength and the types of particle they act on.

Table 10.4 The fundamental forces and their exchange or field particles

Interactions	Relative strength	Range of force	Mediating field particle	Mass of field particle ($GeV\,c^{-2}$)
Strong	1	Short (≈1 fm)	Gluon	0
Electromagnetic	10^{-2}	∞	Photon	0
Weak	10^{-5}	Short (≈10^{-3} fm)	$W^{\pm}$, Z^0 bosons	80.4, 80.4, 91.2
Gravitational	10^{-39}	∞	Graviton	0

Relative strengths of forces

The relative strength of a force is a way of describing how strong the interaction is that it causes between particles. As an illustration, let's consider the forces between two electrons. We already know from Chapter 3 that electrons interact via the electromagnetic force. The electromagnetic force exerted by one electron on another is described by Coulomb's law:

$$F_{\text{ee, EM}} = \frac{k_e q_e^{\,2}}{r^2} = \frac{2.307 \times 10^{-28}}{r^2}\ \text{N m}^{-2}$$

But electrons have mass, and so as well as exerting an electrostatic force, they will also exert a *gravitational* force on each other:

$$F_{\text{ee, G}} = \frac{G m_e^{\,2}}{r^2} = \frac{5.535 \times 10^{-71}}{r^2}\ \text{N m}^{-2}$$

So we don't worry about the gravitational force between two electrons because the electromagnetic force is much stronger than the gravitational force.

As we can see from the equations, the exact magnitude of the forces depends on the properties of the particles that are interacting (charge for electromagnetism and mass for gravity), the separation between them, and factors relating to the force itself. These factors are represented by the constants k_e and G in the equations above. To compare the strengths of the fundamental forces alone, independently of properties such as mass and charge, physicists have introduced dimensionless quantities called coupling constants or relative strengths, given in Table 10.4. The exact definition of these constants depend on some quite complex theory.

The strong force acts between quarks and is mediated by gluons. The strong nuclear force acts between nucleons and is mediated by pions. Other sources may refer to the strong nuclear force as the strong force. Make sure you know which is meant.

The strong force and the strong nuclear force

The nuclear force was historically called the strong nuclear force. Once the quark theory was established, however, the phrase 'strong force' was reserved for the force between quarks. We shall follow this convention: the strong force is between quarks or particles built from quarks, and the nuclear force is between nucleons in a nucleus. The nuclear force is a result of the strong force. Other books and resources may refer to the nuclear force as the strong force.

The Standard Model, in which quarks interact via gluons, gives us a different way of looking at nuclear interactions. The proton–neutron interaction was described in Chapter 9 as being the result of an exchange of a pion between the two particles (see Figure 9.8, page 268).

Let's look at the same interaction from the viewpoint of the quark model, shown in Figure 10.5. Figure 10.5(a) shows the same process as Figure 9.8, although we now represent the process in three steps. First, the emission of a π^- particle by the neutron. The emission of a π^- changes the neutron to a proton. The absorption of the π^- by the proton results in it being converted into a neutron. The net effect is a force acting between the two nucleons. In Figure 10.5(b), the proton and neutron are represented by their quark constituents. Each quark in the neutron and proton is continuously emitting and absorbing gluons. The energy of a gluon can result in the creation of quark–antiquark pairs. This process is similar to the creation of electron–positron pairs in pair production. When the neutron and proton approach to within 1 fm of each other, these gluons and quarks can be exchanged between the two nucleons, and such exchanges produce the nuclear force.

Figure 10.5(b) shows one possibility for the process shown in Figure 10.5(a). A down quark in the neutron on the right emits a gluon. The energy of the gluon creates a $u\bar{u}$ pair. The u quark stays within the neutron (which has now changed to a proton), and the recoiling d quark and the $\bar{u}$ antiquark are transmitted to the proton. Here the $\bar{u}$ annihilates a u quark within the proton and the d is captured. The net effect is to change a u quark to a d quark, and the proton on the left has changed to a neutron.

As the d quark and $\bar{u}$ antiquark in Figure 10.5(b) transfer between the nucleons, the d and $\bar{u}$ exchange gluons with each other and hence are bound to each other by means of the strong force.

Looking back at Table 10.2, we see that this combination is a π^-, or Yukawa's field particle, as described in Chapter 9. Therefore, the quark model of interactions between nucleons is consistent with the pion-exchange model.

The pion was introduced in Chapter 9 as the field particle that mediates the strong nuclear force.

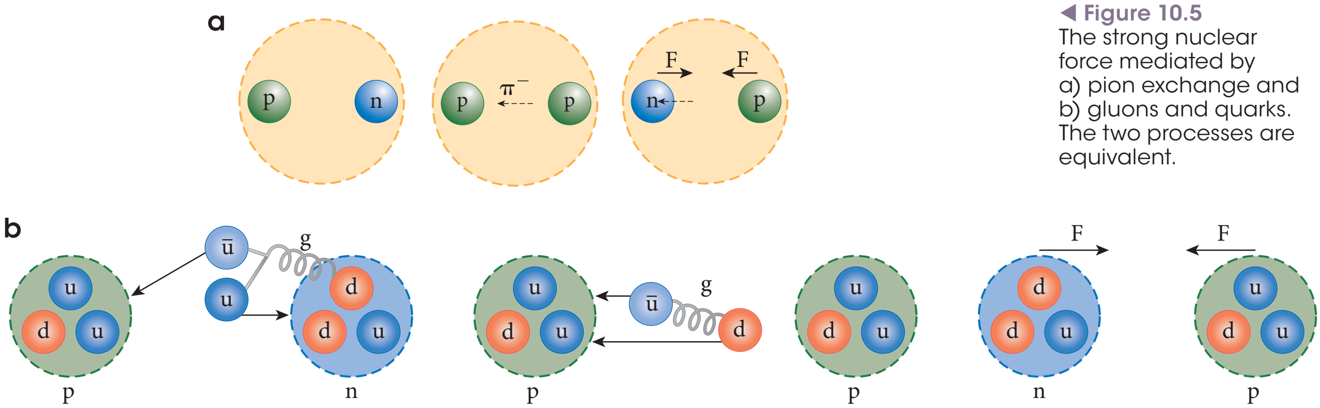


Figure 10.5 The strong nuclear force mediated by a) pion exchange and b) gluons and quarks. The two processes are equivalent.

The electroweak force

In 1979 Sheldon Glashow, Abdus Salam and Steven Weinberg won the Nobel Prize in Physics for developing a theory that unifies the electromagnetic and weak interactions. This **electroweak theory** postulates that the weak and electromagnetic interactions have the same strength when the particles involved have very high energies. Because of the mass difference between photons and the W and Z bosons, the electromagnetic and weak forces are quite distinct at low energies, but become similar at very high energies, when the rest energy is negligible relative to the total energy. This behaviour, as one goes from high to low energies, is called *symmetry breaking* because the forces are similar, or symmetric, at high energies but are very different at low energies.

The two interactions are viewed as different manifestations of a single unifying electroweak interaction. The theory makes many concrete predictions, but perhaps the most spectacular is the prediction of the masses of the W and Z particles at approximately 82 GeV c^{-2} and 93 GeV c^{-2}, respectively. These predictions are close to the masses in Table 10.4 determined by experiment.

The electroweak theory is an example of a unifying theory. Maxwell's theory of electromagnetism is another example of a unifying theory. Unifying or unified theories are those that explain more than one force within a single theoretical framework.

We shall see that this separation of the electromagnetic and weak forces was an important step in the evolution of the universe. The electroweak theory was important in the development of the Big Bang theory.

Mass and the Higgs boson

One of the questions raised by the Standard Model is why all the field particles except the W and Z bosons are massless. Or, put another way, why do the W and Z bosons have mass, and in fact what is the origin of the mass of all the other massive elementary particles?

To resolve this problem, a hypothetical particle called the Higgs boson, which provides a mechanism for breaking the electroweak symmetry, has been proposed. When particles interact via the Higg's field, they gain mass from this interaction. The field particle is the Higgs boson. The Standard Model modified to include the Higgs boson provides a logically consistent explanation of the massive nature of the W and Z bosons.

The Higgs boson was named for Peter Higgs, one of a group of physicists who proposed its existence in 1964. The existence of the Higgs boson was confirmed in March 2013 by CERN (the European Organization for Nuclear Research) after a particle with properties corresponding to those predicted for the Higgs boson was observed in July 2012, as described in the Scientific literacy box on the next page.

The Standard Model: a summary

Matter is made up of the elementary particles, quarks and leptons. Interactions between these particles are mediated by the third type of elementary particle, gauge bosons or field particles.

Particle physicists now believe that all matter is made up of two types of elementary particles: quarks and leptons. There appears to be a symmetry between the quarks and leptons, in that there are six types of each, along with their antiparticles.

Quarks combine to form mesons and hadrons. Mesons are made up of two quarks, as shown in Table 10.2. Baryons are made up of three quarks each (see Table 10.3).

In addition to these particles, there are also the exchange particles associated with the four fundamental forces, as shown in Table 10.4.

Although the details of the Standard Model are complex, its essential ingredients are summarised in Figure 10.6. This diagram shows that quarks participate in all the fundamental forces and that leptons participate in all except the strong force.

Note that the Standard Model *does not* include the gravitational force at present. However, we include gravity in Figure 10.6 because physicists hope to eventually incorporate this force into a **unified theory**.

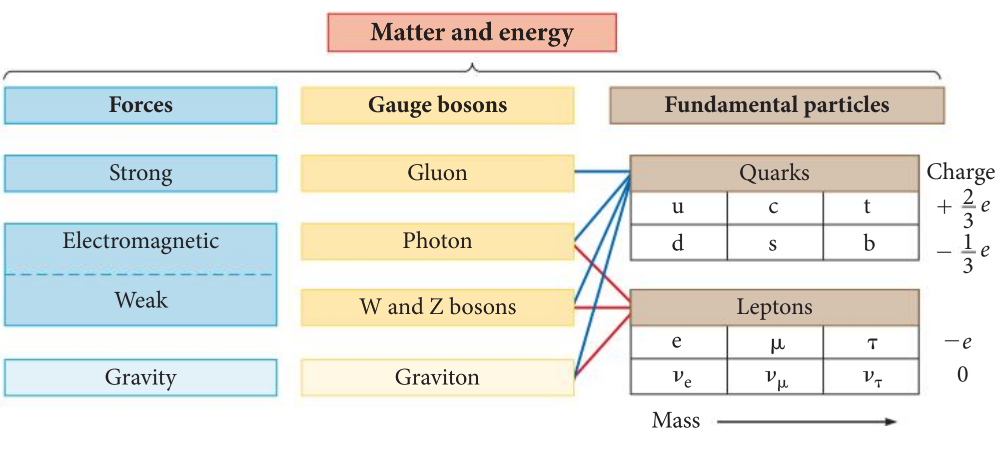


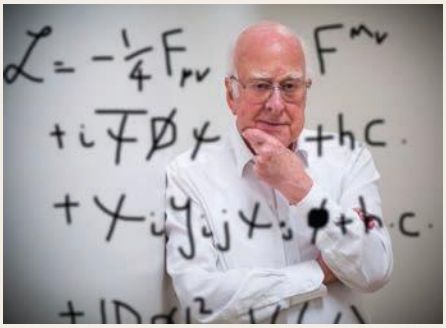
Figure 10.6 ▶ The Standard Model of particle physics

Scientific literacy: Discovery of the Higgs boson

As we have seen above, the mass of the Z and W bosons that mediate the electroweak force is something of a mystery in the Standard Model. Related to this, is the question as to how particle masses arise and what determines the mass of a given particle.

The theories behind this are extremely complicated mathematically and well beyond the scope of this text. They deal with fundamental symmetries in nature, and are referred to as *gauge theories*. They are based on quantum field theories, in which exchange particles are the mechanism by which particles interact and reactions occur.

In 1964, British physicist Peter Higgs proposed the existence of a field, now called the Higgs field, with which particles with mass interact. It is this field that gives them the property of mass. The field is composed of particles, now called Higgs bosons.



Alamy/James Davies

▲ **Figure 10.7** Peter Higgs

In the same year, and published in the same journal (*Physical Review Letters*), were two other papers that also described a mechanism by which particles acquired mass via interactions with a field. The first paper was by Francois Englert and Robert Brout, who were working in Brussels, and the other was by Gerald Guralnik, Carl Hagen and Tom Kibble, a collaboration of British (Kibble) and American (Guralnik and Hagen) physicists working in London.

Although the popular literature generally only mentions Higgs, all three papers are considered of equal importance in the development of the relevant theory by particle physicists. Many physicists, including Higgs himself, have argued that the name 'Higgs boson' is inappropriate because it does not acknowledge the contribution made by the others. However, given the widespread and long-term use of the name, it is unlikely to change now.

Remember that for a model to be considered a good one, it must *explain existing phenomena* and *produce testable predictions*. The model developed by Englert, Brout, Higgs, Guralnik, Hagen and Kibble explained why the weak force has a short range, while the electromagnetic force has an infinite range. It also explained why particles have mass. The model predicted the existence of a gauge boson – the Higgs boson – as well as its approximate mass. The detection of this particle is an important test of the theory.

However, detecting a Higgs boson is not a simple task. It requires massive energies of colliding particles to produce a reaction in which isolated Higgs bosons can be detected. It also requires massive computing power to analyse the data generated in such collision events. The Large Hadron Collider (LHC) was constructed, among other purposes, to search for the Higgs boson.

In July 2012, almost 50 years after its prediction, the detection of a particle with the properties and approximate mass of the Higgs boson was announced by physicists at CERN. It is interesting to note that the announcement *did not* claim that the particle found was the Higgs boson. Such a finding is of enormous importance in confirming the Standard Model and in understanding the nature of matter, forces and energy – the basic concepts in physics. This explains why researchers were hesitant to make such a significant claim without further careful analysis and measurement. Nonetheless, the announcement caused enormous excitement.

Eight months later, in March 2013, a press release from CERN stated that 'the measured interactions of the new particle with other particles strongly indicates that it is a Higgs boson'. The same press release (see weblink) quotes a spokesperson for one of the experiment teams as saying: 'The preliminary results with the full 2012 data set are magnificent and to me it is clear that we are dealing with a Higgs boson though we still have a long way to go to know what kind of Higgs boson it is.'

JULY 2012 PRESS RELEASE

Read the announcement of the discovery of a particle with the properties of the Higgs boson.

INTERVIEW WITH PETER HIGGS

You can read about Peter Higgs' reaction to the news of the discovery in this interview.

MARCH 2013 PRESS RELEASE

This is the official confirmation that the particle is the Higgs boson.

LATEST ON THE HIGGS BOSON FROM CERN

Catch up on the latest news on the Higgs boson here.

2013 NOBEL PRIZE FOR PHYSICS

Read more about the 2013 Nobel Prize for physics at this weblink.

In 2013 the Nobel Prize for physics was awarded to Peter Higgs and Francois Englert (Robert Brout died in May, 2011, and the Nobel Prize is not awarded posthumously). The citation by the Nobel society states that the award was given 'for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider'.

Although many physicists believe that Guralnik, Hagen and Kibble should also be recognised, the Nobel Prize is awarded to a maximum of three people.

Questions

1 Why is the detection of the Higgs boson of such importance to the Standard Model?

2 The Higgs boson detected has a mass between 125 GeV c^{-2} and 126 GeV c^{-2}. What is this in kilograms? Give your answer in the form (_ ± _) × 10^{-} kg.

3 Why could the Higgs boson not be detected with earlier particle physics experiments?

4 The Higgs boson was called 'the God particle' by Leon Lederman. Do some web searches and find out why. Comment on Higgs's response that it was embarrassing and 'the kind of misuse ... which I think might offend some people'. Why do you think he said this?

Limitations of the Standard Model

The Standard Model has helped us make sense of the huge number of particles. It has provided a sort of 'periodic table' to help us understand particle properties as arising from their underlying quark composition.

The Standard Model has also been very successful in explaining the origin and nature of three of the four fundamental forces – the strong, weak and electromagnetic. However, the Standard Model does not include the gravitational force.

The inclusion of the Higgs boson in the Standard Model gives a mechanism by which particles have mass. Particles acquire mass because of interactions with Higgs bosons. As we have seen, mass is the property that causes the gravitational field and hence the gravitational force. However, the Standard Model does not explain how the gravitational force is mediated. Many physicists believe that the gravitational field is mediated by massless exchange particles called gravitons. This is consistent with the way other fields are mediated by exchange particles (see Table 10.4), but the Standard Model does not include the graviton; therefore, the Standard Model is incomplete.

Gravity is the force that is most significant in the large-scale structure and evolution of the universe. Theories such as the Big Bang theory, discussed in the next section, base their predictions and explanations of cosmological phenomena largely on our current understanding of gravity. For example, the predicted existence of **dark energy** and **dark matter** is based on observed gravitational effects. The development of a more fundamental explanation for the gravitational field, or a **grand unified theory (GUT)** that explains all four forces, is therefore of great importance not only to particle physicists but also to cosmologists.

The structure and rate of expansion of the universe can be explained by the presence of dark matter and dark energy, as described in the next section. Dark matter neither radiates nor reflects energy – hence it is dark. Its presence is inferred from gravitational effects. Cosmologists believe that dark matter and dark energy consist of subatomic particles of a type as yet undiscovered. The Standard Model does not predict such a particle – there is a mismatch between current theories in cosmology and the Standard Model of particle physics.

Finally, the Standard Model in its current form has been found to be consistent with experimental results over the last 50 to 60 years, with one exception. The Standard Model predicts that the neutrino should be massless.

In 1998, at the Super-Kamiokande neutrino detector in Japan, it was discovered that neutrinos can change from one type to another. This implies that they have mass.

STANDARD MODEL AND NEUTRINO MASS

Find out more about the mass of neutrinos.

Astronomical observations including 'galactic lensing' imply a neutrino mass of about $0.2\,\text{eV}\,c^{-2}$ to $1.5\,\text{eV}\,c^{-2}$. The masses of the various types of neutrinos are not yet known, other than that they are very, very small. However, there is general agreement that neutrinos are not massless. Various theories have been put forward modifying the Standard Model to allow for non-zero neutrino masses.

The standard model is considered incomplete because it:
- does not explain gravity, and
- predicts that neutrinos are massless.

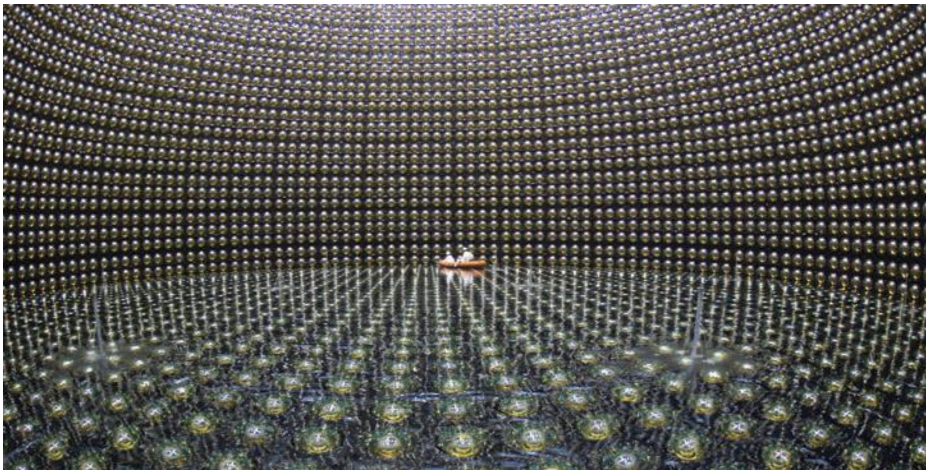
WOW

The Super-Kamiokande

The Kamioka Neutrino Observatory is located 1 km under Mount Kamioka in Japan. The Super-Kamiokande detector consists of a huge steel cylinder, 41 m tall and 39 m in diameter that contains 50 000 tonnes of highly purified water.

When neutrinos from sources such as the Sun or supernovae enter the water, they can interact with electrons or nuclei in the water molecules. These interactions produce particles that travel faster than the speed of light in water. (Note that this is much slower than the speed of light in vacuum.) This produces shock waves in the water, which are seen as Cherenkov radiation – a blue-green glow in a cone shape. This cone of light shows up as a ring, incident on detectors on the cylinder structure. The information from the detectors is analysed to determine the type of particle that interacted with the water. One of the reactions observed at the Kamioka observatory is $n + \nu_e \rightarrow p + e$, as described in the previous chapter.

The facility is located deep underground to shield the detector from more strongly interacting particles, allowing neutrinos, which interact only very weakly with matter, to be detected.



Kamioka Observatory, ICRR (Institute for Cosmic Ray Research), The University of Tokyo

▲ **Figure 10.8** The Super-Kamiokande neutrino detector

QUESTION SET 10.2

Remembering

1 Which forces are included in the Standard Model, and which are not?

2 Who won the Nobel Prize for physics in 2013, and for what?

Understanding

3 Why is the strong nuclear force not considered to be a fundamental force?

4 Describe two limitations of the Standard Model.

▶

Applying

5 Calculate the gravitational and electrostatic forces between two neutrons 1 fm (1×10^{-15} m) apart.

6 Calculate the gravitational and electrostatic forces acting between two protons 1 fm apart. What is the minimum force provided by the strong nuclear force between the protons?

Analysing

7 What is a field particle? Briefly describe how the existence of field particles explains 'action-at-a-distance'. What other models have we previously used to explain 'action-at-a-distance'? Explain how these models differ.

Reflecting

8 If you were in charge of the Nobel committee for physics, to whom would you have awarded a Nobel Prize for the prediction of the Higgs boson? Justify your answer.

9 Why do you think a unified theory that incorporates all four forces is important to both particle physicists and cosmologists? How has your understanding of the way in which different areas of physics interact changed? Draw a concept map to illustrate your understanding.

The Standard Model and cosmology

As you have seen, the Standard Model of particle physics deals with elementary particles, which are the smallest building blocks of matter. In contrast, cosmology deals with the large-scale structure of the universe, such as galaxies and nebulae. So it is not immediately obvious how the two are connected.

To understand the current structure of the universe, and predict its future, we need to understand the evolution of the universe. The currently accepted theory of how the universe began, the Big Bang theory, states that the universe began as an infinitesimally small and dense singularity about 14 billion years ago. This singularity exploded, forming all the elementary particles described by the Standard Model of particle physics. The interaction of these particles, which led to the formation of other particles – atoms, molecules, and eventually stars and galaxies – is governed by the fundamental forces.

Hence, to understand the evolution of the universe, we need to start with particle physics. But to understand where the particles came from in the first place, we need to use cosmology: the two are intimately connected.

The Big Bang theory is an expansion theory of the universe. It says that the universe was once much smaller and has been expanding since some time in the past. There are other expansion theories in cosmology, as well as steady state theories (which say that the universe is not expanding) and even oscillatory theories which say that the universe is oscillating in size and happens to be expanding at the moment. However, the Big Bang theory is currently the most widely accepted cosmological theory.

There are two main experimental observations that support the Big Bang theory. The first is the **redshift** of light from distant galaxies, described in the next section. If we assume that this means that galaxies are moving away from us and extrapolate backwards in time, then the universe must once have been much smaller. The second important piece of evidence is the existence of the **cosmic microwave background radiation**, which is discussed later.

In addition to the redshift and cosmic microwave background radiation, the different characteristic spectra of distant stars and the structure of distant galaxies compared to close ones implies that the universe is changing with time, and hence is not in a steady state. This argument is based on the idea that looking at distant galaxies is equivalent to looking backwards in time. This is because light takes a finite time to travel a given distance. If a star is 1000 light-years away, then the light from that star we observe today left the star 1000 years ago and what we are actually observing is what that star looked like 1000 years ago.

The expanding universe

In 1912 Vesto Melvin Slipher reported that most galaxies are receding from Earth at speeds of up to several million kilometres per hour. Slipher used Doppler shifts in spectral lines to measure the velocities of various galaxies.

Recall from Chapter 8 that at this time spectra had already been used to identify specific atomic species. Slipher observed spectra from distant galaxies that had the same line spacings as known species, but with the frequency of each line shifted by a fixed amount.

When the source of a wave is moving, the frequency of the wave detected by an observer is shifted. A higher frequency is detected if the source of the wave is travelling towards the observer. A lower frequency is detected if the source is travelling away from the detector. You may have noticed this effect when an ambulance or police car goes past you with its siren on. Initially, as the vehicle approaches, you hear a higher frequency. As the vehicle passes you, the frequency drops and you hear a lower frequency as it moves away.

The Doppler shift for sound is described in Nelson Physics Units 1 & 2 for the Australian Curriculum Chapter 10. Recall that when a sound source is moving away from you, you hear a lower frequency.

The spectral lines from distant galaxies that Slipher observed were consistently shifted to lower frequencies. The shift varied for different galaxies, but in each case the shift was towards a lower frequency (or longer wavelength). This is called a redshift, because red is at the lower frequency end of the spectrum. From these shifts Slipher deduced that the galaxies were moving away from us.

When light of wavelength λ_{emitted} is emitted from a source moving at velocity v away from an observer, the wavelength observed is given by:

$$\lambda_{\text{observed}} = \lambda_{\text{emitted}}\sqrt{\frac{1+\frac{v}{c}}{1-\frac{v}{c}}}$$

The shift in wavelength is therefore:

$$\Delta\lambda = \lambda_{\text{observed}} - \lambda_{\text{emitted}} = \lambda_{\text{emitted}}\left(\sqrt{\frac{1+\frac{v}{c}}{1-\frac{v}{c}}} - 1\right)$$

The factor $\sqrt{\frac{1+\frac{v}{c}}{1-\frac{v}{c}}}$ is also sometimes written as:

You have met the ratio β in Chapter 6 when you studied relativity.

$$\sqrt{\frac{1+\frac{v}{c}}{1-\frac{v}{c}}} = \sqrt{\frac{1+\beta}{1-\beta}} \text{ where } \beta = \frac{v}{c}$$

WORKED EXAMPLE 10.3

Calculate the observed wavelength of the hydrogen spectral line that has wavelength 434 nm, from a galaxy receding from us at 0.25c. Is this light still in the visible region of the spectrum? (4 marks)

Answer	Logic	
$\lambda_{\text{observed}} = \lambda_{\text{emitted}}\sqrt{\frac{1+\frac{v}{c}}{1-\frac{v}{c}}}$	Relate the observed wavelength to the emitted wavelength.	1 mark
$\lambda_{\text{observed}} = 434\,\text{nm}\sqrt{\frac{1+\frac{0.25c}{c}}{1-\frac{0.25c}{c}}}$	Substitute wavelength and speed in units of c.	1 mark
$\lambda_{\text{observed}} = 434\,\text{nm}\sqrt{\frac{1.25}{0.75}} = 560\,\text{nm}$	Calculate the final value.	1 mark
The observed wavelength is in the yellow-orange region of the spectrum.		1 mark

Try this yourself

If this same spectral line was observed with a wavelength of 650 nm, what speed would the source galaxy need to have? (5 marks)

Subsequent observations by other astronomers consistently showed that the spectra of stars in all observed galaxies were redshifted. It appeared that everything was moving away from us!

The atomic spectra of light from distant galaxies is redshifted such that each spectral line has a longer wavelength than the same line for that atomic species on Earth. This redshift is a result of the Doppler shift in the frequency of the light due to the source of the light moving away from us. This observation implies that distant galaxies are moving away from us.

Hubble's law is empirical. It is a mathematical description of experimental observations. It is not based on any underlying theory. NASA's current best estimate of Hubble's constant has an uncertainty of 3%.

Not only did observation show that all galaxies are moving away from us, it also seems that the more distant the galaxy, the greater the redshift. This implies that the further away a galaxy is, the faster it is moving.

In the late 1920s, Edwin P. Hubble put forward the theory that the entire universe is expanding. Observations showed that the speeds at which galaxies are receding from Earth increase in direct proportion to their distance from us.

This is described mathematically by Hubble's law:

$$v = HR$$

where v is the speed of recession, R is the distance from us and H, is the Hubble constant, $H \approx 22 \times 10^{-3}\,\text{m s}^{-1}\,\text{ly}^{-1}$. (1 ly = 1 light-year = 9.5×10^{15} m)

This is possible if the entire universe is expanding, with all points getting further away from each other. So the universe must have started out much smaller.

The current theory in cosmology, the Big Bang theory, says that it started with an explosion of all matter from a single point, a singularity.

The Big Bang theory and the evolution of the universe

Recall that in the Standard Model of particle physics the electromagnetic and weak forces are unified into the electroweak force at high energies.

According to the Big Bang theory, the universe erupted from an infinitely dense singularity about 14 billion years ago.

For the first few moments after the Big Bang the universe was at such extremely high energy (temperature) that all matter was contained in a quark–gluon plasma.

It is thought that in these first moments all four fundamental forces were unified; the strong, electroweak and gravitational forces were joined to form a single force.

The evolution of the four fundamental forces from the Big Bang to the present is shown in Figure 10.9.

Then, about 10^{-35} s after the Big Bang, when the temperature had dropped to about 10^{29} K, symmetry breaking occurred for gravity. This symmetry breaking meant that the properties of the gravitational force became distinct from those of the other forces. At this time the strong and electroweak forces remained unified.

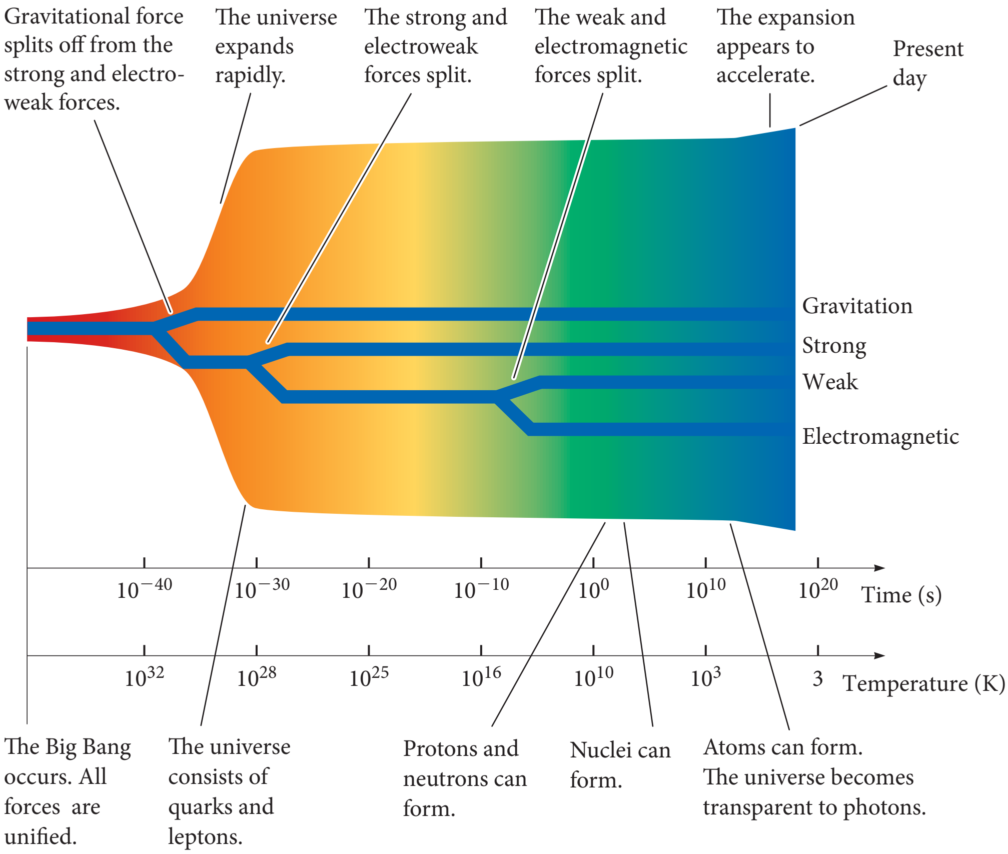
It was a period when particle energies were so great that very massive particles as well as quarks, leptons and their antiparticles existed. For some reason not yet understood, far more matter than antimatter particles formed. The amount of matter far exceeds the amount of anti-matter in our universe to this day. This is not explained by the Standard Model of particle physics, or by the Big Bang theory.

Then the universe rapidly expanded and cooled, and the strong and electroweak forces became distinct.

The universe continued to cool and, at approximately 10^{-10} s after the Big Bang, the electroweak force split into the weak force and the electromagnetic force.

After a few minutes, protons and neutrons condensed out of the plasma. For half an hour, the universe underwent thermonuclear detonation, exploding as a hydrogen bomb and producing most of the helium nuclei that now exist. The universe continued to expand, and its temperature dropped.

Until about 700 000 years after the Big Bang, the universe was dominated by radiation. High-energy radiation prevented matter from forming neutral hydrogen atoms because



◀ **Figure 10.9** The Big Bang and the evolution of the fundamental forces

collisions would instantly ionise any atoms formed. Photons were continuously scattered from the vast numbers of free electrons, resulting in a universe that was opaque to radiation.

This is an important point in the history of the universe. Astronomers claim that when you look through a telescope into space, you are looking into the past. This is because light takes a finite time to travel through space. Hence, the further away you look into space, the further back in time you are seeing because the light has taken a long time to reach you.

But no matter how far away you look with a telescope, you cannot see further back than the time when the universe was 700 000 years old. This is because the universe was opaque to light before then, so no light can reach us from that time.

Astronomers like to say that they are looking backwards in time when they look at light from a distant star, because the light has travelled many years to get to them.

Astronomical observations cannot give us information about the universe when it was less than 700 000 years old. We rely on particle physics to tell us about the state of the universe before this time.

The 700 000-year-old universe had expanded and cooled to approximately 3000 K and protons could now bind to electrons to form neutral hydrogen atoms. As we know from Chapter 8, atoms have quantised energy levels. Now that atoms existed as the main state of matter, far more wavelengths of radiation were *not* absorbed by atoms than were absorbed, and the universe suddenly became transparent to photons.

Radiation no longer dominated the universe, and clumps of neutral matter steadily grew: first atoms, then molecules, gas clouds, stars and finally galaxies.

The universe has continued to expand and cool since the Big Bang.

One prediction of the Big Bang theory is that residual radiation from the Big Bang should currently be observable. This was predicted in 1948 by Ralph Alpher and Robert Herman, although their work was largely ignored at the time.

This radiation comes from the time when the universe was at a temperature of about 3000 K and photons were first able to pass through the matter in the universe. We would expect a

The universe became transparent to photons when atoms formed, because the quantised energy levels of atoms mean they can absorb only fixed, discrete frequencies. See Chapter 8 for more details of atomic spectra.

radiation spectrum from the early universe to look like that of a 3000 K black body curve.

The universe has since expanded and cooled and the predicted radiation spectrum now corresponds to a 3 K black body curve. This background radiation should have a peak intensity at a wavelength of a few millimetres, which is in the microwave region of the spectrum. Hence this residual radiation is known as 'cosmic microwave background radiation'.

Cosmic microwave background radiation was first observed in 1965 by Arno Penzias and Robert Wilson. The measurement of the radiation was important in establishing the Big Bang theory as the accepted theory in cosmology.

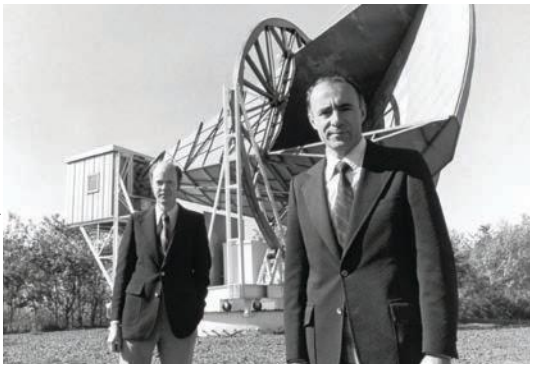
WOW

3K Cosmic microwave background radiation, not pigeons

In 1965, Arno A. Penzias and Robert W. Wilson of Bell Laboratories were testing a sensitive microwave receiver for satellite communications. A faint background hiss was interfering with their experiments. They noticed that the hiss was the same regardless of the direction they pointed the antenna. They cooled the microwave detector and went outside to chase a flock of pigeons out of the horn-shaped antenna, but the signal remained.

In a casual conversation with colleagues they realised that what they had taken to be interference caused by pigeons was actually the residual radiation from the Big Bang.

Subsequent measurements confirmed that the radiation they measured corresponded to that of a back body at 2.7 K. They were awarded the Nobel Prize for their discovery in 1978.



Science Photo Library/Emilio Segre Visual Archives/American Institute of Physics

▲ **Figure 10.10** Wilson and Penzias with the antenna first (accidentally) used to detect cosmic background microwave radiation

Dark matter, dark energy and the future of the universe

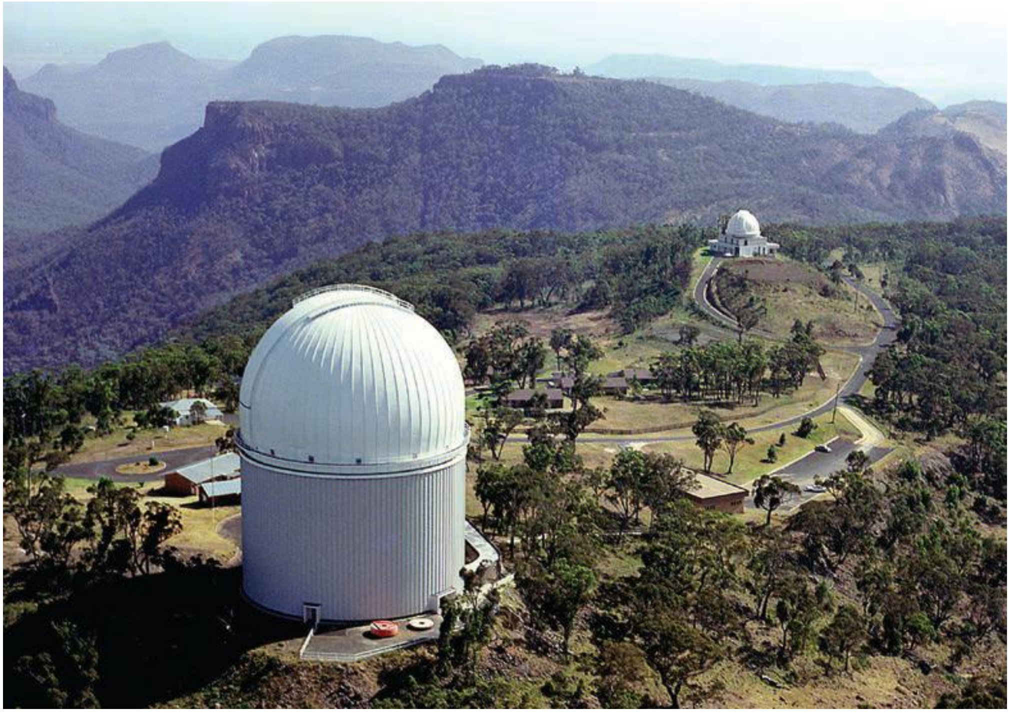
The Big Bang theory gives us a model for the beginning of the universe. But what will happen to the universe in the future? Will it continue to expand and cool forever, or will it at some time begin to contract again? What happens will depend on the amount of mass in the universe, and on the rate at which the universe is expanding.

If there is enough mass, then eventually the gravitational force will cause the expansion to slow down and stop, and then reverse it. Gravity will pull all the galaxies, stars and planets back together into a 'big crunch'. If there is not enough mass, or if the rate of expansion is too great, then the universe will continue to expand and cool forever.

In any gas, such as a star, there are two competing effects. Gravity tries to pull the particles together, while pressure due to thermal energy acts to push the particles apart. Hence, the hotter the gas, the more it expands. The denser the gas, the more gravity compresses it.

In 1998 observations of the apparent brightness and the redshift of supernovae were used to measure their distance and speed of recession from the Earth. These observations led astronomers to the conclusion that the expansion rate is increasing.

To explain this acceleration, physicists proposed the existence of dark energy, which is energy possessed by the vacuum of space. The theory is that in the early universe, gravity dominated over the dark energy. As the universe expanded and the gravitational force between galaxies became smaller because of the great distances between them, the dark energy became comparatively more important. The dark energy results in an effective repulsive force that causes the expansion rate to increase. This is a similar mechanism to that which causes hot gases to expand.



◀ **Figure 10.11** The Anglo-Australian Telescope at the Siding Spring Observatory, NSW

Other observations carried out between 2006 and 2011 at the Anglo-Australian Telescope at the Siding Spring Observatory have provided strong evidence that the expansion rate is indeed increasing. The new results imply that dark energy is likely to account for about 72% of the energy in the universe!

The dark energy may be acting to increase the rate of expansion of the universe. If the balance between dark energy and gravity favours dark energy, then the universe will continue to expand forever. Based on the observable mass of the universe, dark energy should win out. However, based on observable gravitational effects, there appears to more than just visible matter in the universe.

'Dark matter' was first postulated by Jan Oort in 1932. Recall from your study of gravity and orbital motion (Chapter 2) that the orbital velocity of a planet depends on the mass of the star about which it orbits. Similarly, the orbital velocity of stars about the galactic centre depends on the mass in the galaxy. Oort observed that the measured velocity of stars in the Milky Way did not correspond to that predicted by the observable mass in the Milky Way. The observable mass is that calculated from the objects that we can see – mainly stars.

Observations by other astronomers also provided evidence that the total mass in galaxies, including our own, is far greater than that due to visible matter. The 'missing matter' needed to explain these observations was given the name 'dark matter' because we cannot see it.

Cosmologists now believe that most of the matter in the universe is this mysterious dark matter. If there is enough dark matter then gravity will eventually win over dark energy. If this happens the universe will contract and end in a big crunch.

The nature of both dark energy and dark matter remains a mystery that many physicists hope to solve.

Limitations of the Big Bang theory

The Big Bang theory is the model of the evolution of the universe that is most widely accepted by cosmologists today. However, as with all the models and theories we have examined so far, it has limitations.

One of the predictions based on the combination of particle theory and the Big Bang theory is the creation of magnetic monopoles (isolated magnets with only one magnetic pole). However, no magnetic monopoles have ever been observed.

Case study

Professor Elisabetta Barberio

Elisabetta Barberio is a Professor in the School of Physics at the University of Melbourne. She is considered the world expert in measurements of b-quark (bottom quark) properties. She has contributed to two major experiments that confirmed key pieces of Standard Model theory, which led to two Nobel Prizes in 1999 and 2008.

Professor Barberio is a chief investigator at the ARC Centre of Excellence for Particle Physics at the Terascale (CoEPP) where she leads a team of experimental particle-physics researchers. CoEPP is a Centre of Excellence funded by the Australian Government through the Australian Research Council. CoEPP provides the necessary funding so that Professor Barberio and her colleagues can go to work at amazing places like CERN (European Organization for Particle Physics) in Switzerland. She and her team work on the ATLAS experiment at CERN.

Courtesy University of Melbourne/ Casamento Photography

▲ **Figure 10.12**
Professor Elisabetta Barberio

Professor Barberio became a physicist because she has always been very curious about the world around us, and trying to understand natural phenomena. She chose physics because she wanted the opportunity to contribute to our understanding of the universe. Physics tries to address fundamental questions, and often can give an answer. Perhaps not the final answer, but a step forwards. She enjoys physics because physicists need to be inventive, to have ideas and to have creativity.

Professor Barberio studied physics as an undergraduate at the University of Bologna in Italy. She then went on to a Masters degree and a PhD in Germany.

Since then she has led many international research teams in Europe and Japan. She and her teams have played a crucial role in data analysis for the OPAL experiment at the Large Electron Positron Collider. The OPAL detector made measurements of W and Z bosons. Recall that these are the exchange particles for the electroweak force. Precision measurements made with the OPAL experiment have confirmed the theory describing fundamental particle behaviour (the Standard Model) to an extraordinary degree.

Professor Barberio's research continues to look at some of the fundamental questions in science. Her current group had an important role in the discovery of the Higgs boson at the Large Hadron Collider (LHC) (see 'Scientific literacy: Discovery of the Higgs boson' on page 295). The group works on the ATLAS experiment at the LHC and are foundation members of the experiment.

The CoEPP was also a major sponsor and organiser of ICHEP2012, the international physics conference at which CERN announced its Higgs boson discovery via two-way webcast with Melbourne.

The acronym ATLAS stands for A Toroidal Large hadron collider ApparatuS. ATLAS is 46 metres long, 25 metres high, 25 metres wide and weighs 7000 tonnes. It has six different detecting subsystems and a huge magnet system that bends the paths of charged particles. It is housed at CERN and its goal is to explore the fundamental nature of matter and the basic forces that shape our universe.

ATLAS is one of the largest collaborative efforts ever attempted in the physical sciences. There are 3000 physicists participating from more than 150 universities and laboratories in 38 countries. CoEPP experimentalists take turns in going to CERN and working on the detector.

As we have seen, the Standard Model of Particle Physics describes our current understanding of matter and forces. The fundamental matter and force particles form a 'periodic table' of the building blocks of the universe. These observed patterns or symmetries have been key to the Standard Model and many of its predictions. However, the symmetries are not perfect. Some particles, such as the photon, are massless, while others, such as the electron, are massive. Particle masses vary over many orders of magnitude.

1999 NOBEL PRIZE IN PHYSICS

This Nobel Prize was awarded for elucidating the electroweak interactions.

2008 NOBEL PRIZE IN PHYSICS

This Nobel Prize was awarded for the discovery of the origin of a broken symmetry.

ATLAS

ATLAS investigates a wide range of physics. Read more about ATLAS here.

A key goal of the CoEPP is to look at the origin of mass, and the symmetry breaking that generates the variety of masses of the elementary particles. The simplest of these, the Higgs mechanism, demands the existence of the now famous Higgs boson.

Through their involvement in the giant ATLAS experiment at the LHC, Professor Barberio's team participated in the discovery of the Higgs boson. This involved both gathering the data from ATLAS and interpreting that information in light of our current theories.

In parallel with the Standard Model, a 'standard model' of cosmology has emerged, leading to the astounding conclusion that only 4% to 5% of the universe is made of the 'normal' matter so well described by the Standard Model. Approximately 23% of the universe is made of 'dark matter', which interacts via gravity, but otherwise remains unknown to us. Even stranger is the remaining 72% that results in the accelerating expansion of the universe – dark energy.

The current Standard Model does not include dark matter or dark energy. However, a feature of some extensions to the Standard Model is the prediction of dark matter particles that would be discoverable at the LHC. Professor Barberio's research program aims to draw together the Standard Model and cosmology, leading to a genuine understanding of the nature of dark matter. They are hoping to be able to detect the presence of these theorised dark matter particles in experiments using the LHC at CERN.

Professor Barberio summarises the main questions at the frontiers of particle physics:

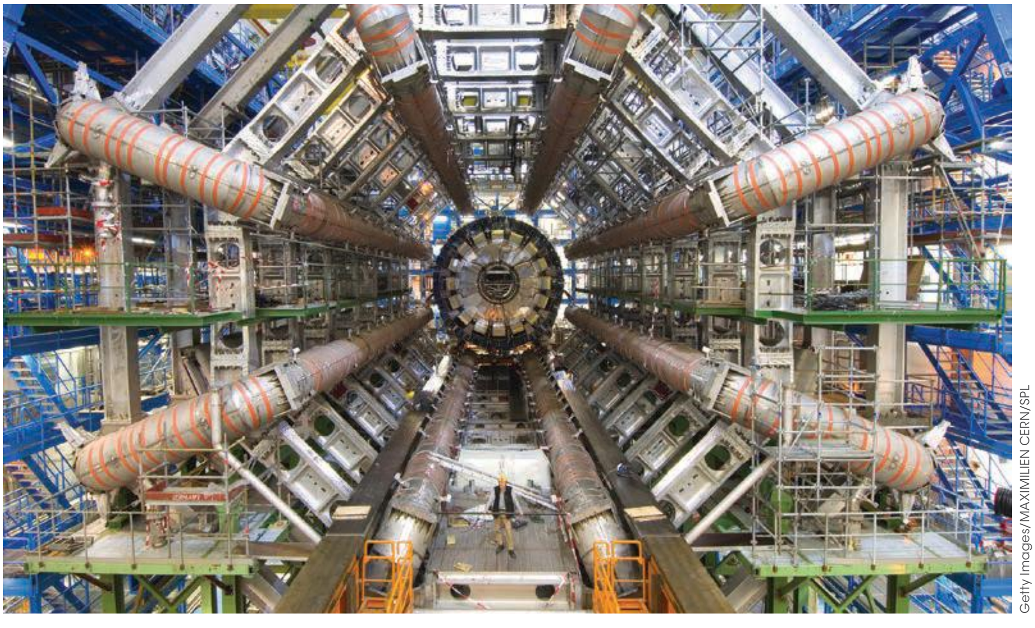
What is dark matter and dark energy?

How do quarks bind to form protons, neutrons, pions and other hadrons?

How do neutrinos gain mass and why are those masses so small?

Why is the universe filled with matter but essentially no antimatter?

Professor Barberio and her team, along with the many theoretical and experimental physicists with whom they collaborate, hope to answer some of these questions.



Getty Images/MAXIMILIEN CERN/SPL

▲ **Figure 10.13** The ATLAS detector at CERN

Questions

1 Why did Professor Barberio choose to study physics?

2 What fraction of the matter in the universe is estimated to be 'dark matter'?

3 Use the tables earlier in this chapter to answer these questions.

  a What are the properties (charge, spin, mass) of the b quark?

  b Name one meson and one baryon that contain a b quark.

4 Why was the discovery of the Higgs boson so important to particle physics?

5 Follow the weblink to the 1999 or 2008 Nobel Prize in Physics page. Prepare a five-minute talk or a poster on one of these discoveries for which the Prize was awarded.

A second problem is known as the 'flatness problem'. This is based on the idea that gravity curves space–time, an idea that comes from the general theory of relativity. It seems that the universe is very 'flat', which is a very unlikely outcome if it started as a singularity and evolved in the way described by the Big Bang theory.

The theory is also limited in what it explains. It does not, for example, explain how stars and galaxies form. In fact, it makes an assumption that matter is uniformly distributed through the universe on all scales – this is known as the **cosmological principle**. This assumption is actually at odds with the existence of galaxies, and even with the formation of atoms. In fact, matter appears to be 'clumped' at all scales rather than uniformly distributed.

The cosmic microwave background radiation is also not uniform, as predicted by the Big Bang theory, but varies slightly depending on the direction in which you look – another observation that cannot be explained by the Big Bang theory.

Particle physics and the Big Bang theory

Astronomical observations will continue to provide evidence to help cosmologists refine their theories. However, if the Big Bang theory is correct, then no matter how big the telescopes we build, we will not see further back in time than when atoms formed. Astronomical observations cannot give us information about the very hot, early stages of the universe or the singularity from which the universe started.

For this information we need to turn to particle physics and the accelerators and colliders that can create conditions similar to those of the early universe. Recall that at very high energies the four fundamental forces are believed to be unified into a single force. Theories that attempt to explain what this force was, and how it split into the four forces we know today are called grand unified theories. These theories can only be tested in experiments at enormously high energies – the sort of energies only available in massive particle accelerators. The Big Bang theory relies on the existence of grand unified theories to explain the first stages in the evolution of the universe from a tiny, enormously hot singularity.

Astronomical observations are limited in what they can tell us about the early universe. Information about the early universe will need to come from high-energy particle physics experiments.

It is also only in high-energy particle physics experiments that we can hope to actually *make* particles of dark matter and measure their properties. This is one of the future goals of particle physicists at CERN and other facilities.

Hence, the Big Bang theory will eventually be tested in experiments at particle accelerators like the LHC at CERN.

These experiments will not only help us build better models of the beginning of our universe, but they may also tell us what the ultimate fate of our universe will be – whether it continues to expand forever or contracts to end in a 'big crunch'.

QUESTION SET 10.3

Remembering

1 Why is dark matter called dark matter?

2 What is the evidence for the existence of dark energy?

3 What causes the redshift in observed spectra?

Understanding

4 Explain how the observation of spectra allows us to determine the speed of a star.

5 Why is the point at which the universe became transparent to photons significant?

Applying

6 Use Hubble's law to estimate the speed of recession of the nearest star to us, Alpha Centauri, 4.4 light-years away.

7 The redshift of light from distant galaxies implies that some of them are moving at a significant fraction of the speed of light. If a distant galaxy is found to be moving at 0.45*c*, how far away from us is it according to Hubble's law?

8 Use the information on black body radiation in Chapter 7 to find the peak wavelength radiated by:
   a a black body at 3000 K (the universe at 700 000 years after the Big Bang).
   b a black body at 3 K (the universe today).

Analysing

9 The hydrogen emission spectrum has a series of visible lines called the Balmer series, as described in Chapter 8. This series corresponds to transitions to the $n = 2$ energy level. For a distant galaxy, the line corresponding to the $n = 4$ to $n = 2$ transition is observed at a wavelength of 700 nm.
   a Find the wavelength of the line corresponding to the $n = 4$ to $n = 2$ transition on Earth.
   b Find the speed at which this galaxy is receding from us.
   c Find the distance of this star from us.

Reflecting

10 Draw a concept map to illustrate your understanding of the Big Bang theory. Include the evidence for the theory, as well the limitations of the theory, on your diagram. Include the things that the theory predicts and explains. Add links showing how it draws on ideas from other areas of physics.

CHAPTER SUMMARY

- The Standard Model of particle physics is the currently accepted model that describes the elementary particles of which matter is made and explains the forces via which they interact.
- The Big Bang theory is currently the most widely accepted cosmological model for the evolution of the universe, beginning with a singularity around 14 billion years ago. It relies on particle physics to explain the evolution of matter and forces from this initial singularity.
- More than 300 distinct types of particles have been observed. These can be classified as leptons and hadrons. Hadrons can be further classified as mesons and baryons.
- The Standard Model says that all hadrons are composed of elementary particles called quarks.
- Quarks come in six 'flavours': up, down, top, bottom, charmed and strange. For each quark type there is an antiquark.
- Mesons are composed of one quark and one antiquark.
- Baryons are composed of three quarks. Their corresponding antiparticles are composed of the corresponding three antiquarks.
- Interactions between particles occur via the exchange of field particles. According to the Standard Model of particle physics, these field particles are photons for the electromagnetic force, W and Z bosons for the weak force and gluons for the strong force.
- The four fundamental forces are the strong force (not the same as the strong nuclear force), the weak force, the electromagnetic force and the gravitational force.
- The strong nuclear force, which is mediated by pions, is a result of the strong force interaction between quarks, which is mediated by gluons.
- The Standard Model does not currently include gravitation, although some theories predict the existence of the graviton as the exchange particle for the gravitational force, and the existence of gravitational waves (analogous to photons and electromagnetic waves).
- In the Standard Model, particles acquire mass via interactions with the Higgs field. The mediating particle is the Higgs boson.
- The Higgs boson was first observed at CERN in July 2012.
- The Standard Model has limitations. It does not include a theory for gravity, and it predicts that neutrinos should be massless. Experimentally, neutrinos have been found to have very small but non-zero mass.

- The Big Bang theory explains the universe as having originated from the explosion of a singularity containing all the energy and matter of the universe.
- Astronomical evidence such as the redshift points to the expansion of the universe, which is compatible with the Big Bang theory.
- Hubble's law is an empirically derived relationship between the distance to any given star and its rate of recession from us.
- In the first stages of the formation of the universe, the four forces were unified as a single force. Symmetry breaking as the universe cooled resulted in the separation of the forces.
- Particles formed in the first second after the Big Bang.
- To understand the earliest stages of the universe and its current structure and composition we need to use particle physics as well as cosmology.

CHAPTER GLOSSARY

Big Bang theory the theory that the universe began with a massive explosion of matter from a single point

cosmic microwave background radiation the observed radiation coming from all points in space corresponding to radiation from a black body at 3 K; it is believed to come from an earlier, much hotter stage of the universe's evolution

cosmological principle the assumption that matter is uniformly distributed throughout the universe on all scales; it is one of the assumptions underlying the Big Bang theory

dark energy energy that is predicted from the increasing rate of expansion of the universe, but which is not identifiable as any currently known energy form

dark matter matter that is postulated to explain gravitational effects but which is not observable by the emission or reflection of light

electroweak theory theory combining electromagnetic and weak interactions

flavours (quarks) the six classifications of quark types: up, down, strange, charm, top and bottom

gauge bosons the field or exchange particles predicted by gauge theory

grand unified theory (GUT) a theory that unites all four fundamental forces in a single model and explains the symmetry-breaking mechanism that caused them to separate into the four distinct forces we know of. There is as yet no widely accepted GUT

quark type of elementary particle (along with leptons and gauge bosons). Quarks come in six flavours and are fermions

redshift the observed shift to longer wavelength of spectral lines in distant stars

singularity a point in space and time at which the density of matter/energy is infinite and the volume is infinitesimally small

Standard Model the current most widely accepted model of particle physics that uses quarks and leptons to explain the nature of matter and exchange particles to explain the origins of three of the four fundamental forces: strong force, weak force and electromagnetic force; does not include a theory for gravity

unified theory any theory that demonstrates how fundamental forces can be united, and explains the mechanism by which they become distinct, for example the electroweak theory

CHAPTER REVIEW QUESTIONS

Remembering

1 Name the three types of elementary particle.

2 Which forces are included in the Standard Model? Which particles interact via each of these forces?

3 Which force is mediated by the W and Z bosons? How are these exchange particles different from those that mediate the other fundamental forces?

Understanding

4 Why do physicists believe that protons and neutrons are not elementary particles? Describe the evidence that suggests this.

5 Why are we not able to 'see' further back than when the universe was approximately 700 000 years old?

6 The emission spectrum of a hydrogen lamp is observed to have a line with a wavelength of 486 nm. This line corresponds to a transition from the $n = 4$ state to the $n = 2$ state. Would you expect the line corresponding to this transition in hydrogen from a distant star to have the same wavelength, a longer wavelength or a shorter wavelength? Explain your answer.

7 Why is the Super-Kamiokande detector located so far underground?

8 The strong force and the strong nuclear force are often confused. Make a table comparing the two. Include columns headed 'acts on', 'exchange particle' and 'range'.

9 Why does charge always add arithmetically, but spin can add in two different ways?

10 The redshift is evidence that distant stars are receding from us. How would the wavelength shift be different if instead those stars were moving towards us?

Applying

11 Which of these quark combinations is a possible particle? For those that are possible, identify the particle.

- a cd
- b $\bar{u}$ud
- c $\bar{u}\bar{s}\bar{s}$
- d c$\bar{c}$

12 For each of these quark combinations, give the resulting particle's charge, strangeness and baryon numbers and all possible spins.

- a $\bar{d}\bar{s}\bar{s}$
- b bbb
- c uuu
- d $\bar{u}\bar{u}\bar{s}$

13 What is the quark composition of a D^- meson? Show that this composition gives the characteristics of a D^- particle.

14 Show that the characteristics of a Δ^{++} are consistent with its quark composition, as given in Table 10.3.

15 What is the largest possible charge of a baryon? Give an example of a combination of quarks that gives this charge.

16 A physicist claims to have detected a meson with a charge of $\frac{1}{3}$. Why is this impossible in the Standard Model?

17 A distant star is moving away from us at 0.15*c*.

- a How far away is this star?
- b What is the observed wavelength of the 656 nm red line in the Balmer series for this star?

Analysing

18 Consider the β^- emission process in which the following reaction occurs: $n \rightarrow p + e^- + \bar{v}_e$

a What is the quark composition of each particle before and after the reaction?

b Is quark flavour conserved in this reaction?

19 It has been observed at the Super Kamiokande observatory that neutrinos can change from one type to another. What does this imply about conservation of lepton number, *L*?

20 The NGC157 galaxy is 70 Mly away. Calculate the wavelength of the longest and shortest wavelength lines in the Balmer series for light coming from NGC157. Which of the lines are in the visible region?

21 A distant star is observed to have a hydrogen spectral line that is shifted from 103 nm to 201 nm. How far away is this star?

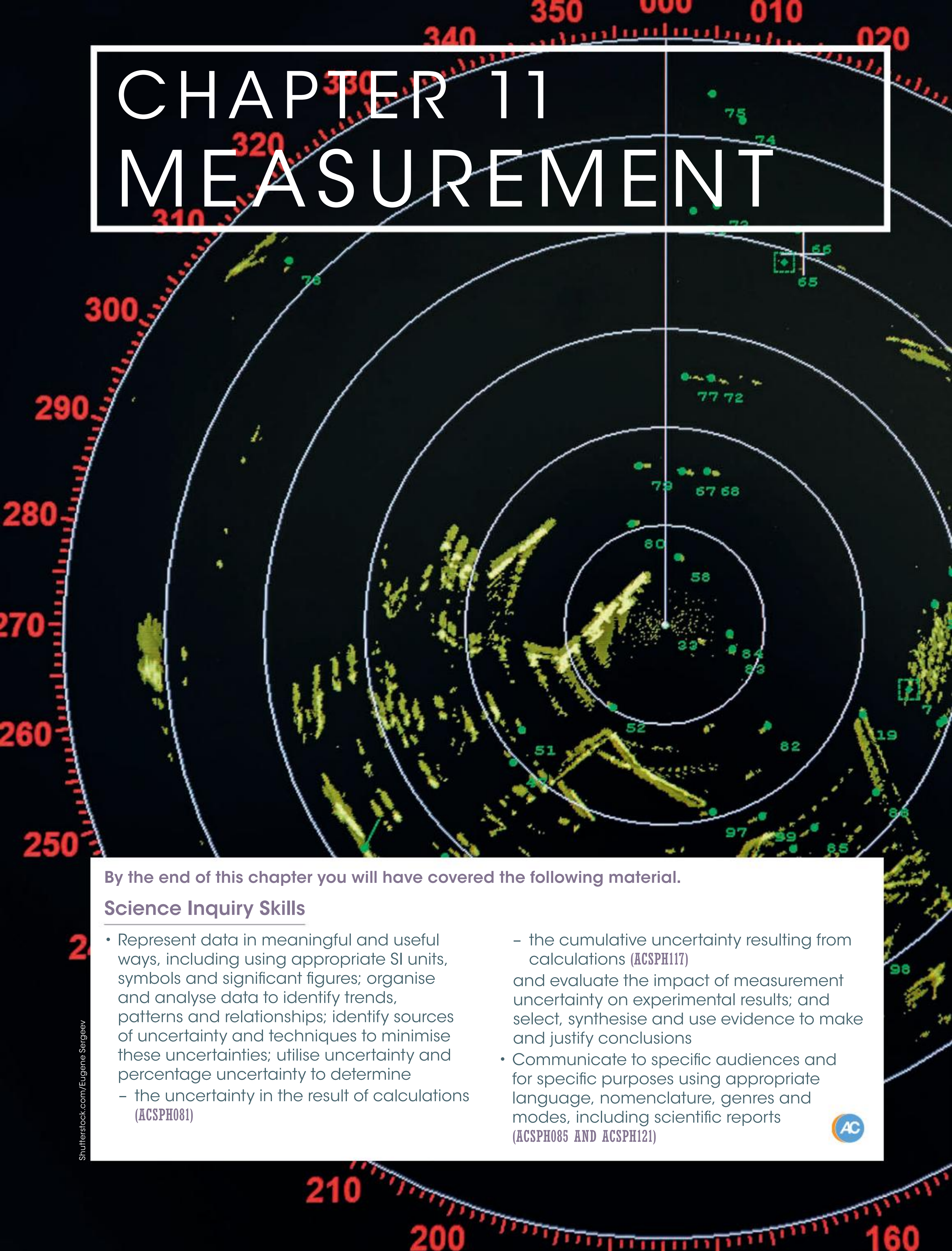
Reflecting

22 Reflect on the interaction between particle physics and cosmology. Compare and contrast the contributions each has made to our understanding of the universe.

23 Draw a concept map summarising your understanding of the Standard Model. Compare your diagram with that of another student, and add anything that you have missed. Do this at least three times, or until you are no longer adding new things.

24 How has your understanding of the way models and theories are developed and used in physics changed from your studies of the topics in Unit 4? Write a short summary describing your understanding of how models and theories develop and change. Describe the development of two models or theories as examples to illustrate your summary.

CHAPTER 11
MEASUREMENT



Shutterstock.com/Eugene Sergeev

By the end of this chapter you will have covered the following material.

Science Inquiry Skills

- Represent data in meaningful and useful ways, including using appropriate SI units, symbols and significant figures; organise and analyse data to identify trends, patterns and relationships; identify sources of uncertainty and techniques to minimise these uncertainties; utilise uncertainty and percentage uncertainty to determine
  - the uncertainty in the result of calculations (ACSPH081)
  - the cumulative uncertainty resulting from calculations (ACSPH117)

  and evaluate the impact of measurement uncertainty on experimental results; and select, synthesise and use evidence to make and justify conclusions
- Communicate to specific audiences and for specific purposes using appropriate language, nomenclature, genres and modes, including scientific reports (ACSPH085 AND ACSPH121)

Introduction

Physics, and science more generally, is the study of phenomena and their causes. We ask two questions of phenomena: What happens? Why does it happen?

By careful investigation, we attempt to make sure we know what happens, and with what degree of certainty we can say that the phenomenon happens. That is the purpose of experimentation. Once we are confident that we know what actually does happen, we search for a satisfactory **explanation**. There may be competing explanations, so we try to test the explanations to see whether they stand up in all circumstances. In this way, the explanations enable us to look for further data: data and theory inform each other.

In the case of competing explanations, we try to find **critical tests** of the explanations. A critical test will show that one explanation is superior to the other. A superior explanation is able to explain more of the evidence and has greater power to predict, and then explain, novel results. This should mean that one of the explanations is discarded. In reality, both explanations tend to exist side by side until it becomes obvious that one is superior, or an even more effective explanation takes hold.

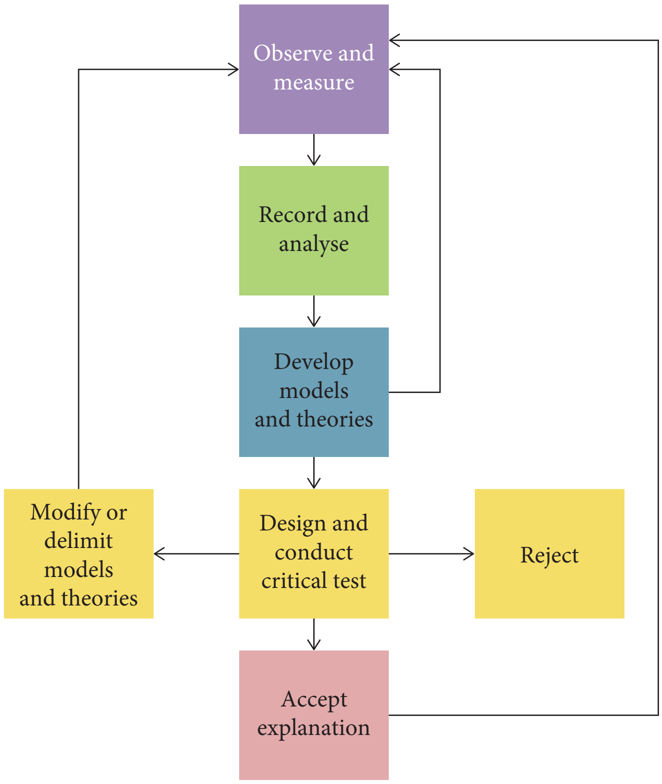


Figure 11.1 ▶ Flowchart representing the enterprise of doing physics

Philosophy and science

ROGER BACON

Find out more about the incredible breadth of Roger Bacon's scholarship.

Physics was known from classical times as 'Natural philosophy'. It was the study of, and explanation for, the physical world. Aristotle (384 BCE–322 BCE) was one of the most famous and influential of the classical natural philosophers. In the 13th century, Roger Bacon (c.1214–c.1292) integrated new interpretations of Aristotelian ideas into university teaching of theological and philosophical ideas. He argued successfully for reform in university education to include experiential science. He was well-versed in Greek and Muslim thought, especially in optics, which he used to develop a clear description of empirical science as well as its purposes – to gain understanding of phenomena, to explain it and to make predictions.

Aristotle's ideas came under increasingly intense scrutiny during the 15th through 17th centuries. Careful data collection and analysis provided the impetus for the gradual development of a new form of natural philosophy. Finally, the work of Nicolaus Copernicus (1473–1543), Francis Bacon (1561–1626), Galileo Galilei (1564–1642), Isaac Newton (1642–1727) and many others radically changed the philosophical basis of natural philosophy. Physics, the academic discipline we now study, was born.

Karl Popper (1902–94) was a philosopher of science. He argued that science proceeds by the way of **falsifiability**. That is, a theory is formed that starts off with a set of concepts, models, connected ideas and consequences linked to data, just like the solution to a mathematical problem. This construction of a theory fits some data and suggests new experiments. All experiments that fit with the theory and related models serve only to prove what we already know. However, if a theory is to be refined or changed, then we need to search for data from experiments that make the consequences of the theory false (falsifiability). Only then can we improve our understanding.

Falsifiability
Science proceeds by way of critical tests of explanations. A critical test is an experiment that, having been rigorously conducted, shows that one or other of competing theories is false.

Popper's ideas have been challenged by other philosophers of science. They disagree in three ways. First, they argue that falsifiability is too strict a criterion. Scientists gain confidence that they are on the right track when they confirm their theories through experiment.

Second, they point to the way scientists hold firm to useful ideas even when those ideas are under threat. It is not sufficient for new data to outweigh older data. Theories do not fail completely simply because some data falsifies them.

Often, the emerging theory does not seem to be better than the accepted one. This is a common experience for physics students: they do not accept what they are told unless it is better than what they already think. For example, Newton's laws are well accepted. They predict that all masses falling near Earth's surface will accelerate at the same rate. Yet, students who observe different freefalling masses striking the ground simultaneously frequently say, 'I've seen this many times. I've seen it, but I don't believe it.' Students, like scientists, are human. They hang on to dearly held ideas well beyond their use-by date. That slows progress towards better understanding and explanations.

The third way in which falsifiability is used is to constrain or put limits on the applicability of a theory. A theory may be found to be less general than previously claimed. In this way, the theory is saved but reduced in applicability. This happens for Newton's laws, which work well for big, slow things, but not for speeds approaching the speed of light.

Physics and philosophy are always in dialogue. The development of the quantum theory, which described the universe as fundamentally probabilistic, caused consternation to those who favoured a philosophically deterministic physics. It can be argued that Einstein's relativity arose in the 19th century philosophical ferment that resulted in deconstructionism and post-modernism. And let us not forget the possibility that the philosophical underpinnings of mathematics itself may be holding back further developments in science and philosophy.

Models in science

Data represent physical systems. Data can be organised in tables, graphs, images, diagrams and words. It can be in **quantitative** (numerical) or **qualitative** forms. All of these data sets represent the actual physical system. Relationships between data sets can be represented using mathematical relationships, geometric constructions and diagrams, words, computer programs, physical models, etc. These representations model the situation – they are not the physical reality any more than an architectural drawing is a building.

Models in science demonstrate relationships between measureable quantities. In the process, they are used to explain things. Models or **representations** can show multiple relationships between parameters so that we can predict new observations and relationships. New observations, whether made as a result of experiments carried out to test a model or otherwise, shed light on models. That is, representations are subject to data, and data is described and explained by models. Neither the data nor the model is identical with the situation being observed.

Models are central to science because scientists use them to describe, explain, relate and predict phenomena. Models can be expressed in a range of ways – via words (with language that is commonly metaphorical), images (actual or imagined), mathematics, computer simulations or physical constructions (including some machines). Models help scientists to frame physical laws and theories, and these laws and theories are also models of the world. Models are not static – as scientific understanding of concepts or physical data or phenomena evolves, so too do the models scientists use to describe, explain, relate and predict these.


iStockphoto/npix

Figure 11.2 ▲ 'My love is like a red, red rose' (Robert Burns). The rose, the girl and the poet's love are not identical with the poet's image or model of love. This photo is also a model; it is neither the girl nor the flowers.

BUREAU INTERNATIONAL DES POIDS ET MESURES

Visit the home page of the organisation that manages all international agreements about weights and measures. Look for SI base units, SI derived units and SI prefixes. You could even download and read the brochure that explains all.

In the case of light, there are competing explanations of data concerning the interactions between light and matter. These explanations are based on a particle model or a wave model. Interestingly, we can show that both these explanations are useful, but for different data sets. There are clear, critical tests where one or the other of the models fails to explain some of the data.

This causes us to reflect on the nature of explanation itself. When we try to explain something, we use our common experience as a starting point. We say that something '*is like*' something else. A picture, an image, a word, a photograph and a computer-generated simulation are examples of models. So are mathematical equations. Theories model reality. They are not themselves reality. Poets use imagery to model their vision of the world.

Lord Kelvin (1824–1907), of kelvin temperature fame, said about models: 'I am never content until I have constructed a mechanical model of the subject I am studying. If I succeed in making one, I understand; otherwise I do not.'

A model is an aid to understanding. But a model is not the same as the thing itself. A model is a model. The thing to which a model refers is the thing to which it refers.

Units and standards

Physicists are required to make very careful and accurate observations and measurements of physical quantities. Good measurements are crucial to good science. It is the responsibility of experimenters to know about measurement and to cover all possible ways in which a measurement could fail to be accurate and precise.

An internationally agreed system of units called the Système Internationale or SI system is used in science. It provides definitions of quantities, lists the most up-to-date values for important quantities, codifies measurement theory and practice, and provides standards for the reporting of measurements. No country uses only SI units. If it did, it would use kiloseconds instead of hours.

Fundamental units

Seven units are defined for the fundamental or basic quantities – length, mass, time, current, temperature, luminous intensity and amount of matter.

The fundamental or basic SI units and their definitions are:

- *Length:* The unit of length is the metre (m), which is defined as 1 650 763.73 wavelengths of the orange-red line of the spectrum of ^{86}Kr (krypton) in a vacuum.
- *Mass:* The unit of mass is the kilogram (kg), which is based on a cylinder of platinum–iridium alloy kept by the International Bureau of Weights and Measures in Paris.
- *Time:* The unit of time is the second (s), which is defined as the duration of 9 192 631 770 periods of the radiation corresponding to the transition between the two hyperfine levels of the ground state of the ^{133}Cs (caesium) atom at 0 K.
- *Electric current:* The unit of current is the ampere (A), and this is defined as the current that, if maintained in two straight parallel conductors of infinite length and negligible cross-section, separated from each other by a distance of 1 metre in a vacuum, will produce a force equal to 2×10^{7} newton per metre of length between the conductors.
- *Temperature:* The unit of temperature is the kelvin (K), which is defined as $\frac{1}{273.16}$ of the thermodynamic temperature of the triple point of water.
- *Luminous intensity:* The unit of luminous intensity is the candela (cd), which is defined as the luminous intensity in the perpendicular direction of a surface of $\frac{1}{600\,000}$ square metre of a black body at the freezing temperature of platinum (2042 K) under a pressure of 101 325 pascals.

- *Amount of substance:* The unit of amount of substance is the mole (mol), which is defined as the amount of substance of a system that contains as many elementary entities as there are atoms in 0.012 kilogram of carbon-12. (This number is approximately 6.023×10^{23}.)

These seven definitions are obviously quite impractical for everyday use, but they have been chosen carefully to provide invariable standards from which more practical devices can be manufactured; for example, a ruler or a stopwatch.

Derived units

Derived units are formed by combinations of the fundamental units. A simple example is the unit for area, the square metre (m^2). Other examples are:

- Volume – cubic metre (m^3)
- Speed – metre per second ($m\,s^{-1}$)
- Density – kilogram per cubic metre ($kg\,m^{-3}$)

A number of derived units have been given special names to commemorate notable scientists. These include frequency (hertz, Hz); force (newton, N); work and energy (joule, J); power (watt, W); current (ampere, A).

Prefixes

Consider the measurement of length. The metre is too large a unit with which to measure the thickness of this page. It is too small a unit to measure the distance to the Moon. For this reason, multiple or submultiple units may be formed by adding a prefix to the SI unit. The prefix is combined with the unit name and is written as one word.

millimetre (mm) equal to 10^{-3} m	megawatt (MW) equal to 10^6 W
centimetre (cm) equal to 10^{-2} m	kilogram (kg) equal to 10^3 g
kilometre (km) equal to 10^3 m	gigajoule (GJ) equal to 10^9 J

The preferred prefixes relate to the SI units usually by powers of three. The common prefixes are given in Table 12.1.

Table 11.1 Common prefixes

Multiple	Prefix	Symbol	Multiple	Prefix	Symbol
10^{18}	exa	E	10^{-2}	centi	c
10^{15}	peta	P	10^{-3}	milli	m
10^{12}	tera	T	10^{-6}	micro	μ
10^9	giga	G	10^{-9}	nano	n
10^6	mega	M	10^{-12}	pico	p
10^3	kilo	k	10^{-15}	femto	f

All measurements require at least a value and a unit. Some also require a direction. In more advanced measurements an indication of accuracy is also given. The value is often written in a standard form.

You should get into the habit of writing numerals and units correctly. The unit is separated from the numeral by a space. Each separate part of the unit is also separated by a space. For example, a speed of 5.3 metres per second is written as $5.3\,m\,s^{-1}$. Notice the space between 5.3 and 'm', and the space between 'm' and 's' (ms^{-1}, no space, means 'per millisecond').

Standard or scientific form

Standard or scientific form is widely used in science. It enables very large and very small numbers to be expressed relatively simply. Numbers are written as a number between 1.0 and 10 multiplied by the relevant power of 10. For example:

- The distance from Earth to the Sun is 150 000 000 km. In standard form this is written as 1.5×10^8 km.
- The average distance between two atoms is 0.000 000 000 16 m. This is written as 1.6×10^{-10} m.

In some cases, it is not sensible to write measurements in this form. For example, the measurement 0.8 m is better left as it is than written as 8×10^{-1} m. As a general rule, it is not usual to express the numbers between 0.01 and 1000 in standard form. It is not wrong to do it; it is just not the preferred method.

Dimensions

In the SI system, the unit of speed is the metre per second. This is not the only unit that can be used. Some alternative, equally correct units are mile per hour, kilometre per hour and centimetre per minute. All of these have one thing in common – each is a unit of length per unit of time. The quantity length per time is called the dimension of the property speed.

When writing the dimensions of a physical quantity, we use the symbols [L], [M], [T] and [I] to represent the dimensions of length, mass, time and current, respectively. Hence, we can write: $[\text{speed}] = [\text{length}]\,[\text{time}]^{-1} = [L]\,[T]^{-1}$, where the brackets are read as 'dimensions of'.

All of the quantities used in this book can be expressed in terms of a few fundamental quantities. Examples include $[\text{area}] = [L]^2$, $[\text{volume}] = [L]^3$, $[\text{density}] = [M]\,[L]^{-3}$.

The unit of all derived quantities can be found by substituting the fundamental unit for each dimension in the dimension equation. For example, in the SI system the unit of mass is the kilogram (kg) and the unit of length is the metre (m). The unit of density – dimension $[M]\,[L]^{-3}$ – then becomes $\text{kg}\,\text{m}^{-3}$ in the SI system, but $\text{lb}\,\text{ft}^{-3}$ in the British Imperial System.

The examination of the dimensions of an expression can be used to check whether an equation is likely to be correct. The dimensions of both the left-hand side and right-hand side of all equations must be dimensionally correct and equal, so that they balance. If they are not, *then the equation cannot be correct*. For example, the circumference of a circle, $C = 2\pi r$ has dimensions of length, [L] on both sides: $[\text{circumference}] = [L]$, $[2\pi r] = [\text{radius}] = [L]$ – the constant, 2π has no dimensions. Hence, the equation is dimensionally correct. Note, however, that the equations $C = 2r$ and $C = \pi r$ are also dimensionally correct, but are in fact wrong. The dimension check will only tell us if the equation is wrong – not that it is necessarily correct.

Making and reporting measurements

When we make a measurement we have some idea that what we measure is relevant and appropriate to our interest. We may also expect there to be a **'true value'**, an exact number that represents what we are measuring. This cannot, however, be known with 100% certainty. Whenever a measurement is undertaken two things need to be guaranteed. The **measurand**, the quantity being measured, should be clearly specified. The **measurement result**, the best estimate of the 'true value', should be both **accurate** and **precise**.

Measurand

Mostly, it will be obvious which quantity is to be measured: length on a ruler, mass on a weighing machine, time on a clock, current on an ammeter, potential difference on a voltmeter. In student-designed practical investigations it may be necessary to decide what to measure: vertical angle of a cone, 'squash' of a crumple zone, extension of a spring, density of a fluid, rate of temperature change of a solid. These quantities are generally measured using length, mass and time quantities. Sometimes, the measuring device obscures this. Digital meters often rely on length, mass and time to produce the number on the dial, but this is not nearly as clear as on an analogue meter. For example, a digital radar gun measures the time between successive sound waves and internally converts this to a speed (distance/time). An analogue thermometer has an obvious, linear scale printed on it.

Many students think that because they get a numerical readout on a digital meter it is more accurate than an analogue meter. In general, this belief must be questioned. When comparing the readings from an analogue meter and a digital meter, the analogue meter is often more accurate. For example, if both give a measurement to the second decimal place, the uncertainty for the analogue meter will be a maximum of ± 0.005. For the digital meter, the uncertainty is ± 0.01.

Measurement result

The measurement result is the best estimate of the 'true value'. Except for discrete, countable things, the 'true value' is an ideal that can never be completely and unambiguously known. It is logically impossible to know the exact value of a continuous variable, even for standard values, such as the speed of light.

Measurand is a specified quantity to be measured.

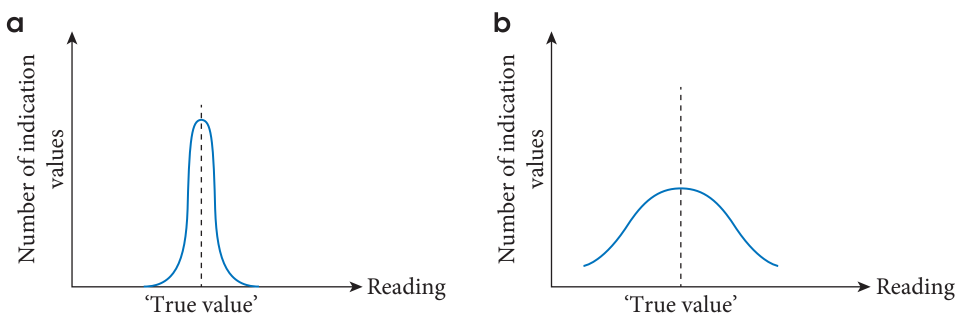
Measurement result is the best estimate of the 'true value' of a measurand.

'True value': for continuous variables this is an unknowable, ideal value that represents the measurand.

It is possible to make better or worse estimates of the 'true value'. This depends on the environment, the quality of the equipment and the skill of the person making the measurement.

Accuracy

An accurate measurement result is one that represents the 'true value' of the measurand as closely as possible. How can this be achieved? Every measurement provides an indication of the 'true value'. If we take repeated measurements, there will always be a spread of these **indication values** or results. However, the mean of these indication values should be a very good estimate of the 'true value'. Notice this is an agreed procedure. We agree that the mean is the most likely or best estimate of the 'true value'. It is not certain that it is the 'true value'. A plot of the number of indication values versus reading shows the spread of results around the supposed 'true value'.



◀ **Figure 11.3**
In a plot of number of indication values versus reading, results may: a) cluster close to the 'true value', or b) be spread significantly around the 'true value'.

The means of Figures 11.3(a) and (b) are the same. The two measurement results are identical. Both report the same mean as the best estimate of the 'true value'. In this sense, both measurement results are accurate.

It is possible to get plots that look similar yet not achieve a proper measurement result for the 'true value'. If the measurand is specified incorrectly, then the measurement result will cluster around another value that is not the 'true value'. This is shown in Figure 12.4.

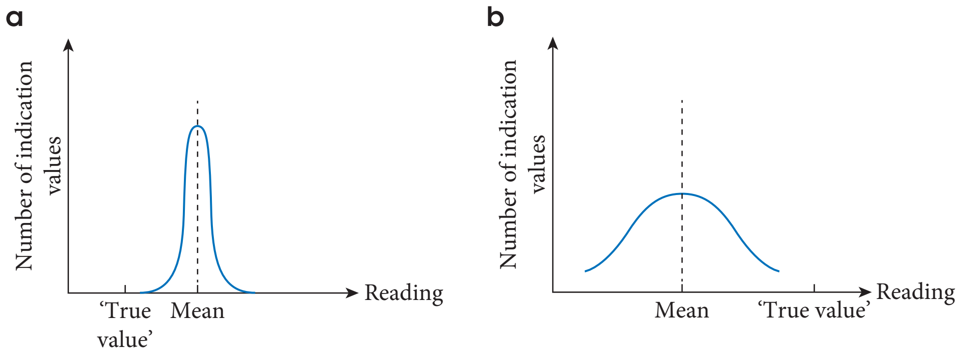


Figure 11.4 ▶ In a plot of number of indication values versus reading, a) many results cluster close to a value that is not the 'true value'. b) The results may be spread significantly around a value that is not the 'true value'.

The means of Figures 11.3(b) and 11.4(b) are the same. The two measurement results are identical. Both report the same mean of the indication values, but they do not represent the best estimate of the 'true value'. In this sense, both measurement results are inaccurate.

Precision

Precision relates to the skill of the experimenter within the environment and the quality of both the equipment and the measuring techniques used. If the quality of these is high, the measurement result is highly precise; if the quality is low, the result is imprecise. Figures 11.3(a) and 11.4(a) show precise measurement results because the individual indication values cluster closely around the mean. But Figures 11.3(b) and 11.4(b) show imprecise measurement results because the individual indication values spread significantly around the mean.

Accurate: value that represents the 'true value'

Precise: small spread around a mean value

Accurate and precise: small spread around the 'true value', which is taken to be the mean of indication values

Uncertainty

No measurement is exact. There are always effects that contribute to each measurement of a quantity being a bit different. They add to the **uncertainty** with which an indication value or measurement result can be reported. Good experimental observers always ensure that they can estimate the uncertainty, which is the range of values between which they are confident the 'true value' lies. Figures 11.3(a) and (b) both show the dispersion of indication values around the 'true value'. This dispersion is the result of random effects such as small air currents or localised temperature changes. This dispersion of indication values is called **random error**, although this term is starting to be replaced by more precise concepts. The term 'random error' alerts us to the possibility that, if the dispersion of the results is not considered, a mistake will be made in reporting the value correctly.

It is not possible, in principle, to predict the next indication value from the previous measurement. If it were, the effect could be taken into account in making the estimate of the 'true value'. Figures 11.4(a) and (b) show the mean of the indication values, which are offset from the 'true value'. Some of this offset can be accounted for by careful consideration of the situation and the measurement activities. These **systematic errors** can be identified and indication values adjusted for these known, regular effects. For example, when you weigh ingredients for a recipe, the scales may read 200 g before you put the ingredients in the bowl. After weighing the ingredients you would adjust the result by taking 200 g from the scale measurement to work out the actual quantity. This is referred to as calibration. Often calibration requires correction for a **zero error**, as in this example. The term, systematic error, alerts us to the possibility that, if they are not considered, a mistake will be made in reporting the value correctly. When drawing graphs of data, you should never force a line of best fit through the origin, even if you believe it should pass through the origin. A line of best fit that does not pass through the origin when it is expected to do so may be alerting you to a systematic error in your data. The systematic error should not be ignored.

The uncertainty in a measurement is a quantity that makes apparent the quality of the measurement. Some people refer to measurement uncertainty as measurement error. This is unhelpful. An error is a mistake. There should be no mistake in making the measurement. The uncertainty represents a careful estimate of doubt about the value.

Reporting uncertainty

Any measurement should be reported with two numbers: the **best estimate** of the value observed in the measuring system, and an estimate of how much uncertainty there is in this value. For example, the radius of a small ball can be measured with greater or less precision by different measuring devices. Using a ruler, its diameter might be recorded as (2.5 ± 0.1) cm. A different system might enable the diameter to be recorded as (2.513 ± 0.001) cm. Both these results would be accurate if there were no zero error. In general, the measurement with the least uncertainty is the most precise.

Managing uncertainty in practice

Measurements are used to derive new quantities. The best estimate of the measurand is used, for example, to fix a point on a graph or substitute into an equation. The uncertainty comes along as well. We need to manage the effect of this uncertainty on the overall quantities produced from such operations.

Precision of a measurement

We use place value to report the precision of a measurement. For (2.5 ± 0.1) cm, the final place in the value, the 5 in the tenths column for 2.5, is uncertain by 0.1 cm. We are saying, 'The best estimate of the value is 2.5 cm, and we are confident the 'true value' lies between 2.4 cm and 2.6 cm'. Similarly, for the more accurate and precise measurement (2.513 ± 0.001) cm, we are reporting confidence that the 'true value' lies between 2.512 and 2.514. The measurement is uncertain in the thousandths column.

The number of **significant figures** in a measurement can be found by counting the number of reported digits. There are rules for deciding on the number of significant figures in a reported value:

1. Zeros in front of the integer part of a numeral are not significant; for example, 00346 has 3 significant figures.
2. All non-zero figures are significant, for example in 25.4 there are 3 significant figures.
3. All zeros between non-zero digits are significant, for example in 203.4 and 27.6002 the zeros are significant. They have 4 and 6 significant figures respectively.
4. All zeros to the right of a decimal point, which follow a non-zero digit, are significant, for example in 21.000 the zeros are significant (5 significant figures).
5. For numbers less than 1, the zeros before the first non-zero digit are not significant, for example, 0.003 682 has 4 significant figures, starting at the digit 3.

Numbers between zero and 0.01 and numbers between 10 and 1000 are not usually written in standard form. Between 10 and 1000, a zero in the units column is then treated as significant. Thus, 200 and 850 both have 3 significant figures.

Some examples are shown below (the number of significant figures is reported in the bracket).

700 046 (6), 901.040 (6), 5403.2, (5), 350 (3), 67.2, (3), 64.0 (3), 0.409 (3), 0.0038 (2)

Other than 100–1000, a number such as 350 000 has 2 significant figures, because it can be written as 3.5×10^5; however, if written as $3.500\,00 \times 10^5$ it has 6 significant figures.

WORKED EXAMPLE 11.1

How many significant figures are in the following numbers?

a 0.0098 (1 mark)

b 758.0 (1 mark)

c 650 (1 mark)

d 3.0002 (1 mark)

e 21 000 (1 mark)

Answers	Logic	
a 2	Use the rules for significant figures.	1 mark
b 4		1 mark
c 3		1 mark
d 5		1 mark
e 2 (2.1×10^4)		1 mark

Try these yourself

How many significant figures are in the following numbers?

a 0.0241	(1 mark)
b 6.098	(1 mark)
c 0.0045	(1 mark)
d 0.006 20	(1 mark)
e 650 000	(1 mark)

Relative and percentage uncertainty

Uncertainties can be reported as absolute or relative uncertainties. The examples above are all absolute uncertainties. They always have the same units as the measurement and are usually given to only 1 significant figure.

Uncertainties can also be given as relative (proportional or fractional) or percentage uncertainty. **Percentage uncertainty** is calculated by finding the **relative uncertainty** or **proportional uncertainty**, then converting it to an equivalent fraction out of 100:

$$\text{Relative uncertainty} = \frac{\text{uncertainty}}{\text{value}}$$

$$\text{Percentage uncertainty} = \frac{\text{uncertainty}}{\text{value}} \times \frac{100}{1}\%$$

Relative and percentage uncertainties are usually given to two significant figures. They have no units because they are the ratio of the absolute uncertainty in the measurement. The absolute uncertainty and the measurement have the same units, so when you divide one by the other, the units cancel out.

WORKED EXAMPLE 11.2

Calculate the relative and percentage uncertainties for these values.

a 2.5 ± 0.1 cm (2 marks)

b 2.513 ± 0.001 cm (2 marks)

Answers	Logic	
a Relative uncertainty $= \frac{0.1\text{ cm}}{2.5\text{ cm}} = 0.04$	Use the correct fraction, calculation and number of significant figures.	1 mark
Percentage uncertainty $= \frac{0.1\text{ cm}}{2.5\text{ cm}} \times \frac{100}{1}\% = 4\%$	Use the correct fraction, calculation and number of significant figures.	1 mark
b Relative uncertainty $= \frac{0.001\text{ cm}}{2.513\text{ cm}} = 0.00040$	Use the correct fraction, calculation and number of significant figures.	1 mark
Percentage uncertainty $= \frac{0.001\text{ cm}}{2.513\text{ cm}} \times \frac{100}{1}\% = 0.040\%$	Use the correct fraction and calculation.	1 mark

Try these yourself

1 Calculate the relative and percentage uncertainties for these values.

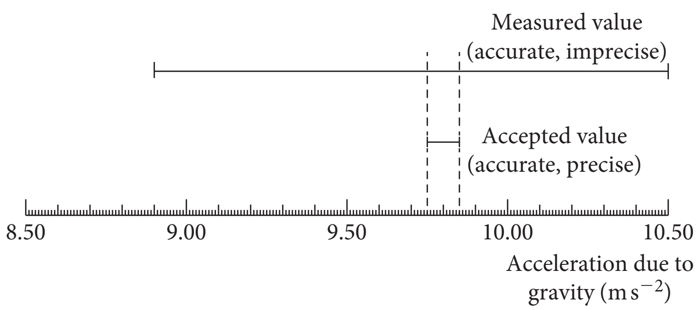
- **a** (2.5 ± 0.1) s (1 mark)
- **b** (0.0524 ± 0.003) μm (1 mark)
- **c** (250.0 ± 0.8) kg (1 mark)

2 In radiation counting, the uncertainty in N counts is $\pm\sqrt{N}$. Calculate the percentage uncertainties for these counts.

- **a** 100 counts (1 mark)
- **b** 1000 counts (1 mark)
- **c** 10 000 counts (1 mark)

Proportional error

Every measurement, including highly precise measurements, should be reported with an uncertainty value. When two measurements of the same quantity are compared, the range of uncertainties must be considered. For example, the accepted value for the acceleration due to gravity near Earth's surface might be reported at a location to be (9.80 ± 0.02) m s^{-2}. This means that the value could be anywhere between 9.78 m s^{-2} and 9.82 m s^{-2}. In a laboratory experiment, a student measured the acceleration due to gravity to be (9.7 ± 0.8) m s^{-2}. This means that the student's value could be anywhere between 8.9 m s^{-2} and 10.5 m s^{-2}. The student's result is clearly less precise than the accepted value. However, the student has managed to measure a value that fits within the range of uncertainty of the accepted value. With a less precise measurement, the student has nevertheless produced a confirming instance of the same value as the accepted value. The accepted value and the student's value both encompass a region in which the 'true value' could actually lie. This is within the overlap region, between 9.78 m s^{-2} and 9.82 m s^{-2}. Hence the two values agree.



◀ **Figure 11.5** The uncertainty of a student-measured value overlaps that of the accepted value. The two values agree.

Some measurements are extremely precise, well beyond the precision generally possible in a school laboratory. These accepted values can be considered *as though* they are 'true values'. For example, the charge on an electron is $1.602\,176\,565 \times 10^{-19}$ C. The uncertainty is $0.000\,000\,035 \times 10^{-19}$ C. The relative uncertainty is an incredibly precise 2.2×10^{-8}.

In these cases, the concept of **proportional error** can be defined. It is the difference between a measurement result and an accepted value, expressed as a fraction of the accepted value:

$$\text{Proportional error} = \frac{|\text{measured value} - \text{accepted value}|}{\text{accepted value}}$$

PHYSICAL CONSTANTS

Search for particular constants, such as the electronic charge, speed of light and mass of a proton.

Proportional error is a way for students to compare their measurement result with an accepted value.

Percentage error is the ratio of the magnitude of the proportional difference between an accepted value and a measurement result, expressed as a percentage:

$$\%\ \text{error} = \frac{|\text{measured value} - \text{accepted value}|}{\text{accepted value}} \times \frac{100}{1}\%$$

Proportional error and percentage error are useful for helping students compare their accuracy with far more precisely measured quantities; however, neither is defined in the international measurement standards from the Bureau International des Poids et Mesures (BIPM).

WORKED EXAMPLE 11.3

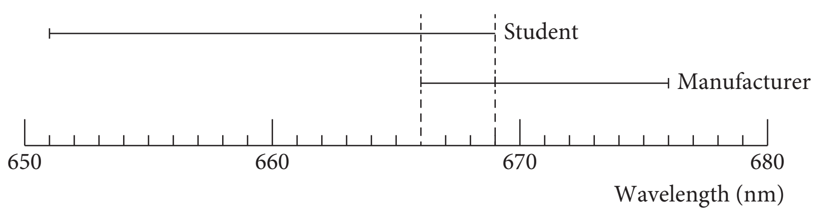
A manufacturer gives the wavelength of a light as $(671 \pm 5) \times 10^{-9}$ m. A student measured the wavelength to be $(660 \pm 9) \times 10^{-9}$ m.

a Does the student's measurement result fit within the accepted value measurement result? Sketch a graph of range of indication values versus reading to show any differences between the measurement results. (5 marks)

b Calculate the percentage error in the student's measurement result. (2 marks)

Answers

a The two measurement results overlap in the region $(666–669) \times 10^{-9}$ m.



▲ **Figure 11.6** Overlap of measurement results for manufacturer's accepted value and student measurement result

b $\%\ \text{error} = \frac{|\text{measured value} - \text{accepted value}|}{\text{accepted value}} \times \frac{100}{1}\%$

$\Rightarrow \%\ \text{error} = \frac{|660\ \text{m} - 671\ \text{m}| \times 10^{-9}}{671 \times 10^{-9}\ \text{m}} \times \frac{100}{1}\%$

$\Rightarrow \%\ \text{error} = \frac{11}{671} \times \frac{100}{1}\%$

$\Rightarrow \%\ \text{error} = 1.6\%$

As both measurements incorporate the region in which the 'true value' may be meaningfully said to exist, this figure is of little relevance. This is why the BIPM does not define percentage error.

Logic

Identify the overlap.	1 mark
Label the horizontal axis and mark the correct overlap values.	4 marks
Substitute the correct values into the formula.	1 mark
Calculate the answer.	1 mark

Try these yourself

The colour code on a resistor claims the resistance to be $1000 \pm 50\,\Omega$. A student who uses a specially designed circuit to determine the resistance finds the value to be $1025 \pm 15\,\Omega$.

a Does the student's measurement result agree with the accepted value of the measurement result? (2 marks)

b Calculate the percentage error in the student's measurement result. Comment on its meaningfulness. (2 marks)

Effect of uncertainty in derived quantities

Measured **raw data** are used to find **derived data**. Any uncertainty in the raw data must be taken into account in reporting derived data. In many instances data are given without uncertainty values. In these cases, the last decimal place represents the first uncertain figure. The number of significant figures in each figure is the number of digits, taking into account the rules for significant figures.

Adding and subtracting raw data

When adding and subtracting numbers with uncertainties, it is useful to take advantage of place value. Perform the operation with the best estimate values. Add the individual uncertainties to find the uncertainty in the sum or difference. Claim the best estimate only as far as the sum of uncertainties will allow; that is, the place value of the first digit of the sum of the uncertainties. For example, consider the following sum:

$$(17.23 \pm 0.02) + (5.1 \pm 0.4)$$

Add the best estimates of the value: $17.23 + 5.1 = 22.33$
Add the uncertainties for each value: $0.02 + 0.4 = 0.42$

The answer could be given as 22.33 ± 0.42, but the result becomes uncertain at the first decimal place in the uncertainty, 4 in 0.42. Hence, the best estimate, with uncertainty, of the sum is 22.3 ± 0.4.

When there is no uncertainty given in a data point, the *last decimal* place in each number is regarded as the first uncertain figure. Then the result must have no more decimal places than the number with the least number of decimal places. The addition or subtraction is performed with the numbers as given, then rounded up from 5 or down from 4. For example:

$$\begin{array}{r} 32.2187 \\ +\ 126.3 \\ +\ \ \ 3.132 \\ \hline 161.6507 \end{array}$$

This is rounded up (0.65 to 0.7) to 161.7 because the number 126.3 is given to only one decimal place.

$$\begin{array}{r} 3052.3 \\ -\ \ 235 \\ \hline 2817.3 \end{array}$$

This is rounded down (0.3 to 0) to 2817 because the number 235 is given to the nearest whole number.

Multiplying and dividing raw data

When multiplying and dividing numbers with uncertainties, perform the operation using the best estimate values. Add the individual relative or percentage uncertainties to find the relative or percentage uncertainty in the product or quotient. To find a value for the uncertainty, the percentage or proportional uncertainty is used. For example, consider the following product:

$$(50.6 \pm 0.8) \times (123.63 \pm 0.91)$$

Multiply the best estimates: 50.6×123.63

$$= 6255.678$$

Find the percentage uncertainties for each number:

$$= \frac{0.8}{50.6} \times \frac{100}{1}\% \quad \text{and} \quad = \frac{0.91}{123.63} \times \frac{100}{1}\%$$

$$= 1.581\% \qquad\qquad = 0.736\%$$

Add the percentage uncertainties:

$$= 1.58\% + 0.736\%$$

$$= 2.371\%$$

$$= 2.3\% \text{ by convention}$$

The product can be given as (6255.68 ± 2.3)%, but this hides a problem. The product has far more significant figures than the individual raw data numbers from which it was calculated. That is, the derived quantity is claimed to be more accurate than the raw data from which it was derived! That cannot be.

Let us look at the actual value of the uncertainty:

$$2.3\% \text{ of } 6255.68 = 143.88$$

The result could now be reported, inaccurately, as (6255.68 ± 143.88). The result, 6255.68, is shown to be uncertain in the hundreds column. Using standard form to express the significant figures, the result should now be written, correctly, as:

$$(6.2 \pm 0.1) \times 10^3$$

As long as the raw data measures what is intended, this is an accurate and precise statement of the derived quantity. Notice, however, that the derived quantity has 2 significant figures. This is less than the number of significant figures in the number with the least significant figures, namely 50.6 (3 significant figures). The number of significant figures in a derived quantity is always equal to or less than the number of significant figures in the least precise piece of raw data. In experiments, physicists try to ensure that uncertainties in the raw data do not accumulate to the point where derived quantities become meaningless. This takes high-level thinking, planning and effort.

When there is no uncertainty given in the data, calculations are performed using the unrounded data. The result is then rounded to a figure that has the same number of significant figures as the data with the *least number of significant figures*.

Two examples follow.

i
```
    45.71    (4 significant figures)
 ×  34.1     (3 significant figures)
 ---------
  1558.711   (7 significant figures on calculator)
```

This is rounded to 1.56×10^3 (3 significant figures) because 34.1 has the least number of significant figures (i.e. 3).

ii $\frac{5465.48}{2.4}$ a 6 significant figure number divided by a 2 significant figure number

$= 2277.283333$ (10 significant figures on calculator)

$= 2.3 \times 10^3$ (2 significant figures)

This is rounded to 2.3×10^3 because 2.4 has the least number of significant figures (i.e. 2).

Numerical methods for calculating uncertainties involving functions

Sometimes, measurements are used in functions such as sin, cos and tan. An uncertainty value can be found numerically, rather than by applying the rules outlined above. Consider the way the function works before deciding how to proceed.

WORKED EXAMPLE 11.4

A student conducts an experiment to find the refractive index of an organic biochemical. Angles of incidence, i, and angles of refraction, R, are measured several times and an average, with uncertainty, is found for each. The refractive index is then calculated from the equation:

$$n = \frac{\sin i}{\sin R}$$

Use the following data to compute the refractive index, with uncertainty:

$i = (43.6 \pm 0.5)°$, $R = (32.1 \pm 0.5)°$ (6 marks)

Answer	Logic	
Calculate the refractive index from the best value estimate: $n = \frac{\sin(43.6°)}{\sin(32.1°)} = 1.298$	Substitute the known values and calculate the answer.	2 marks
Decide what is the worst case possible. As the sine function is increasing in the range of possible values, calculate: $n = \frac{\sin(43.6 + 0.5°)}{\sin(32.1 - 0.5°)} = \frac{\sin(44.1°)}{\sin(31.6°)} = 1.328$	Substitute the known values and calculate the answer.	2 marks
The difference between these two values is: $1.328 - 1.298 = 0.030$. This shows the second decimal place is uncertain, so the answer must not go beyond the second decimal place.	Calculate the difference.	1 mark
Thus, after rounding, the refractive index, n, of the liquid is 1.30 ± 0.03.	Give the answer and its uncertainty.	1 mark

CHAPTER SUMMARY

Model

- Models are used to describe, explain, relate and predict phenomena.
- Models can be represented by words, images, mathematics, or physical constructions.
- Models help scientists to frame physical laws and theories.
- Laws and theories are models of the world.
- Models change: Data and ideas affect models.
- Falsifiability is the process that uses a critical test to determine whether a model needs to be replaced or constrained.

Units and standards

- SI units are an internationally agreed consistent set comprising seven fundamental units: length (m), mass (kg), time (s), electric current (A), temperature (K), luminous intensity (cd) and amount of a substance (mol).
- All units for other quantities are derived from these units.
- Standard or scientific form: Numbers less than 0.01 and greater than 1000 are written as a number between 1 and 10 multiplied by 10 raised to an integral power.
- Dimensional analysis can be used to:
  - check if an equation is possibly correct.
  - determine units in different systems.

Measurement

- Purpose: To find a value for a quantity of interest (measurand)
- Value is never exact. It is affected by system, procedure, operator skill and environment.
- Experimenter responsibility: All systematic errors should be quantitatively specified and measurements adjusted accordingly. The effect of random errors should be reported in the value reported for the measurand. No errors, that is, no mistakes, including mistakes about systematic and random errors, should be allowed to affect results.
- Mean of indication values is the best estimate of the 'true value'.
- Uncertainty is the measure of the dispersion of indication values around the mean; hence, it is an estimate of the quality of an indication value or measurement result.

Managing uncertainty in practice

- Significant figures: Place value is used to show the significance of a measure. The number of digits in a measurement is an indication of its accuracy.
- Percentage uncertainty is the relative uncertainty expressed as a percentage:

$$\text{Percentage uncertainty} = \frac{\text{uncertainty}}{\text{value}} \times \frac{100}{1}\%$$

- Percentage error is a quantity produced for learning purposes, but which is not defined by the BIPM:

$$\%\ \text{error} = \frac{|\text{measured value} - \text{accepted value}|}{\text{accepted value}} \times \frac{100}{1}\%$$

- Raw data are directly measured quantities, including uncertainty. Derived data is computed from raw data. The effect of uncertainty in raw data is accumulated in derived quantities.

CHAPTER GLOSSARY

accurate the degree to which a measurement result approaches the 'true value'

best estimate value chosen to represent the indication value or measurement result of a measurand

critical test an experiment that, having been rigorously conducted, shows that one or other of competing theories is false

derived data data that is deduced from raw data by mathematical manipulation, such as graphs, algebraic equations and geometric constructions

explanation generalised account of why a body of data occurs

falsifiability principle used to determine the experimental data that would disprove a model, law or theory; data from a critical test

indication value a single result of a measurement; the indication value gives a hint as to the 'true value'

measurand quantity being measured

measurement result best estimate of a 'true value'; numerical value based on judgements about one or more attempts to measure the 'true value'

model a representation of a system or phenomenon that explains the system or phenomenon; a model may be mathematical equations, a computer simulation, a physical object, words or other form

percentage error proportional error expressed as a percentage; not defined by BIPM

percentage uncertainty proportional uncertainty, expressed as a percentage

precise the degree to which individual measurements cluster around the mean

proportional error difference between a measurement result and an accepted value, expressed as a fraction of the accepted value; not defined by BIPM

proportional uncertainty relative uncertainty

qualitative non-numerical data; descriptive information

quantitative numerical data; specific amount

random error a variation that affects a measurement in a random way so that the measurement is as likely to change in any one direction as in any other

raw data original data taken directly from a measurement system

relative uncertainty ratio of uncertainty to value

representation model of reality

significant figure digit reported in a measurement result; the number of significant figures is the number of meaningful digits in a measurement result

systematic error an error that acts to give a consistent offset in data; for example, a zero error

true value the exact value of a measurand; the 'true value' is an ideal that can never be known with certainty

uncertainty estimate of the range of values within which the 'true value' of a measurement or derived quantity lies; the extent to which the result of an experiment is unknown or unpredictable

zero error scale is not zero when measurements are taken; also called a calibration error

CHAPTER 12 SCIENTIFIC INVESTIGATIONS

By the end of this chapter you will have covered the following material.

Science Inquiry Skills

- Identify, research and construct questions for investigation; propose hypotheses; and predict possible outcomes (ACSPH078 AND ACSPH114)
- Design investigations, including the procedure to be followed, the materials required, and the type and amount of primary and/or secondary data to be collected; conduct risk assessments; and consider research ethics (ACSPH079 AND ACSPH115)
- Conduct investigations, including
  - the manipulation of force measurers and electromagnetic devices (ACSPH080)
  - use of simulations and manipulation of spectral devices (ACSPH116)

  safely, competently and methodically for the collection of valid and reliable data
- Represent data in meaningful and useful ways, including using appropriate SI units, symbols and significant figures; organise and analyse data to identify trends, patterns and relationships; identify sources of uncertainty and techniques to minimise these uncertainties; utilise uncertainty and percentage uncertainty to determine
  - the uncertainty in the result of calculations (ACSPH081)
  - the cumulative uncertainty resulting from calculations (ACSPH117)

  and evaluate the impact of measurement uncertainty on experimental results; and select, synthesise and use evidence to make and justify conclusions
- Interpret a range of scientific and media texts, and evaluate processes, claims and conclusions by considering
  - the accuracy and precision of available evidence (ACSPH082)
  - the quality of available evidence (ACSPH118)

  and use reasoning to construct scientific arguments
- Select, construct and use appropriate representations, including text and graphic representations of empirical and theoretical relationships
  - vector diagrams, free body/force diagrams, field diagrams and circuit diagrams (ACSPH083)
  - simulations, simple reaction diagrams and atomic energy level diagrams (ACSPH119)

  to communicate conceptual understanding, solve problems and make predictions
- Select, use and interpret appropriate mathematical representations, including linear and non-linear graphs and algebraic relationships representing physical systems, to solve problems and make predictions (ACSPH084 AND ACSPH120)
- Communicate to specific audiences and for specific purposes using appropriate language, nomenclature, genres and modes, including scientific reports (ACSPH085 AND ACSPH121)

Introduction

Performing investigations is your chance to experience what doing science is really like. Science is about finding things out through observation and experiment, which is what doing investigations is all about. This is why investigations are central to science, *and* why they are so much fun.

Figure 12.1 ▲ Student working on an investigation into rocket propulsion

Sometimes an important advance in science begins with a casual observation or a lucky accident. This was the case when Davisson and Germer first observed electron diffraction after an accident with a vacuum system. This eventually led to advances in quantum theory and to devices such as electron microscopes. However these developments could not have followed if Davisson, Germer and many others had not carried out further investigations. This sort of lucky accident may begin a new field of research, but it then proceeds by carefully planned investigation.

Scientific investigations can take years to complete and may involve collaboration among many scientists. They may require access to special equipment in Australia or overseas. They may cost a lot of money, sometimes millions of dollars, to complete. Hence scientists invest time in *planning* investigations before they begin. When scientists apply for grants to carry out investigations they need to show that they have carefully planned what they will do and how any money provided will be spent. Good planning is crucial to the success of the investigation.

They then make careful *measurements and observations* and record their *results*. They *keep records* of all their experiments. This is a legal requirement. Typically experimental results need to be kept for 5–7 years. There are also requirements on how and where data is stored.

Once data is collected it needs to be *analysed*. There are various ways this is done, but in the physical sciences (physics, chemistry and geology) it almost always involves constructing graphs. Once a relationship is established graphically, an algebraic relationship can be derived.

Finally, the results of the investigation must be *communicated*. Usually this involves publishing a scientific paper either in a journal or conference proceedings. It often includes presenting the results in talks or posters at conferences. If the result is funded by a grant then a research report must be submitted. If the results are really exciting, then the scientists may write a media release. However the results are communicated, this step must happen for the investigation to be completed.

When *you* perform scientific investigations you will also need to plan carefully. You will find out about the topic by reading about what other people have done. You will collect data from your own experiments and secondary sources. You will then analyse that data and draw conclusions about what it means. Finally, you will communicate your findings. Each of these steps is outlined below.

Planning your investigation

There are many things to consider when planning an investigation. You need to think about how much time you will have inside and outside class. You will also need to think about what space and equipment you will need and where you will go if you want to make measurements or observations outside.

You may be working in a group or on your own. Most scientists work in groups. If you can choose who you work with, think about it carefully. It is not always best to work with friends. Think about working with people who have skills that are different from your own.

Finally, and probably the first thing that most students think about, is the topic of the research. You will need to come up with a **research question** or **hypothesis**.

Choosing a research question

▲ **Figure 12.2**
Brainstorm as many ideas as you can in your group.

Obviously, it is a good idea to investigate something that you find interesting. If you are working in a group try to find something that is interesting to everyone in the group.

A good way to start is by 'brainstorming' for ideas. This works whether you are working on your own or in a group. Write down as many ideas as you can think of. Don't be critical at this stage. Get everyone in the group to contribute and accept all contributions uncritically. Write every idea down.

After you have run out of ideas, it is time to start being critical. Decide which questions or ideas are the most interesting. Think about which of these it is actually possible to investigate given the time and equipment available. Make a shortlist, but keep the long list too for the moment. Once you have your shortlist it is time to start refining your ideas.

Researching and refining your question

The next step is to find out what is already known about the ideas on your list. Use the Internet, your text books and the library to find out. Make sure you *keep a record* of the information that you find as well as *the sources*. You should start a **logbook** at this stage. You can write in references, or attach printouts to your logbook. This can save you a lot of time later on! Many research students forget to do this when they first start reading about their topic and then have to search all over again.

> **Good record keeping is important in scientific research, and it begins at this stage of the investigation.**

▲ **Figure 12.3**
Start researching your topic and make sure you keep a record of all your references.

Be critical of what you read. Do not assume that everything you read online or even in books is true. Try to find **reliable** sources of information. Textbooks, websites from universities and government research agencies are usually very reliable. Publications and web pages from professional associations, such as the Australian Institute of Physics and equivalent international organisations are also good sources. Blogs and homepages of other students are not usually reliable, although they are useful to give you ideas. Websites that are trying to sell you something should also be treated sceptically. Talk to your teacher about sources of information as well. They will be able to tell you if a website is reliable, and suggest sites that they know are suitable.

You may find examples of similar investigations to the one you are thinking of. It is a good idea to look at these, so you can learn from the experience of other researchers. However, in general, it is better not to try to replicate someone else's investigation exactly. If you do decide to replicate someone else's investigation then you need to acknowledge and carefully **reference** their work. (See the section on referencing on page 394). If you do not do so, it is **plagiarism**. This is a very serious form of academic misconduct. Talk to your teacher about how original your research needs to be, and how closely it can be based on someone else's work. It is much better to do this at the start than to be accused of cheating later on!

Finally, talk to your teacher about your ideas. They will be able to tell you whether your ideas are likely to be possible given the equipment available. They may have had students with similar ideas in the past and can make suggestions.

After you have researched your questions and ideas, you will hopefully be able to narrow the shortlist down to the one question that you want to tackle. If none of the questions or ideas look possible (or still interesting), then you need to go back to the long list.

Proposing a research question or hypothesis

Once you have decided on what you will investigate you need to turn it into a research question or a hypothesis.

A research question is one that can be answered by performing experiments or making observations. A hypothesis is a prediction of the results of an experiment, which can be tested by performing experiments or making observations.

You may also be able to do a design, build and test project. These are described later. Make sure you check what sort of investigation or project you are supposed to be doing with your teacher.

AUSTRALIAN INSTITUTE OF PHYSICS

This site contains a lot of useful information on physics-related matters.

AMERICAN INSTITUTE OF PHYSICS

This is a useful resource for keeping up with physics news.

INSTITUTE OF PHYSICS

This is another useful resource on physics-related matters.

Research questions

A research question may be of the form 'How does the height attained by a bottle rocket depend upon the volume of water in the rocket?'. The aim of your research is then to answer the question.

You need to frame the question carefully. It needs to be specific enough that it guides the design of the investigation. A specific question rather than a vague one will make the design of your investigation much easier. Asking 'what volume of water gives the maximum height for a water rocket?' tells you what you will be varying and what you will be measuring. It also gives a criterion for judging whether you have answered the question.

Asking 'How can we make a water rocket fly the best?' is not a good question. This question does not say what will be varied, nor does it tell you when you have answered the question. 'Best' is a vague term. What you mean by 'best' may not be what someone else means.

A good research question identifies the **variables** that will be investigated. Usually you will have one **dependent variable** and one **independent** or **controlled variable**. For a lengthy investigation you may have two or more independent variables. Variables are discussed in more detail later.

Finally, a good research question should be answerable with the time and equipment available.

Alamy/Marmaduke St John

Figure 12.4 ▲ You need to develop the question you are researching very carefully. These students are investigating the launch angle at which their water rocket will achieve maximum distance.

Hypotheses

A hypothesis is a tentative explanation or prediction not yet confirmed by experiment, such as 'The height attained by a water rocket will increase with the amount of water contained in the rocket'. Your hypothesis should give a prediction that you can test, ideally quantitatively.

A hypothesis is usually based on some existing model or **theory**. It is a prediction of what will happen in a specific situation based on that model. For example, kinematics describes the trajectory of any projectile. A hypothesis based on the kinematics model predicts the range of a specific projectile launched at a given angle and speed.

A hypothesis should give you a prediction that you can test by performing an experiment. This means it should at least be **falsifiable**. A good hypothesis should be able to be disproved. However, you will *not* generally be able to claim that you have proved your hypothesis.

If your experiments agree with predictions based on your hypothesis, then you can claim that they support your hypothesis. This *increases your confidence* in your model, but *it does not prove that it is true*. Hence an aim for an experiment should not start 'To prove …', as it is not possible to actually prove a hypothesis, only to disprove it.

If your experimental results disagree with your hypothesis, then you may have disproved it. This is *not* a bad thing! Often the most interesting discoveries in science start when a hypothesis based on an existing model is disproved. This means that the model it was based upon either is not a good model, or does not apply to the particular situation. You could then try to work out why the model does not apply, or try to formulate a better model. What to do when your hypothesis is not supported is discussed further in the analysis section.

In summary, a good research question is a question that is specific and can be answered by performing experiments and making measurements. A good hypothesis is a statement that predicts the results of an experiment and can be tested using measurements.

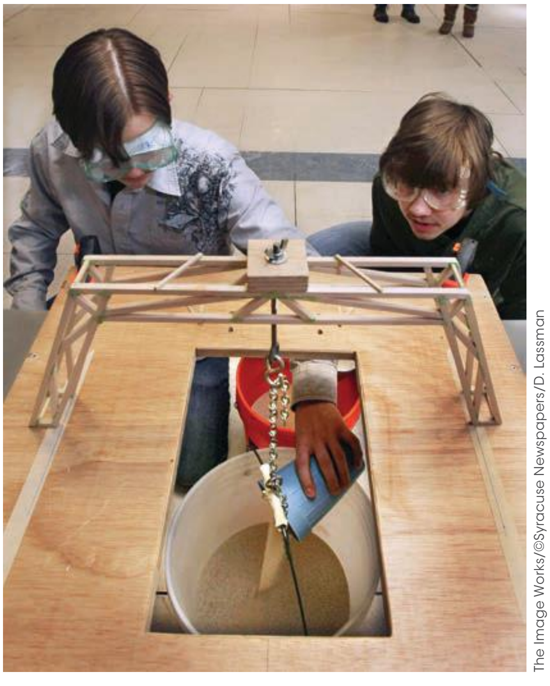
Even if your question or hypothesis meets these criteria, do not be surprised if you change or modify it during the course of your investigation. In scientific research, the question you set out to answer is often only a starting point for more questions.

Design briefs

Sometimes the aim of the investigation may be to design, build and test something. In this case, rather than a research question or a hypothesis, you will need a **design brief**. The design brief will specify what you intend to build, and some criteria by which to judge whether you have succeeded. You need a clear aim, and quantitative criteria by which to judge the final product.

The criteria will usually specify minimum performance characteristics of the end product. For example, if it is a water rocket, it may need some minimum range. A model bridge may need to span some distance and carry some minimum weight without breaking. The criteria may also include limits on what may be used or on the specifications of the product. For example, there may be a maximum weight or cost.

Just like a good research question or hypothesis, a well-written design brief tells you what to do, and how to know if you have succeeded.


The Image Works/©Syracuse Newspapers/D. Lassman

▲ **Figure 12.5** Testing how much weight a model bridge can carry

Designing your investigation

Once you have a specific research question or hypothesis you need to design your investigation. It is fun to start making measurements or observations immediately, but it is also important to spend time learning how to use the equipment, and experimenting to find the best way to set up your investigation. You may also discover that you need different or more equipment. This may save you time later on.

It is also important not to get distracted playing and forget the purpose of your investigation. At the end of the process, you need good data that answer your question or test your hypothesis. Having a plan allows you to ensure that you make the measurements that you need. The longer the investigation, the more important it is that you have a clear plan. There are several things to consider.

- What data will you need to collect?
- What materials and equipment will you need?
- When and where will you collect the data?
- If you are working in a group, who will collect the data?
- Who will be responsible for record keeping?
- How will the data be analysed?

The data that you collect will always include **secondary data**, and will usually include **primary data**. Secondary data is data that has been collected by someone else.

You will already have collected some secondary data when you investigated your research topic to formulate your question or hypothesis. You will probably want to collect more secondary data. If your topic is not one for which you can collect primary data, then you will need to rely on secondary data. Remember that when you collect secondary data it is important to use reliable, reputable sources.

Primary data is data that you collect yourself. You can collect data by performing experiments or making observations in the field. You should be able to measure distances, times, temperatures, forces including weights, potential differences and currents. You will have had practice at measuring some or all of these things already. You need to decide which variables you will measure and which variables you will control. Consider which variables you can control, and which you cannot.

If you are working on a design-build-test project then your primary data includes the performance tests of your product.

Consider how you will analyse the data. Will you need access to specific software such as a graphing or statistics package? If so, make sure that you know how to use it. If you are using software to draw graphs then you need to know how to produce a scatter graph and fit a **line of best fit** and add **uncertainty bars**. Note that a line of best fit is *not* the same as joining the dots. You should *never join the dots*, even though this is often the default setting in spreadsheet software. You should consult a reference guide, the 'help' menu for your software, or ask your teacher. Graphs are discussed in more detail in the analysis section on page 341.

Keep a record of your planning. This should go in your logbook. Writing down what you plan to do, and why, will help you stay focused during the investigation. If you are working in a group, then a record of what each person agrees to do during the investigation can be very important.

Variables and measurements

Anything that can vary in an experiment is a variable. An independent or controlled variable is one whose value you can control. For example, if you are doing an experiment to measure the voltage–current characteristics of an unknown circuit, then you would control the potential difference input into the circuit and measure the current in the circuit as a result. In this case the potential difference (or voltage) is the independent variable. The current, which is what varies as a result of the independent variable changing, is the dependent variable.

In the question 'What volume of water gives the maximum height for a water rocket?', the volume of water is the independent variable. The dependent variable is the height attained. A second possible independent variable, not mentioned in this question, is the air pressure inside the bottle. The air pressure should be kept constant, so it is not a variable in the investigation. If it was a long investigation, the air pressure could be a second controlled variable. If you decide to have two independent variables then it is important to keep one constant while you vary the other, if at all possible. Then you take multiple sets of measurements, keeping one variable at a fixed value for each set of data while you vary the other.

When variables have a numerical value, you make **quantitative measurements**. You measure that numerical value in the appropriate units. For example, you may measure a current of 15 mA or a height of 15 m.

Continuous variables may take any possible value, usually within some range. Length, time and current are continuous. In the water rocket example, the volume of water is a continuous variable, as it may take any value up to the volume of the bottle. A variable that may take only fixed values is called a **discrete variable**. Often these are whole numbers of things that cannot be broken into smaller parts, such as electrons or students. In the water rocket example, if the rocket has stabilising fins then the number of fins is a discrete variable.

Your measuring equipment will sometimes restrict you to only measuring discrete values. This is always the case with **digital** equipment. A set of digital scales that measures in grams gives you discrete values. It does not, however, mean that mass and weight are discrete variables. The weight of water in a bottle rocket is a continuous variable, but digital scales will only give you discrete measurements of the mass.

In some investigations you may use **qualitative measurements** or data. For example, a chemical reaction may lead to a colour change. You would usually describe the colour in words, such as 'pink' or 'green', rather than using a number. Sometimes you use a combination of qualitative and quantitative data. For example, you may describe the flight of a water rocket as reaching some maximum height in metres (quantitative) but following a spiralling path (qualitative).

Once you have decided on the variables you will be measuring you will be able to identify the equipment and other resources you will need.

Identifying the resources required

If you are going to collect primary data, make a list of all the equipment that you need. Consider how precise the measurements will need to be. If your hypothesis predicts a temperature change of 0.1°C, but you can only measure to a precision of 0.5°C, then you will not be able to test your hypothesis. You may need to think carefully about how you measure some things. For example, in a water rocket investigation, measuring the height attained can be very difficult. You may not be able to measure it directly.

If you are doing a design-build-test project then make sure you list all the materials and tools you will need. It is a good idea to allow for some extra materials in case mistakes are made during construction. For example, if you are testing a water rocket, the rocket could get stuck in a tree or lost. Consider who will supply the materials and how much they might cost. Scientists and engineers generally have tight budgets that they have to work within.

The equipment you plan to use must be safe. Will you need special protective equipment, such as lab coats, safety glasses or ear protectors? For a rocket project, you might need some temporary fencing or witches hats to mark off the launch area. There is a section on risk assessment on page 334. Make sure that you include any safety equipment needed in your equipment list.

Consider where you will perform your experiments or observations. Can you use normal classroom space, or do you need to be outside? If you are outside, what provisions can be made for ensuring that you can work without interference? Will you need to consider the convenience or safety of others? Talk to your teacher about what space is available.

When you have your list, talk to your teacher about what space and equipment is available. You might find that you need to modify your question or hypothesis at this stage.

Planning the experimental procedure

The most common problem that students have when doing research is time management. It is important to plan to have enough time to perform the experiments, *and* to analyse them, *and* to report on them. You also need to allow time to learn how to use the equipment if you have not used it before.

If you are doing a design-build-test project then allow plenty of time for all three stages. Often in this sort of project the design needs to be modified, sometimes several times. Building may not go as smoothly as you expect, and problems are often discovered during testing. If you allow plenty of time and start work early, you will be able to go back and modify your design if needed.

In any investigation you will need to collect reliable and precise data. You cannot do this if you do not know how to use the equipment. Always ask if you are unsure. Reading the user manual is also a good idea. It will usually specify the precision of the device, and let you know of any potential safety risks.

Whenever possible you should make repeat measurements, so allow time for this. This allows you to check that your measurements are **valid**. Valid results require that each independent variable gives similar results each time. If the results are similar each time, then your results are likely to be valid. If a result is not **reproducible**, it is probably not a valid result. A result is reproducible if you make exactly the same measurement more than once and get the same result, within the limits of experimental uncertainty. If a result is not reproducible, then a variable other than the one you are controlling is affecting its value. If this is the case, you need to determine what this other variable is, and control it if possible.

Think about how you can minimise uncertainties. Minimising uncertainty is not just about using the most precise equipment you can find, it is also about clever experimental technique. Very precise measurements are possible using simple equipment. For example, in 1862 Léon Foucault measured the speed of light with an uncertainty of 0.2%, without a computer, data logger or even

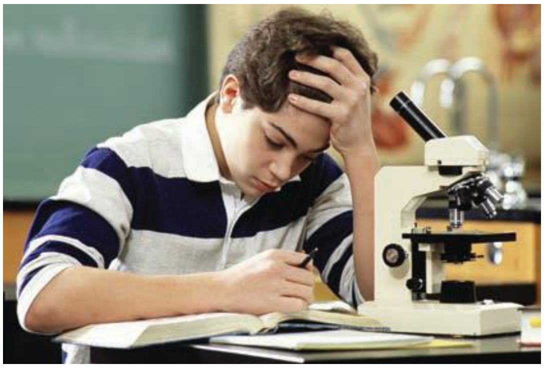


Figure 12.6 ▲ Sometimes experiments just don't work.

a digital stopwatch. Remember that it is a poor workman who blames his tools! See the section on uncertainties on page 337, and also in Chapter 11.

Sometimes experiments simply don't work or can't be done for some reason such as equipment failure or bad weather. A water rocket is hard to launch in high winds, and not pleasant to use in heavy rain. Try to think of all the things that could go wrong. If possible, come up with backup plans. Allowing plenty of time helps with this, as does starting your experiments as soon as possible.

Make sure you allow time for analysis. Ideally, do as much analysis as you can while you collect results. If you plot graphs as you take measurements, then you will be able to identify **outliers** early. An outlier is a data point that does not fit the pattern of the rest of the data. If you identify an outlier while you still have access to equipment and space, you can check the measurement and make sure that you didn't make a mistake.

After you have analysed your results, you need to write your report or communicate your findings in some other form. You need to plan ahead how this will be done. If you are working in a group, who will write which part of the report and when? Who will proofread it? Who will be responsible for making sure all the parts fit together?

You may find a timeline useful. A timeline helps keep you on track, and reminds everyone of their responsibilities. If you are working in a group get everyone to agree on it.

You can use the following table as a template.

Date and place	What will be done	By whom	Outcomes

Risk assessment

You may be required to complete a risk assessment before you begin your investigation. Even if this is not a requirement, it is a good idea to think about it. You need to think about three things.

1 *What are the possible risks* to you, to other people, to the environment or property?
2 *How likely is it* that there will be an injury or damage?
3 If there is an injury or damage to property or environment, *how serious are the consequences* likely to be?

A 'risk matrix', such as Table 12.1, can be used to assess the severity of a risk associated with an investigation. The consequences are listed across the top from negligible to catastrophic. Negligible may be getting clothes dirty or a very minor injury such as a scratch. Marginal might be a bruise from falling off a bike, or a broken branch in a tree. Severe could be a more substantial injury or a broken window. Catastrophic would be a death or the release of a toxin into the environment. In general, you need to ensure that your investigation is low risk. You can use a risk matrix either for individual identified risks, or for the investigation overall. If there are multiple experiments, then you would use a risk matrix for each one.

Table 12.1 Risk matrix for assessing for severity of risk

Consequences→ Likelihood↓	Negligible	Marginal	Severe	Catastrophic
Rare	Low risk	Low risk	Moderate risk	High risk
Unlikely	Low risk	Low risk	High risk	Extreme risk
Possible	Low risk	Moderate risk	Extreme risk	Extreme risk
Likely	Moderate risk	High risk	Extreme risk	Extreme risk
Certain	Moderate risk	High risk	Extreme risk	Extreme risk

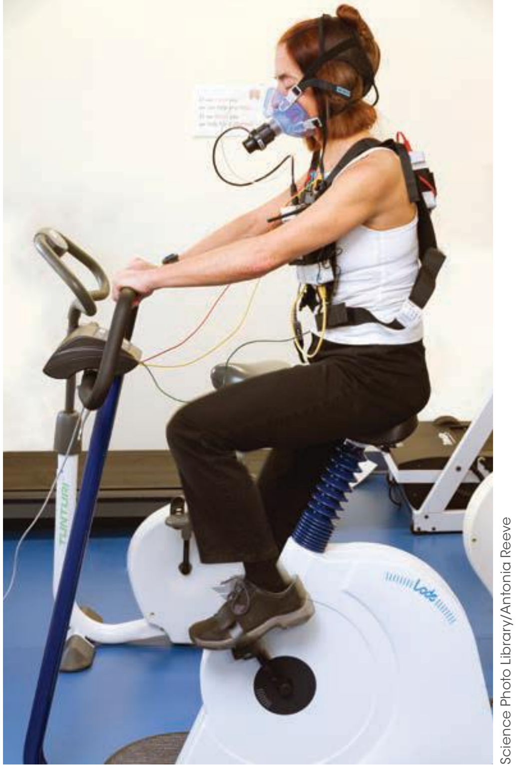
Once you have considered what the possible risks are, you need to think about what you will do about them. What will you do to minimise them, and what will you do to deal with the consequences if something does happen? This may be as simple as, 'Keep everyone at least 10 m away from the water rocket launch site.' You can use a risk assessment table like the one shown.

What are the risks in doing this experiment?	How can you manage these risks to stay safe?
Water from the rocket may be spilled on the ground and someone might slip.	Clean up all spills immediately. Keep the area around the rocket clear.

Ethics

Ethics in research can be controversial. More than one scientist has lost their job for unethical research behaviour. Being ethical in your research has two aspects. The first is about being honest as a scientist. This means recording data accurately, and not ignoring, hiding or changing any data that doesn't support your hypothesis. It means acknowledging and referencing sources of information including books, websites, articles and people who have helped you. It means not using other people's ideas or data without their knowledge and permission. Put simply, it is showing integrity or 'doing the right thing'. A good rule is that if you wouldn't want someone to know what you are doing, you probably shouldn't be doing it. It is no different from behaving ethically in any other area of your life.

The other aspect to ethics is treating animals, other people and the environment with care and respect. If your investigation will be using humans or animals, then you need to make sure you do not harm them, either physically or psychologically. If you are working with animals, then you need to make a strong case for any investigation that harms or could potentially harm them. When scientists want to use humans or animals in their research, they need to be able to show that the benefits to the environment, other animals or humans significantly outweigh the negative effects on the animals or humans used. The National Health and Medical Research Council (NHMRC) has guidelines on the ethical use of humans and animals in experimentation.



Science Photo Library/Antonia Reeve

▲ **Figure 12.7**
When working with humans you need to make sure that they are not harmed in any way.

NHMRC

You can find NHMRC human and animal ethics guidelines here.

Collecting your data

Once you have planned what you are going to do, collected your equipment and set it up, it is time to start collecting data. This is usually the fun part of any investigation. Don't forget you have a question to answer or a hypothesis to test! To do this, you need to make sure that you think carefully about what you do and keep good records. Good logbook records are just as important for design-build-test projects. Make sure you carefully record the process, including all the changes to the design that you make when you come to actually build it.

Record keeping – your logbook

You will need to keep a record of what you do during your investigation. You do this in a logbook. Even if you are using data loggers to collect your results and doing your analysis electronically, you should still keep a hard-copy logbook.

Scientists keep a logbook for each project that they work on. It is a record of what they did, why they did it, and what they found out. A logbook is a legal document for a working scientist. If someone's work is called into question, then the logbook acts as important evidence. Every entry in a scientist's logbook is dated, records are kept in indelible form (pen, *not pencil*), and entries may even be signed. Scientists' logbooks include details of experiments such as methods and results. They include comments and ideas, thoughts about the experiments and

iStockphoto/SteveStone

Figure 12.8 ▲
Make sure you keep an accurate record of what you do *as you do it.*

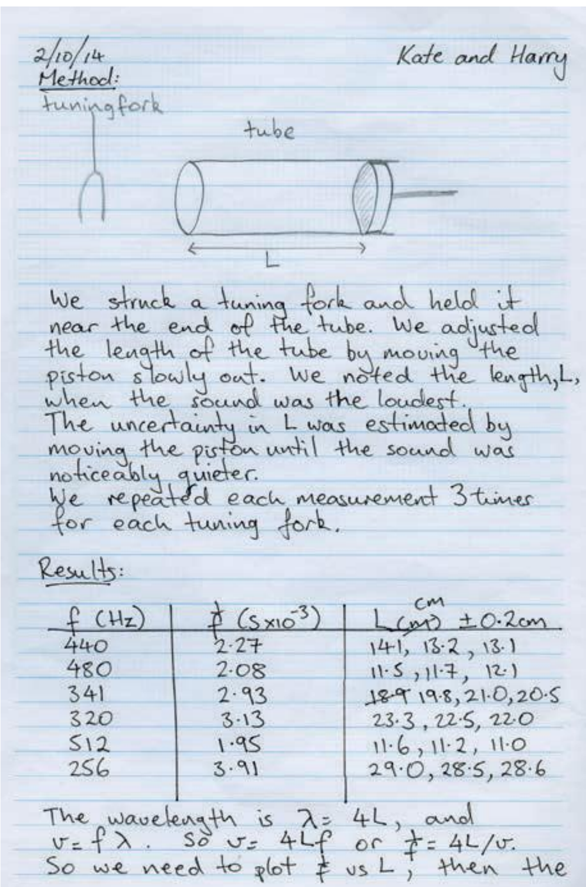
analysis. They frequently include printouts of data, photocopies of relevant information, photos and other items. The logbook is the primary source of information when a scientist writes up their work for publication. Logbooks are even provided as evidence in court cases sometimes; for example, in patent disputes or when a researcher is accused of falsifying data or stealing someone else's results.

Some scientists keep their research records electronically, but most experimental scientists still keep a hardcopy logbook. There are several advantages to a hardcopy logbook over an electronic one. First, electronic records are easy to make changes to, and it is hard to track what was changed, when and by whom. Second, if you are working in a group, it can be hard to keep track of who has the most recent version of the file(s). Third, files can be easily deleted or corrupted. It takes much more care and discipline to maintain a good electronic logbook than a good hardcopy. Remember that the purpose of a logbook is to record and maintain evidence of what you did. Electronic evidence is not as reliable as a signed hardcopy document.

You should talk to your teacher about what form of logbook records they require you to keep.

If you are working in a group then you will need to decide whether to keep one logbook for the entire group, or one each. If you will all be working in the same places at the same times, then one for the whole group is best. If you will be in different places (e.g. doing field observations), then you will need one each. Your teacher may also require each of you to keep your own logbook for assessment, or for authentication purposes.

Your logbook is a detailed record of *what you did* and *what you found out* during your investigation. Make an entry in the logbook *every* time you work on your investigation. At the start of each session you should record the date and the names of all the people with whom you are working at the time.



2/10/14 Kate and Harry

Method:

We struck a tuning fork and held it near the end of the tube. We adjusted the length of the tube by moving the piston slowly out. We noted the length, L, when the sound was the loudest. The uncertainty in L was estimated by moving the piston until the sound was noticeably quieter. We repeated each measurement 3 times for each tuning fork.

Results:

f (Hz)	T (s×10⁻³)	L (~~m~~ cm) ±0.2cm
440	2.27	14.1, 13.2, 13.1
480	2.08	11.5, 11.7, 12.1
341	2.93	~~18.9~~ 19.8, 21.0, 20.5
320	3.13	23.3, 22.5, 22.0
512	1.95	11.6, 11.2, 11.0
256	3.91	29.0, 28.5, 28.6

The wavelength is $\lambda = 4L$, and $v = f\lambda$. So $v = 4Lf$ or $T = 4L/v$. So we need to plot T vs L, then the

Write down what you do as you do it. It is easy to forget what you did if you do not write it down immediately. An accurate record is important if you need to repeat any measurements or if you get unexpected results.

Write down the names, model and serial numbers of any equipment used.

Include large, clear diagrams of any experimental set-up. Label all the parts or pieces of equipment. You can also include photos of experiments. Include large and clearly labelled circuit diagrams of all circuits that you use. Diagrams in your logbook do not need to be neat, but you must be able to understand the diagrams later on.

If you are doing a design-build-test project, then include diagrams of the build process. Flowcharts are very useful for this. Take lots of photos during the building and testing stages. Print these out and attach them to your logbook, and make a note of where the files are kept. These will be useful when you write your report later.

Record the results of *all* measurements *immediately and directly into your logbook, in pen*. *Never* record data onto bits of scrap paper instead of your logbook! Results must be recorded in indelible form. This means using a pen. Never write your results in pencil. Never use white-out or scribble over anything in your logbook. If you want to cross something out, just put a line through it. It is also a good idea to make a note explaining why it was crossed out.

◀**Figure 12.9**
A page from a student's logbook

A good logbook contains:
- notes taken during the planning of your investigation
- a record of when, where and how you carried out each experiment
- diagrams showing the experimental set-ups, circuit diagrams, etc.
- all your raw results
- all your derived results, analysis and graphs
- all the ideas you had while planning, carrying out experiments and analysing data
- printouts, file names and locations of any data not written directly into the logbook.

It is not a neat record, but it is a *complete* record.

Performing experiments

If you have planned carefully and learned how to use the equipment, then hopefully your experiments will go smoothly.

In a design-build-test project, this is the testing stage.

As stated above, *always* record results immediately, with the correct units and with their uncertainty. The raw data should *always* be recorded directly into the logbook unless it is recorded using data loggers connected to a computer. In this case a printout of the data should be attached to the logbook, and the file name and location recorded.

Make sure that you measure and record everything you will need for your analysis. For example, if you are investigating water rockets, you could measure the weight of the rocket, the air pressure used and the wind strength. It is much better to measure something and then discover that you didn't need to, than to start your analysis and realise that you didn't measure something that you do need.

Use SI units. This means metres (m) for lengths, seconds (s) for time and kilograms (kg) for mass. If you are measuring current, use amperes (A) and use volts (V) for potential differences. The uncertainties on raw data will be in the same units as the measurements. Always record these along with the measurements as you go. Do not try to add them in later.

See Chapter 11 for a discussion of SI units.

If you are going to be collecting multiple data points, then it is a good idea to draw a table to record them in. Label the columns in the table with the name and units of the variables. Do not put the units in the table cells. If you know that the uncertainty in all your measurements is the same, then you can record this in the heading cell at the top of the column as well. Otherwise, each data entry should have its uncertainty recorded in the cell with it. Remember that the uncertainty in the raw result should have the same unit as the result.

It is a good idea to start your analysis while you are collecting your data. If you spot an outlier and you are still making measurements then you have the opportunity to repeat that measurement. If you made a mistake, then put a line through the mistake, write in the new data, and make a comment in your logbook. Do not scribble out or remove mistakes, they may turn out to be useful.

If you have not made a mistake, then plotting and analysing as you go allows you to spot something interesting early on. You then have a choice between revising your hypothesis or question to follow this new discovery, or continuing with your plan. Many research projects start with one question and end up answering a completely different one. These are often the most fun, because they involve something new and exciting.

Estimating uncertainties

When you perform experiments there are typically several sources of uncertainty in your data.

Sources of uncertainty that you need to consider are the:
- limit of reading of measuring devices
- precision of measuring devices
- variation of the measurand.

Limit of reading

For all devices there is an uncertainty due to the limit of reading of the device. The limit of reading is different for **analogue** and digital devices.

Analogue devices include swinging needle multimeters, liquid in glass thermometers and clocks with hands. Analogue devices have continuous scales. For an analogue device, the **limit of reading**, sometimes called the **resolution,** is half the smallest division on the scale. We take it as half the smallest division because you will generally be able to see which division mark the indicator (needle, fluid level, etc.), is closest to. You may be able to estimate the measurement to one-fifth or even one-tenth of the smallest division if the spacing between divisions is large; however, the limit of reading uncertainty is still half of the smallest division. So, for a liquid in a glass thermometer with a scale marked in degrees Celsius, the limit of reading is 0.5°C.

Digital devices such as digital multimeters, clocks and thermometers have a scale that gives you a number. It is limited to a specific number of figures, typically three or four, so it is a discrete scale. A digital device has a limit of reading uncertainty of a whole division. So a digital thermometer that reads to whole degrees has an uncertainty of 1°C. For a digital device the limit of reading is *always* a whole division, not a half, because you do not know whether it rounds up or down, or at what point it rounds.

The resolution or limit of reading is the *minimum* uncertainty in any measurement. Usually the uncertainty is greater than this minimum.

Shutterstock.com/Jan van der Hoeven
Shutterstock.com/Ehrman Photographic

Figure 12.10 ▶ a) This digital scale has a limit of reading of 0.001 s. b) This analogue scale has a limit of reading of 0.1 s.

Precision of measuring device

The measuring device used will have a **precision**, usually given in the user manual. For example, a multimeter may have a precision of 0.5% on a voltage scale. This means if you measure a potential difference of 12.55 V on this scale, the uncertainty due to the precision of the meter is $0.005 \times 12.55\text{ V} = 0.06\text{ V}$. This is greater than the limit of reading uncertainty, which is 0.01 V in this case.

Many students think that digital devices are more precise than analogue devices. This is often not the case. A digital device may be easier for you to read, but this does not mean it is more precise. The uncertainty due to the limited precision of the device is generally greater than the limit of reading.

a

Courtesy of FLIR Commercial Systems

b

Function	Range	Resolution	Accuracy	
DC voltage	400 mV	0.1 mV	±(0.3% reading + 2 digits)	
	4 V	0.001 V	±(0.5% reading + 2 digits)	
	40 V	0.01 V		
	400 V	0.1 V		
	1000 V	1 V	±(0.8% reading + 3 digits)	
AC voltage			50 to 400 Hz	400 Hz to 1 kHz
	400 mV	0.1 mV	±(1.5% reading + 15 digits)	±(2.5% reading + 15 digits)
	4 V	0.001 V	±(1.5% reading + 6 digits)	±(2.5% reading + 8 digits)
	40 V	0.01 V		
	400 V	0.1 V		
	750 V	1 V	±(1.8% reading + 6 digits)	±(3% reading + 8 digits)
Frequency	5.000 Hz	0.001 Hz	±(1.5% reading + 5 digits)	
	50.00 Hz	0.01 Hz		
	500.0 Hz	0.1 Hz	±(1.2% reading + 2 digits)	
	5.000 kHz	0.001 kHz		
	50.00 kHz	0.01 kHz		
	500.0 kHz	0.1 kHz		
	5.000 MHz	0.001 MHz	±(1.5% reading + 4 digits)	
	10.00 MHz	0.01 MHz		
	Sensitivity: 0.8 V rms min. @20% to 80% duty cycle and <100 kHz; 5 V rms min. @20% to 80% duty cycle and >100 kHz			
Duty cycle	0.1 to 99.9%	0.1%	±(1.2% reading + 2 digits)	
	Pulse width: 100 µs – 100 ms, Frequency: 5 Hz to 150 kHz			

Note: Accuracy specifications consist of two elements:
- (% reading) – This is the accuracy of the measurement circuit.
- (+ digits) – This is the accuracy of the analog to digital converter.

Courtesy of FLIR Commercial Systems

▲ **Figure 12.11**
a) A typical small digital multimeter; b) A page from the user manual giving the precision on various scales

Variation of the measurand

The measurand itself may vary. For example, the flight of a water rocket is strongly dependent on initial conditions, wind and other factors. Even keeping launch conditions as close to identical as possible, it is unlikely that in repeat experiments you will be able to get a rocket to attain the same height within the limit of reading or equipment precision. Making repeat measurements allows you to estimate the size of the variation.

Sometimes you will be able to see how the measurand varies during a measurement by watching a needle move or the readings change on a digital device. Watch and record the maximum and minimum values. The difference between these is the range:

$$\text{Range} = \text{maximum value} - \text{minimum value}$$

The value of the measurand is the average value, or the centre of the range:

$$\begin{aligned}\text{Measurand} &= \text{minimum value} + \frac{1}{2}(\text{range}) \\ &= \text{minimum value} + \frac{1}{2}(\text{maximum value} - \text{minimum value})\end{aligned}$$

The uncertainty in the measurement is half the range:

$$\text{Uncertainty} = \frac{1}{2}\,(\text{range}) = \frac{1}{2}(\text{maximum value} - \text{minimum value})$$

For example, if you are using an analogue multimeter and you observe that the needle fluctuates between 12.2 V and 12.6 V then your measurement should be recorded as (12.4 ± 0.2) V. Note that the measurement and uncertainty are together in the brackets, indicating that the unit applies to both the measurement and its uncertainty.

For further discussion of uncertainties see Chapter 11.

When you take repeat measurements, the best estimate of the measurand is the average value. If you have taken fewer than 10 measurements then the best estimate of the uncertainty is half the range. If you have more than 10 measurements, the best estimate of the uncertainty is the standard deviation, given by:

$$\text{Standard deviation} = \left(\frac{\sum (x_\text{i} - x)^2}{n - 1}\right)^{\frac{1}{2}}$$

GUM

Read more about uncertainties in the GUM (Guide to expressions of Uncertainty in Measurement).

where x_i is an individual value of the measurand, x is the average value of the measurand and n is the total number of measurements. The sum is over all values of x_i. Most calculators have built-in statistical functions such as standard deviation. Spreadsheet software such as Excel also calculates functions such as standard deviation. Remember that repeat measurements means repeating under the same conditions. It is not the same as collecting lots of data points under different conditions.

Random and systematic errors

These sources of uncertainty all give rise to random errors. That means that repeated measurements will be randomly spread about the 'true value', and centred on that value.

You may also have systematic errors in your data. These typically occur when there is a calibration error, such as a zero error, in a measuring device. Always check that your equipment reads zero when you expect it to. For example, the volts and amps scales on a multimeter should read zero when the leads are not connected to anything. A weighing scale should read zero if no force is applied. If it shows some other value, then all your measurements will be out by this amount.

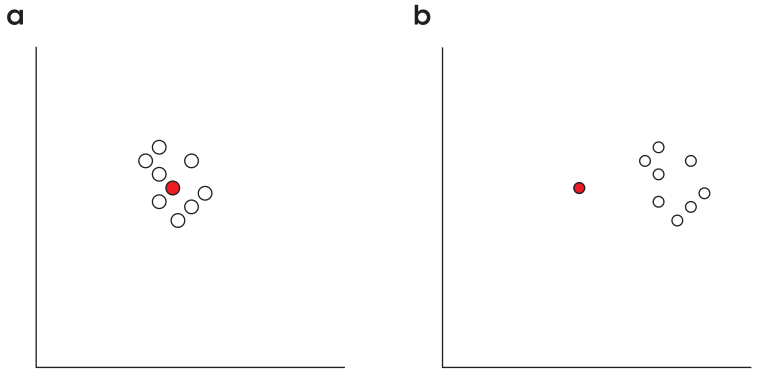


Figure 12.12 ▶ a) Results are clustered about the true value when the errors are random. b) Results are clustered about some other value as a result of systematic errors.

Analysing your data

Once you have collected your data you will need to analyse it. Record your analysis in your logbook. If this is done on a computer, then record the file name and location and attach a printout of the analysis into your book. Many scientists have logbooks that are bulging with printouts.

The first step is organising your data. This will usually involve tabulating it. Plotting graphs is a useful way to begin the analysis of your data. Graphs are a very useful way of representing data so that trends and relationships can be identified. There are many different sorts of graphs that can be used to organise and display data. These are described on the following pages.

You will usually need to do some calculations with your data to be able to answer your question or test your hypothesis. Remember to keep units on all quantities, so that any derived values have the correct units. You will also need to calculate uncertainties on any derived quantities.

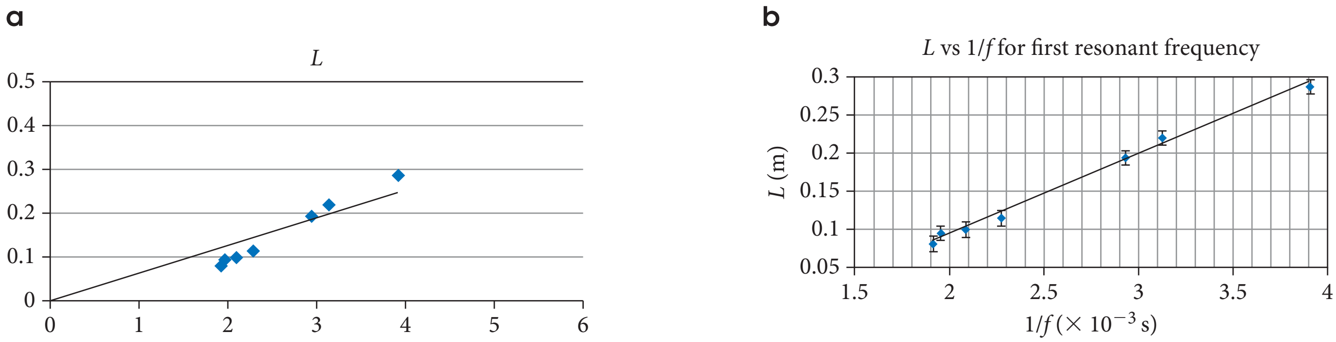
Organising your data

If you have more than a few data points then it is a good idea to display them in a table. You may have several tables, for different experiments. You may also need to do some analysis of the data. For example, imagine you have measured the temperature of different materials as a function of time after heating them. In this case you might have five tables of raw data, each with measurements of temperature as a function of time for a different material. A table summarising the data in some way will be useful. For example, a table showing the time taken for each material to drop from 55°C to 45°C may be useful. Alternatively, a table showing the change in temperature over a given time period may be more useful. You cannot have too many tables or graphs in your logbook. You can decide later, when you write your report, which are the most useful for communicating your results.

Identifying trends, patterns and relationships

You may be able to see a pattern simply by looking at a list of numbers in a table. However, the most reliable way to identify a pattern in data or a relationship between variables is to plot a graph. If you have a hypothesised equation then use it to generate a fit on a graph of your data, as described below. Do not substitute your data into your hypothesised equation and try to show that it fits.

A graph should be large and clear. The axes should be labelled with the names of the variables and their units. Choose a scale so that your data takes up most of the plot area. This will often mean that the origin is not shown in your graph. Usually there is no reason why it should be.



▲ **Figure 12.13**
a) A poor example of a graph; b) A good example of a graph of the same data. How many problems can you identify on the graph in part a?

When you are looking for a relationship between variables, plot a **scatter graph**. This is a graph showing your data as points. *Do not join them up as in a dot-to-dot picture*. Usually the independent variable is plotted on the x axis and the dependent variable goes on the y axis, unless there is a good reason to do otherwise. In the water rocket example, height would be on the y axis plotted against water volume on the x axis.

To determine a relationship you need to have enough data points and the range of your data points should be as large as possible. A minimum of six data points is generally considered adequate if the relationship is expected to be linear, but always collect as many as you reasonably can, given the available time.

For non-linear relationships you need more data points than this. Try to collect more data in regions where you expect rapid variation. Imagine you are measuring an interference pattern from two slits, as shown in Figure 12.14. You may need more than a hundred data points to clearly see the sinusoidal pattern due to the two slit interference.

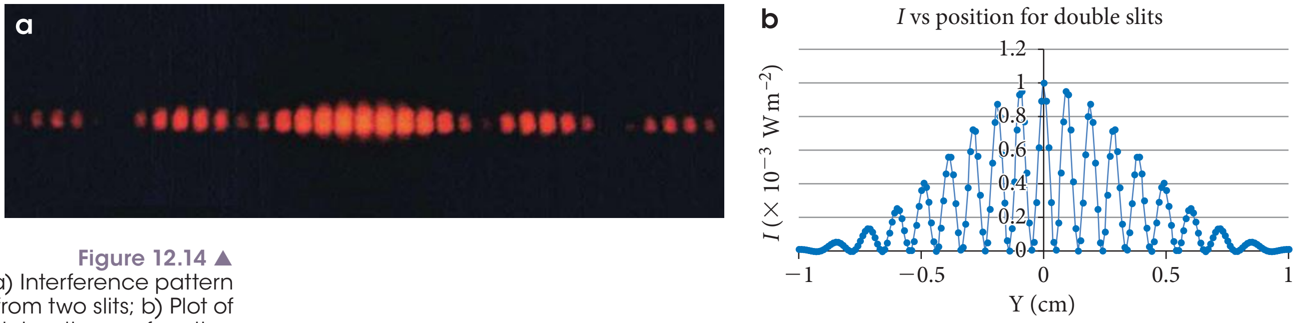


Image reproduced with permission of PASCO scientific, Inc.

Figure 12.14 ▲
a) Interference pattern from two slits; b) Plot of intensity as a function of position for this experiment

DATA POINTS

Some helpful advice on deciding the number of data points

A good graph to start with is simply a graph of the raw data. You will usually be able to tell by looking whether the graph is linear. If it is, then fit a straight line using a graphing package. You can then use a linear regression tool to check how good the straight line fit is. This will give you an R^2 number, which is a measure of 'goodness of fit'. The closer R^2 is to 1 (or –1), the better the fit. If it is not *very* close to 1, then the relationship is not linear. Alternatively, you can calculate the uncertainty in the gradient by using lines of maximum and minimum gradient. If the uncertainty is large, then the relationship may not be linear.

If it is a linear relationship, then finding the equation for the line of best fit may be useful. *Never* force a line of best fit through the origin. Often the intercept gives you useful information. It may even indicate a systematic error, such as a zero error in calibration of your equipment.

When you plot your raw data you may find that one or two points are outliers. These are points that do not fit the pattern of the rest of the data. These points may be mistakes; for example, they may have been incorrectly recorded or a mistake was made during measurement. They may also be telling you something important. For example, if they occur at extreme values of the independent variable then it might be that the behaviour of the system is linear in a certain range only. This is the case for materials under stress. You may choose to ignore outliers when fitting a line to your data, but you should be able to justify why.

When you extend a line of best fit beyond your measured points this is called **extrapolation**. Any data that you read off a graph outside the range of your data points is extrapolated, and should be viewed with caution. You cannot say for sure that the system continues to behave in the same way beyond the bounds of your data. For example, imagine you measure the spring constant of a spring by applying weights and measuring its extension. If you plot extension as a function of weight, you should get a straight line. You could in theory extrapolate your line of best fit to any weight. But you know that in practice if you continued adding weights you would eventually break the spring.

Reading points, other than data points, from a line of best fit within the region in which you have data is called **interpolation**. You cannot be sure that this is exactly what you would find if you measured that point. However, if your line of best fit really represents the behaviour of the system, then you can use interpolated points in your analysis.

Relationships between variables are often not linear. If you plot your raw data, for example height of rocket trajectory as a function of pressure, and it is a curve, then *do not draw a straight line through it*. In this case you need to think a little harder. If your hypothesis predicts the shape of the curve, then try fitting a theoretical curve to your data. If it fits well, then your hypothesis is supported.

If possible, you should **linearise** your data based on your hypothesis. Remember that linear graphs have equations of the form $y = mx + c$. Here y is the variable plotted on the vertical axis, usually the dependent variable. The independent variable x is the variable plotted on the horizontal axis. The gradient is $m = \Delta y/\Delta x$. The constant c is the y intercept.

For example, if your hypothesis is that $h = \frac{1}{2}gt^2$, try plotting your data as a function of t^2. Here h is the initial height of a falling object, g is the acceleration due to gravity and t is the time taken for it to fall:

$$\begin{array}{ccccc} h = & \left(\frac{1}{2}g\right) & t^2 & + & 0 \\ \uparrow & \uparrow & \uparrow & & \uparrow \\ y = & m & x & + & c \end{array}$$

Hence a plot of h vs t^2 should be a straight line with gradient $\frac{1}{2}g$ and a y intercept of zero. So if you plot h vs t^2 and get a straight line of gradient $\frac{1}{2}g$ with a y intercept of zero, then your hypothesis is supported.

A second example is radioactive decay. Your hypothesis could be that the activity of your sample at some time t is $A = A_0e^{-kt}$. If you measure A as a function of time, t, then a plot of $\ln(A)$ vs t will have a gradient of $-k$ and an intercept of A_0. You can see that this is the case if you take the natural logarithm (ln) of both sides:

$$\ln(A) = \ln(A_0e^{-kt}) = \ln(A_0) + (-kt)$$

which is again of the straight line form:

$$\begin{array}{ccccc} \ln(A) = & -k & t & + & \ln(A_0) \\ \uparrow & \uparrow & \uparrow & & \uparrow \\ y = & m & x & + & c \end{array}$$

It is better to linearise your data rather than to try fitting a curve to non-linear data. Often a curve for an exponential relationship can look very much like a curve for a power law. Linearising your data allows you to distinguish between the two.

Log–log graphs are useful for power laws. A log–log graph will give you a straight line if there is a power law relationship between the variables. For example, if the relationship is of the form $y = ax^n$, then if we take logarithms of both sides we get $\log(y) = \log(a) + n\log(x)$. A plot of $\log(y)$ vs $\log(x)$ then has gradient n and intercept $\log(a)$.

Remember that the argument of a logarithm is dimensionless. Hence we ignore the units of a, x and n during this process. It is important to remember to put them back at the end and make sure your final expression is dimensionally correct.

Consider the data shown in Table 12.2.

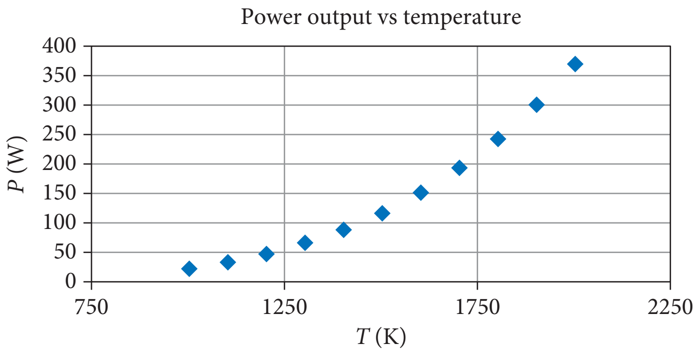
Table 12.2

Temperature (K)	Power (W)
1000	23
1100	34
1200	48
1300	67
1400	89
1500	117
1600	152
1700	194
1800	243
1900	301
2000	370

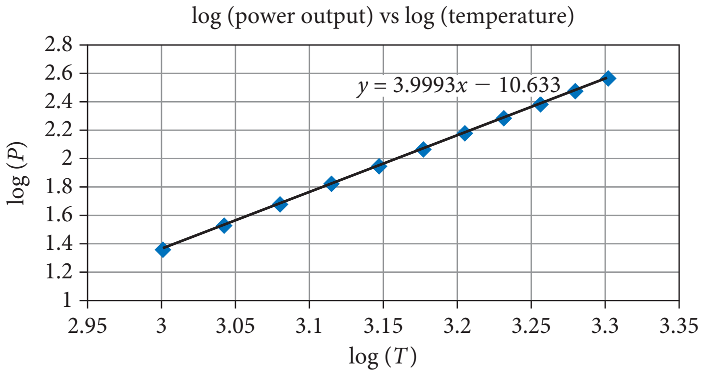
Plots of T vs P and log T vs log P are shown in Figure 12.15.

Figure 12.15 ▶
a) Direct plot of the data in Table 12.2; b) Log–log graph of the same data, with trendline displayed

a



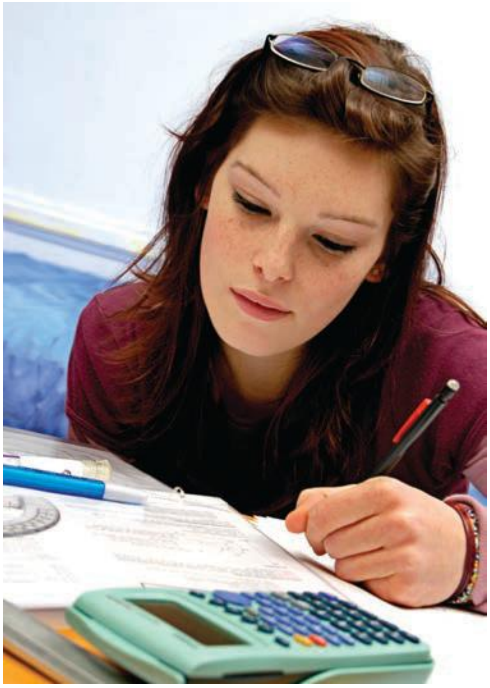
b



We can see from the equation for the line of best fit that the gradient is 3.9993 and the intercept is −10.633. Hence we deduce that the relationship between P and T is $P = (2.7 \times 10^{-11}\,\mathrm{W\,K^{-4}})\,T^{4}$. This is a powerful technique, and one well worth practising.

Sometimes you do not know what relationship to expect between variables, or you may have tried to fit your data and it has not worked. With spreadsheet software it is very quick to generate log–linear, log–log graphs and other plots, so it is worth trying a few.

Figure 12.16 ▼
You will usually need to analyse your raw data in some way.



Science Photo Library/AJ Photo

Sometimes the relationship between variables will be more complicated than a linear, exponential or power law. In this case, a graph is still useful but the most you might be able to say is that one variable increases with another, or that there is a peak at a particular position. This is likely to be the case with the water rocket example we have been using. A graph is still a useful way of identifying trends and patterns, even if you are not able to extract a mathematical relationship from the graph.

Performing calculations with your data

You will usually have to do some calculations with your data as part of your analysis. When you recorded your data you wrote down the units for all your measurements. You may need to convert these to SI units (e.g. cm to m). Include the units with all numbers as you do your calculations. In this way you will make sure you have the correct unit on all derived data. It also allows you to check that any equations you are using are dimensionally correct. For example, if you are using $h = \frac{1}{2}gt^2$ to estimate a maximum height obtained, then you should be using t in s and g in $\mathrm{m\,s^{-2}}$. If you measured time in ms, then you need to convert to s before calculating h. If you make a mistake and try to use $h = \frac{1}{2}gt$ instead, then if you have included units in your calculation you should notice that your answer for height is in $\mathrm{m\,s^{-1}}$. For example, if you measure $t = 3.0\,\mathrm{s}$ and use $g = 9.8\,\mathrm{m\,s^{-2}}$, then using the correct equation:

$$h = \frac{1}{2}gt^2 = \frac{1}{2}(9.8\,\mathrm{m\,s^{-2}})(3.0\,\mathrm{s})^2 = \frac{1}{2}(9.8\,\mathrm{m\,s^{-2}})(9.0\,\mathrm{s^2}) = 44\,\mathrm{m}$$

If you use the incorrect equation you get:

$$h = \frac{1}{2}gt = \frac{1}{2}(9.8\,\mathrm{m\,s^{-2}})(3.0\,\mathrm{s}) = \frac{1}{2}(9.8\,\mathrm{m\,s^{-2}})(3.0\,\mathrm{s}) = 15\,\mathrm{m\,s^{-1}}$$

The incorrect unit for h should alert you to an error, even if the difference in numerical value does not.

It is good practice *in general*, not just in investigations, to include units at each step in all your calculations.

Your raw data should be recorded with uncertainties. All your derived results should also have uncertainties. Chapter 11 shows how to calculate uncertainties on derived values. You can also refer to books on experimental methods and uncertainties in measurement for a more detailed description of how to treat uncertainties.

Interpreting your results

Once you have analysed your results you need to interpret them. This means being able to either answer your research question or state whether your results support your hypothesis. In a design-build-test project this is where you determine whether, or to what extent, your final product meets the requirements of the design brief.

You need to take into account the uncertainties in your results when you decide whether they support your hypothesis. For example, suppose you have hypothesised that the maximum range of a water rocket occurs at a launch angle of 45°. Your results show that the maximum range occurs at an angle of 47°. You may think that this result does not support your hypothesis. To say whether the result agrees with the prediction, you need to consider the uncertainty. If the uncertainty is 1°, then the results disagree with the hypothesis. If the uncertainty is 2° or more, then the results do agree and the hypothesis is supported.

You need to know the uncertainty to be able to interpret the results.

If your hypothesis is not supported

It is not enough to simply say 'our hypothesis is wrong'. If the hypothesis is wrong, *what* is wrong with it?

It may be that you have used a model that is too simple. For example, if you have based your hypothesis on the kinematics model and ignored the effect of air resistance, the range is likely to be shorter and the maximum height lower than you predicted. For many projectiles, including water rockets, air resistance is not negligible. If you find that this is the case, then you may conclude that your situation is better described by a model that includes air resistance.

Before you decide that the model is at fault, however, it is a good idea to check carefully that you have not made any mistakes or ignored any variables.

Think carefully about any factors that you did not take into account but which might have affected your experiment.

Go through your method, results and analysis. Check that your equipment was correctly calibrated, and that you were using it correctly. Check that data is recorded in the correct units, and that units are correctly carried through all calculations during analysis. Check your analysis carefully. If you are working in a group, get another person to repeat the calculations.

It is never good enough to conclude that 'the experiment didn't work'. Either a mistake was made or the model used was not appropriate for the situation. It is your job to work out which.

If you have done a design-build-test project and your final product does not meet the design brief then you should explain why. Try to determine whether the basic design was flawed, or if there was a problem with materials or something else. If you can, offer suggestions for a better design or process.

Communicating your results

If research is not reported on, then no-one else can learn from it. An investigation is not complete until the results have been communicated. Most commonly a report is written.

Writing reports

A report is a formal and carefully structured account of your research. It is based on the data and analysis in your logbook. However the report is a *summary*. It contains only a small fraction of what appears in the logbook. Your logbook contains all your ideas, rough working and raw data. The report typically contains none of this.

A report consists of several distinct sections, each with a particular purpose.

- Abstract
- Introduction
- Method
- Results and analysis
- Discussion
- Conclusion
- Acknowledgements
- References
- Appendices

Reports are always written in the past tense, because they describe what you have done.

The abstract

The abstract is a very short summary of the entire report. It is the most important part, because often it is the only part that people read. Typically an abstract is between 50 and 200 words long. It appears at the start of the report, but is always the last thing that you write. Try writing just one sentence to summarises each part of your report.

Introduction

The introduction tells the reader why you did the investigation and what your research question or hypothesis is. This is the place to explain why this research is interesting.

The introduction also provides any background information needed to be able to understand the rest of the report. This is the place to summarise any existing theories and models. You need to do this to justify your hypothesis. You should also summarise any similar investigations. All of this should be correctly referenced, as described in the section on referencing on page 349.

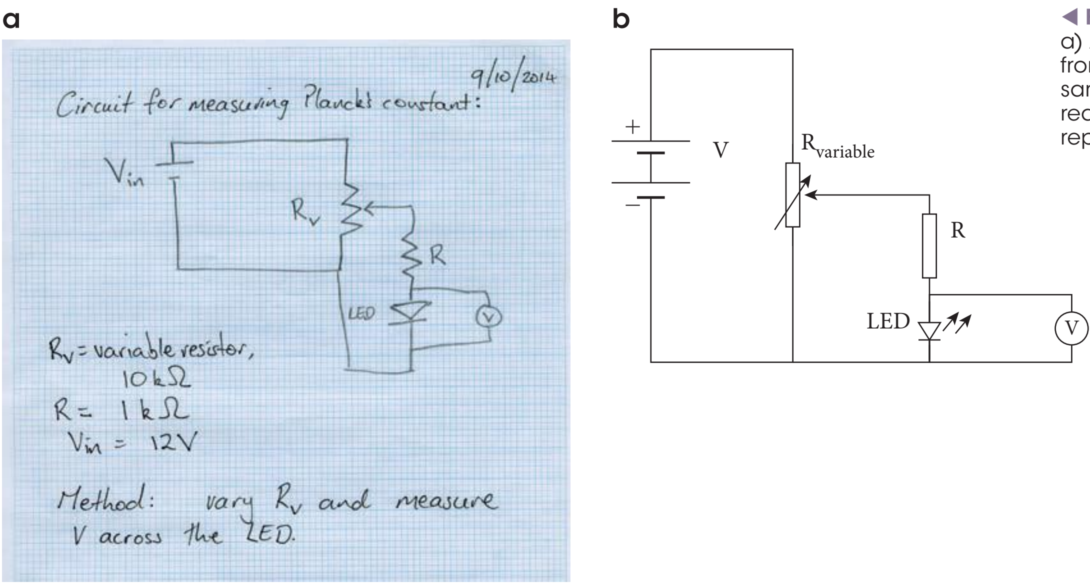
Method

The method describes what you did. It is not a recipe for someone else to follow.

The method summarises what you measured and how you measured it. It also explains, briefly, why you chose a particular method or technique.

Write your method using sentences, not dot points. Remember that these need to be written in the past tense – *it is not a recipe*. You are not commanding anyone to do anything. You are telling people what you did. For example, you would write 'we measured the height' not 'measure the height'.

Include any diagrams, such as circuit diagrams, needed to make your method clear. The diagrams in your logbook will usually be rough sketches (see Figure 12.17). The diagrams in your report should be very neat and carefully labelled. Flowcharts can be useful to describe any procedures in which a series of steps was followed. Each diagram should have a figure number and you should refer to it in the text of your report. Position the diagram close to where it is referred to in the text. You should take the time to learn how to position figures neatly using your word processor software.



◀ **Figure 12.17**
a) A circuit diagram from a logbook; b) The same circuit diagram redrawn in a formal report

Results and analysis

The results section is a *summary* of your results. It is usually combined with the analysis section, although they may be kept separate.

Avoid including tables of raw data in your report unless they compare the results of a few different experiments. For example, you would include a table showing the maximum height attained for a few different designs of water rocket. You would not include a table of raw data showing height attained for a water rocket for many different volumes of water. Wherever possible use a graph instead of a table.

If a table has more than a few rows of data, it is better to represent that data in some other way. Usually this will be a graph.

Think about what sort of graph is appropriate. If you want to show a relationship between two variables then use a scatter plot. Display your data as points with uncertainty bars and clearly label any lines you have fitted to the data. Always make sure you label your axes, including units. Choose an appropriate scale so that the data takes up most of the plot area.

Column and bar charts are useful for comparing two data sets, such as average height attained for different types of water rocket. *Do not* use a column or bar chart to try to show a mathematical relationship between variables.

Figure 12.18 gives examples of the two types of graphs.

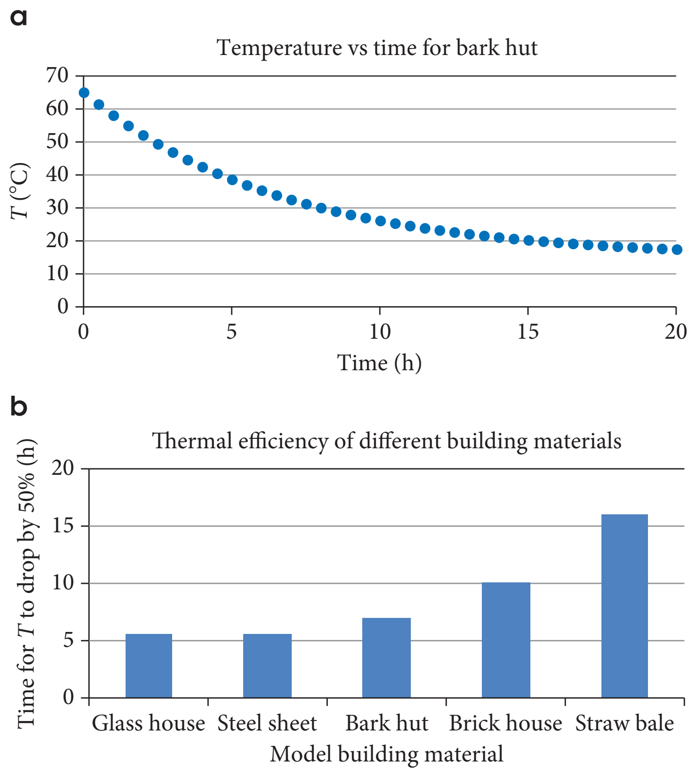


Figure 12.18 ▶ a) A scatter plot demonstrating a mathematical relationship; b) A column graph comparing results from different experiments

Any data and derived results should be given in correct SI units with their uncertainties. If you performed calculations then show the equations you used. You might want to show one example calculation, but do not show more than one if the procedure used is repeated.

Discussion

The discussion should summarise *what your results mean*. If you began with a research question, give the answer to the question here. If you began with a hypothesis, state whether your results support your hypothesis or not. If not, explain why. You may not be able to say why, other than that the model was not suitable for the situation being investigated.

If there are any implications of your work, such as how to build or do something in a better way, put them here.

The discussion is also the place to briefly describe any difficulties that you had and make suggestions for improving the process. Remember that you should never say 'the experiment didn't work' if you didn't get the results you expected. You might choose to make some comments on possible further work that could be done.

Conclusion

The conclusion is a *very* brief summary of the results and their implications. Say what you found out and what it means. A conclusion should only be a few sentences long.

Acknowledgements

You should thank anyone who helped you in your investigation. This includes people who supplied equipment or funding, as well as people who gave you good ideas or helped with the analysis. In science, as in other aspects of your life, it is polite to say thank you; however, this is not a necessary section of a report.

References

A reference list details the source for each piece of information used, and is linked to that information in the report.

A reference list details the sources of all information that were actually used to write the report. Wherever a piece of information or quotation is used in your report it must be referenced *at that point*. This is typically done either by placing a number in brackets at the point [2], or the author and year of publication (Smith, 2014). The reference list is then either provided in a footnote at the end of the page, or a single complete list at the end of the report. Referencing must be done in a consistent style. Check with your teacher what style they prefer. There are several good online guides to referencing.

A reference list is *not* the same as a bibliography. A bibliography is a list of sources that are useful to understanding the research. They may or may not have actually been used by the report authors. You should have a bibliography in your logbook from the planning stage of your investigation. The references will be a subset of these sources.

Appendices

Appendices may be used to provide additional information such as raw data that is not necessary to understanding the report but which might be of interest to some readers. Your teacher might require you to provide raw data in an appendix. Reports do not always have appendices.

Design reports

If you have done a design-build-test project, then the report will be of a slightly different form. It will follow the form of an engineering report rather than a scientific report.

Design reports usually contain the following sections.
- Design brief
- Introduction
- Design
- Testing
- Analysis
- Recommendations
- Conclusions
- Acknowledgements
- References
- Appendices

Many of the sections have the same function as in a scientific report. However, the analysis will deal mainly with the performance of the product. Design reports sometimes also include a recommendations section. The conclusion of the report states whether the final product meets the requirements of the design brief. There are some guidelines for and examples of these reports in the weblinks.

Other ways of communicating your results

You may want to present the results of your investigation in some other way. Scientists communicate their work in many ways. Sometimes a poster is presented or a seminar is given. An article may be written or a website produced. Scientists usually use more than one means, and sometimes several of them, to communicate about a really interesting investigation.

REFERENCING GUIDE

This guide is designed to help you with referencing your sources for assignments.

REFERENCING I-TUTORIAL

This tutorial will help you understand referencing and show you how to avoid plagiarism.

REPORT WRITING EXAMPLE 1

The brochure outlines the key features of a design report.

REPORT WRITING EXAMPLE 2

This online resource guides you through the sections of a typical report.

REPORT WRITING EXAMPLE 3

This online resource will help you write a case study.

▲ **Figure 12.19** A post-doctorate student presenting her poster at a poster session at a conference

Look at examples of articles in the scientific and the popular media, on websites, posters and so on. This will give you an idea of the different styles used in the different modes. Think about the purpose. Is it to inform, to persuade or both? What sort of language is used?

Think about your audience and use appropriate language and style. A poster is not usually as formal as a report. A website may be more or less formal, depending on your audience.

Posters and websites use a lot of images. Images are usually more appealing than words and numbers, but they need to be relevant. Make sure they communicate the information you want them to.

Make sure you keep readability and accessibility in mind if you are creating a poster or website. Posters should use large clear fonts and not have too much text. They should be readable from a few metres away. Fonts also need to be large enough and clear on websites and digital images should have tags. You can follow the weblink for more information on accessibility and web-page design.

However you communicate your work, make sure you know what the message is and who the audience is. Once you have established that, you will be able to let other people know about the interesting things you have discovered in your investigation.

WEBSITE ACCESSIBILITY

The Royal Society for the Blind has information on making websites accessible.

CHAPTER GLOSSARY

analogue a device or scale that gives a continuous measurement; the scale is continuous and may show any value in a range

continuous variable a variable that is able to take any value, sometimes within a fixed range

controlled variable the variable that is controlled by the experimenter, so that its values are chosen; also called the independent variable

dependent variable the variable that changes as a result of changes to the independent or controlled variable

design brief the document that specifies the requirements for a design, including performance of the final product

digital able to measure only a limited number of possible values, usually within a fixed range

discrete variable a variable that is able to take only specific values, not continuous; for example, a line spectrum is a discrete spectrum

extrapolation extension beyond the measured range of data to read or construct new data that has not been measured

falsifiable able to be disproved

hypothesis a tentative prediction, usually based on an existing model or theory; also a tentative explanation of an observation based on an existing model or theory

independent variable a variable upon which another variable depends; the controlled variable

interpolation to read or construct a new data point that has not been measured but is within the range of measured data

limit of reading the minimum uncertainty in a measurement due to the precision with which the scale can be read

line of best fit the line that most accurately fits the data, usually calculated using linear regression

linearise to make linear; to convert into a form that can be described by a straight line

logbook the record of an experiment or investigation kept by the scientist performing the experiments; it is a legal record of the experiments and their results

outlier a data point that does not fit the pattern shown by other measured data points

plagiarism presenting someone else's work, including their words or ideas, as your own

precision the variation in repeated measurements, or the uncertainty of a measuring device

primary data data that you have measured or collected yourself

qualitative measurements measurements with descriptive or non-numerical results

quantitative measurements measurements with numerical values

reference the source of a specific piece of information or quotation; to state the source of information

reliable highly likely to be true; a trustworthy source of information or reproducible data

reproducible giving the same result, within uncertainty, when repeated measurements are made

research question the specific question that a particular experiment or investigation is designed to answer

resolution the limit of reading of a measuring device

scatter graph a graph or plot showing data points, without a line joining the points, and used to demonstrate or determine a mathematical relationship between variables; the axes are defined by the variables

secondary data data or information that has been collected by someone else

theory a collection of models and concepts that explain specific systems or phenomena; scientific theories allow predictions to be made and hence are falsifiable

uncertainty bars bars drawn above and below and/or to left and right of a data point on a graph to indicate the size of the uncertainty in that point

valid results that are affected only by a single independent variable and hence are reproducible

variable something that can change or be changed, as distinct from a constant, which cannot

APPENDICES

Appendix 1: Advice for studying Physics and reading the textbook

Overview

Each chapter has a logical structure. This can be seen in the way the headings are written. The largest headings are used to start a complete section. For a logical segment of the text, such as a complete chapter, do the following, in order, before you do a close reading of the text:

- Read all headings and sub-headings.
- Look at all diagrams and read their captions.
- Read the chapter and section reviews.
- Ask: What is this segment or chapter about? What am I going to be learning about?

Close reading

There is no substitute for careful and specific close reading of all relevant parts of the textbook. This is a student's direct responsibility.

A close reading of the text, plus the notes taken in class, will form the basis for the notes that students generate in order to study, understand and remember.

[High-performing students frequently also read around the subject, for example, by using science-related feeds from the Internet and/or subscribing to a science-type magazine.]

Note taking

All notes should be:

- written in your own words.
- in point form and each point should contain one idea only. Points should not be longer than one line.

(Some people like to use a highlighter on their textbook to emphasise what is important – this makes them too dependent on the author; but it can be a worthwhile, intermediate step towards full understanding.)

Questions

- Keep a list of questions about things you need to check up on.
- BRING THE LIST TO CLASS and use it.
- Seek assistance by using the questions as a basis for your learning.

Seeking help

- Always try to specify what you need help about.
- Ask a precise question.

Reading ahead – always!

- The best learning takes place when you are prepared in advance and have some idea about what is coming up.
- Seek advice from your teacher about the next lesson or series of lessons.

Appendix 2: SI and non-SI units

International System of Units (SI)

The international body that decides the appropriate units to be used for the various physical quantities is the *Conférence Générale des Poids et Mesures* (CGPM). The system of units approved by the CGPM and now widely used by the scientific community throughout the world is known as Système International d'Unites (abbreviated SI).

In your experimental work you should use SI units (or their multiples or submultiples).

The SI consists of seven base units and two supplementary units. All other derived units are based on these nine fundamental units.

The base and supplementary units, together with the derived units with special names that might be relevant to your experimental work are listed in Tables A2.1 and A2.2.

Table A2.1 SI base units

Physical quantity	Name of unit	Abbreviation
Length	metre	m
Mass	kilogram	kg
Time	second	s
Electric current	ampere	A
Thermodynamic temperature	kelvin	K
Luminous intensity	candela	cd
Amount of substance	mole	mol

As you become familiar with each new unit you should make a practice of correctly using its abbreviated form.

The internationally recognised prefixes for the SI units together with their abbreviations are given in Table A2.3.

Table A2.2 SI supplementary units derived from SI base units

Name	Symbol	Quantity	Equivalents	SI base unit equivalents
hertz	Hz	frequency	1/s	s^{-1}
radian	rad	angle	m/m	dimensionless
steradian	sr	solid angle	m^2/m^2	dimensionless
newton	N	force, weight	$kg\ m/s^2$	$kg\ m\ s^{-2}$
pascal	Pa	pressure, stress	N/m^2	$kg\ m^{-1}\ s^{-2}$
joule	J	energy, work, heat	N m	$kg\ m^2\ s^{-2}$
watt	W	power, radiant flux	J/s	$kg\ m^2\ s^{-3}$
coulomb	C	quantity of electric charge	A s	A s
volt	V	electromotive force, electrical potential difference, electric potential voltage	J/C	$kg\ m^2\ s^{-3}\ A^{-1}$
farad	F	electrical capacitance	C/V s/Ω	$kg^{-1}\ m^{-2}\ s^4\ A^2$
ohm	Ω	electrical resistance, impedance, reactance	V/A	$kg\ m^2\ s^{-3}\ A^{-2}$
siemens	S	electrical conductance	1/Ω A/V	$kg^{-1}\ m^{-2}\ s^3\ A^2$
weber	Wb	magnetic flux	J/A	$kg\ m^2\ s^{-2}\ A^{-1}$
tesla	T	magnetic field strength, magnetic flux density	$V\ s/m^2$ Wb/m^2 N/(A m)	$kg\ s^{-2}\ A^{-1}$
degree Celsius	°C	temperature relative to 273.15 K	K – 273.15	K – 273.15
lumen	lm	luminous flux	cd sr	cd
lux	lx	illuminance	lm/m^2	m^{-2} cd
becquerel	Bq	radioactivity (decays per unit time)	1/s	s^{-1}
gray	Gy	absorbed dose (of ionising radiation)	J/kg	$m^2\ s^{-2}$
sievert	Sv	equivalent dose (of ionising radiation)	J/kg	$m^2\ s^{-2}$

Table A2.3 Prefixes for SI units

Prefix	Abbreviation	Value	Prefix	Abbreviation	Value
exa	E	10^{18}	deci	d	10^{-1}
peta	P	10^{15}	centi	c	10^{-2}
tera	T	10^{12}	milli	m	10^{-3}
giga	G	10^9	micro	μ	10^{-6}
mega	M	10^6	nano	n	10^{-9}
kilo	k	10^3	pico	p	10^{-12}
hecto	h	10^2	femto	f	10^{-15}
deka	da	10	atto	a	10^{-18}

Non-SI units

A number of non-SI units are still in use in scientific literature for a variety of reasons. Some of these are being phased out, but others are likely to remain in use. The more common non-SI units that you might come across are listed in Table A2.4.

Table A2.4 Non-SI units

Physical quantity	Unit	Abbreviation	Conversion to SI units
time	minute hour day year	min h d y	60 s 3.6×10^3 s 8.64×10^4 s 3.156×10^7 s
mass	unified mass unit tonne	u t	1.661×10^{-27} kg 1000 kg
angle	degree	°	dimensionless
energy	electron-volt kilowatt hour	eV kW h	1.602×10^{-19} J 3.60×10^3 J
pressure	millimetre of mercury	mmHg	133.3 Pa
charge	elementary or electronic charge	e	1.602×10^{-19} C
source activity	curie	Ci	3.7×10^{10} Bq
radiation absorbed dose equivalent dose	 rad rem	 rad rem	 0.01 Gy 0.01 Sv

Using SI units

There are certain conventions now adopted widely in scientific literature when SI units are being used. Some of the more important ones are given below.

1 When recording a measurement, write the unit in full or use the recommended abbreviation (e.g. 25 metre or 25 m). Using abbreviations save space and time. Notice the space between the numeral and the unit.

2 SI units named after scientists:

 a If the full word is used, it starts with a lower case letter (e.g. 10 newton, 7 joule, 10^5 pascal, 50 hertz)

 b If the abbreviation is used, it is (or at least commences with) a capital letter (e.g. 10 N, 7 J, 10^5 Pa, 50 Hz).

 c Measurements are written as products. '3 kg' means 'the product of 3 and the mass known as a kilogram', just as '3*x*' in maths means the product of 3 and *x*. Therefore 's' is not added to units (e.g. 5 kg or 5 kilogram, not 5 kgs or 5 kilograms).

 d A full-stop is not placed after the abbreviation of a unit, unless it is at the end of a sentence.

 e When units are combined as a quotient (e.g. metre per second), a solidus (/) or negative index may be used. So m/s or m s^{-1} are both acceptable, though the latter is used more widely. Never use more than one solidus in a unit as in m/s/s for acceleration, which should be m/s^2 or m s^{-2}. It is ambiguous, just as writing 36/6/3 in maths is ambiguous. (This could mean 2 or 18.)

Converting between units

Treat the unit as a multiplier. Use the prefixes in Table A2.3 also as multipliers. For example, 4 kg is the same as:

$$4 \times k \times g = 4 \times 1000 \times g = 4000 \text{ g}$$

There are 4000 gram in 4 kg.

1500 cm is the same as 1500 × (10^{-2} m) = 15 m.

CONVERTING BETWEEN UNITS

Learn more about converting between units at this weblink.

Appendix 3: Some important physical quantities

From time to time you will need to find a value of a physical property from a reputable source. These might include finding:

- the value of a physical constant, such as Newton's universal gravitational constant or the electric constant.
- a physical property, such as boiling point or refractive index, which is characteristic of a particular material.
- a conversion factor such as micrometres to metres, electron-volt to joule, unified mass unit to kilogram.

NIST PHYSICAL REFERENCE DATA

The National Institute of Standards and Technology (NIST) provides a wide range of data, including Standard reference data (SRF). For example, click on 'Other Data' to enter the NIST Gateway.

All physical quantities, including physical constants, are measured to very precise levels of accuracy.

NIST PHYSICAL ELEMENT LABORATORY

This website provides atomic and nuclear data for every element.

Some important physical quantities, including some physical constants, are listed alphabetically in Table A3.1. They are given to four significant figures. The uncertainty in most of these figures is better than six-figure accuracy. They are taken from sources such as the National Institute of Science and Technology (NIST). NIST is a specialist organisation dedicated to metrology (study of measurement).

Table A3.1 Physical constants, physical measures and conversion factors

Physical constants	
Avogadro constant, N_A	6.022×10^{23} mol^{-1}
Coulomb law constant, $\frac{1}{4\pi\varepsilon_0}$	8.988×10^{9} N m^{2} C^{-2}
Universal gravitation constant, G	6.674×10^{-11} m^{3} kg^{-1} s^{-2}
Permittivity of free space electric constant, ε_0	8.854×10^{-12} F m^{-1}
Permeability of free space magnetic constant, μ_0	$4\pi \times 10^{-7}$ H m^{-1} = 12.57×10^{-7} H m^{-1}
Planck constant, h	6.626×10^{-34} J s = 4.136×10^{-15} eV s
Speed of electromagnetic radiation in free space, c	2.998×10^{8} m s^{-1}
Physical measures	
Mass of electron	9.109×10^{-31} kg = 5.486×10^{4} u
Mass of proton	1.6726×10^{-27} kg
Mass of neutron	1.6749×10^{-27} kg
Rydberg constant (for hydrogen), R_H	1.097×10^{7} m^{-1}
Gravitational field strength at Earth's surface), g	(9.80 ± 0.3) N kg^{-1}
Acceleration due to gravity at Earth's surface, g	(9.80 ± 0.3) m s^{-2}
Mass of Earth	5.976×10^{24} kg
Mass of Moon	7.348×10^{22} kg
Mass of Sun	1.989×10^{30} kg
Period of rotation of Earth	8.616×10^{4} s
Radius of Earth (equatorial)	6.378×10^{6} m
Radius of Earth (mean)	6.371×10^{6} m
Radius of Earth's orbit about Sun (mean)	1.496×10^{11} m
Radius of Moon's orbit around Earth (mean)	3.844×10^{8} m

Radius of Sun	6.960×10^{8} m
Solar constant (mean)	1.370×10^{3} W m^{-2}
Density of water (pressure and temperature dependent)	9.982×10^{3} kg m^{-3}
Air density (pressure and temperature dependent)	1.292×10^{3} kg m^{-3}
Air pressure (temperature dependent)	1.013×10^{5} Pa
Speed of sound in air at 0°C	331.4 m s^{-1}
Conversion factors	
Absolute zero, 0 K	−273.15°C
Unified mass unit, u	1.661×10^{-27} kg = 931.5 MeV
Electron-volt, eV	1.602×10^{-19} J
Elementary electron charge, *e*	1.602×10^{-19} C
Coulomb	6.242×10^{18} elementary charges

Appendix 4: Analysis of data

Graphical analysis

The purpose of experiments in physics is to find regularities in the physical world. If one thing varies regularly when another is changed, then it may be possible to find the reason behind the relationship. Therefore, physicists try to exclude all variables that might interfere with their investigations. They look for one variable that causes another one variable to change. Sufficient data is collected to draw graphs in order to visualise the relationship. They compare the shape of the graph against known graph shapes in order to determine the form of the relationship.

Look at Figure A4.1. What information can you obtain from this graph of daily solar intensity in Alice Springs? What was the annual average? When could you expect the greatest benefit from solar panels? Could the annual variation be turned into an algebraic equation?

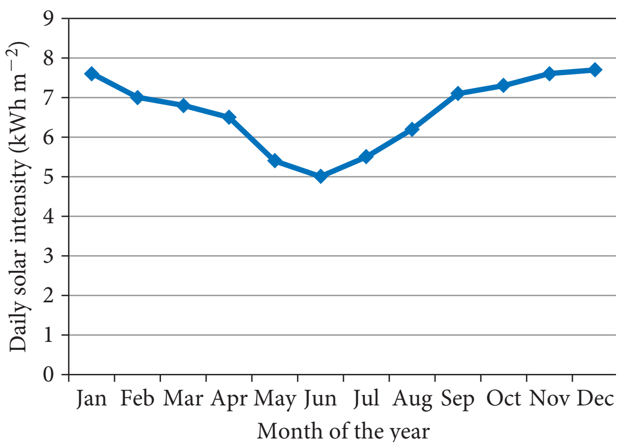


Figure A4.1 Daily solar radiation per area in Alice Springs for one year

Variables

A variable is some quantity that changes or varies. Some changes cause other changes to occur, For example a force acting on an object causes the object to accelerate. The acceleration of the object does not cause the force to be applied to it. The variable that causes some other to change is called the independent variable. The variable that changes is the dependent variable.

Linear graphs

The most important graph in physics is the straight line or linear graph. In its general form the equation is:

$$y = mx + c$$

The shape of this graph is shown in Figure A4.2.

y is the dependent variable.

x is the independent variable.

m is the gradient of the line.

c is the intercept on the y axis.

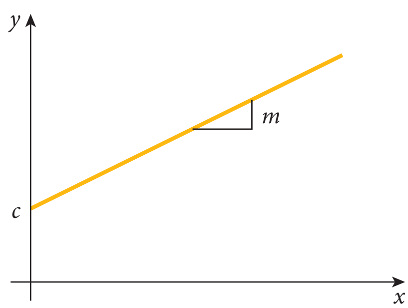


Figure A4.2 Graph of $y = mx + c$

It is usual to plot the independent variable (the one that is controlled) on the horizontal axis (x axis) and the dependent variable (the one that is not controlled) on the vertical axis (y axis). This will become clearer with practice. It is also possible, and indeed common, to plot the dependent variable against the square or inverse values of the independent variable, for example, y vs x^2, y vs $\frac{1}{x}$, etc.

A special case is the straight-line graph that passes through the origin. The variables are then described as being directly proportional to each other. You need to be careful *not* to assume that (0, 0) at the origin is necessarily a data point.

For direct proportionality or direct variation, $y = mx$ or $y = kx$, the value of the independent variable, y, doubles, triples etc. whenever the value of the independent variable, x, doubles, triples etc.

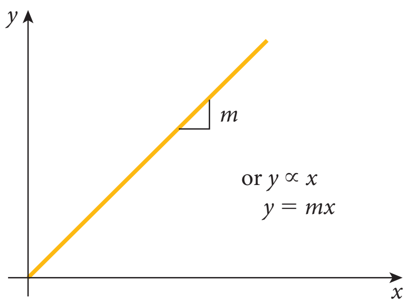


Figure A4.3 Graph of $y = mx$

For example, say that you had three cubes of aluminium. Cube A has sides 1.0 cm in length. Cube B has sides that are twice as long (2.0 cm) and cube C has sides three times as long (3.0 cm). You find the mass of each cube and put the data in a data table. What is the relationship between mass and volume?

Measurements made on the cubes gave the results given in Table A4.1.

Table A4.1

Cube	Length of side (L) (cm)	Mass (m) (g)	L^3 (calculated) (cm^3)
A	1.0	2.7	1.0
B	2.0	21.6	8.0
C	3.0	72.9	27.0

Changes in the length cause the mass to change. If we plot m against L, the shape of the graph is as shown in Figure A4.4.

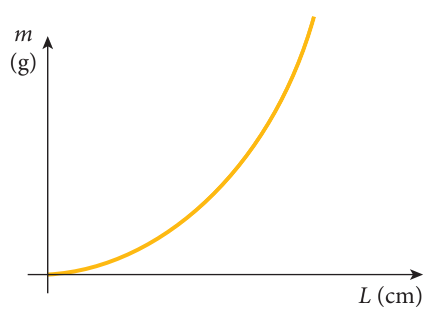


Figure A4.4 Graph of m against L

A little experience with graph shapes and variations in the numbers in data tables suggests that the variation is a cubic. We might expect, therefore, that $m = kL^3$. If we plot m against L^3, we see that the shape of the graph is a straight line. We find the gradient, which enables the equation to be written:

$$\text{gradient, } k = \frac{(72.9 - 2.7)\,\text{g}}{(27.0 - 1.0)\,\text{cm}^3}$$

$$k = 2.7\frac{\text{g}}{\text{cm}^3}$$

$$k = 2.7\ \text{g cm}^{-3}$$

Thus, $m = 2.7\,L^3$ (see Figure A4.5).

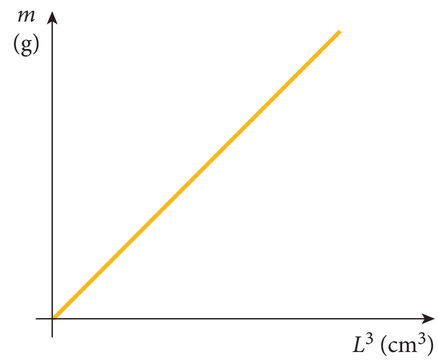


Figure A4.5 Graph of m against L^3

In most cases, the gradient has a physical meaning. In this case, the gradient is the density, in g cm^{-3}, of aluminium.

A similar relationship holds for other graphs. That is, if a straight line is obtained, then the quantity plotted on the vertical axis (y axis) is proportional to the quantity plotted on the horizontal axis (x axis).

If we plot the graph of y against x^2, and find the graph is a straight line, the equation becomes:

$$y = kx^2$$

where k is the gradient of the graph.

The graphs are shown in Figures A4.6 and A4.7.

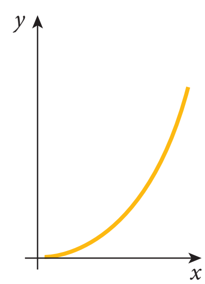


Figure A4.6

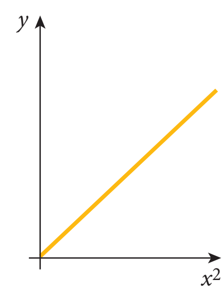


Figure A4.7

If we plot the graph of y against the inverse of x, that is $\frac{1}{x}$, and the graph is a straight line, the equation becomes:

$$y = k\left(\frac{1}{x}\right) \text{ or } y = \frac{k}{x}.$$

The graphs are as shown in Figures A4.8 and A4.9.

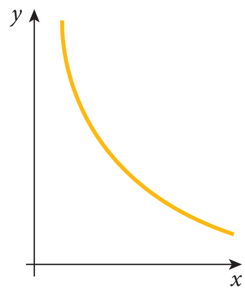


Figure A4.8

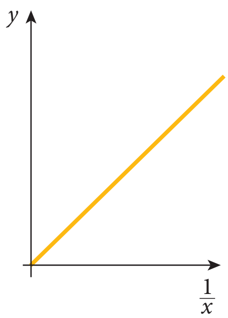


Figure A4.9

In this case, y and x are said to be inversely proportional.

This graphical method is used to analyse experimental data. Measured values are plotted on a graph. The shape of the graph suggests a possible relationship between the quantities. A further graph may need to be plotted to confirm the relationship and obtain the equation.

Table A4.2 shows experimental results, from the measurement of pressure and volume of a fixed mass of gas, and it can be analysed to find a relationship between P and V.

Table A4.2 Experimental results from the measurement of pressure and volume of a fixed mass of gas

V (m^3)	P (Pa)	$1/V$ (m^{-3})
10.00	81.0	0.100
5.00	159.0	0.200
2.00	397	0.500
1.00	792	1.00

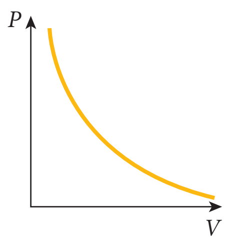


Figure A4.10 Graph of P against V

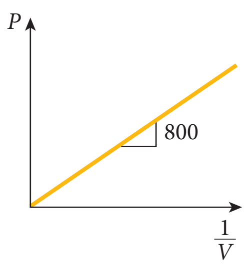


Figure A4.11 Graph of P against $\frac{1}{V}$

Plot P against V. The graph is of the form shown in Figure A4.10.

This suggests that the relation is either inverse $\left(\frac{1}{V}\right)$ or inverse square $\left(\frac{1}{V^2}\right)$.

It could even be to a higher order, but this will not be dealt with here. The next step is to plot the graph of P against $1/V$, as shown in Figure A4.11.

Because this graph is a straight line, the equation can now be given exactly, and is in this case:

$$P = \text{gradient} \times \frac{1}{V} \quad \text{or} \quad P = \frac{800}{V}$$

Recognition of graph shapes is extremely useful. When plotting data, you can quickly see whether you have the expected relationship or can plan further analysis (see Figure A4.12).

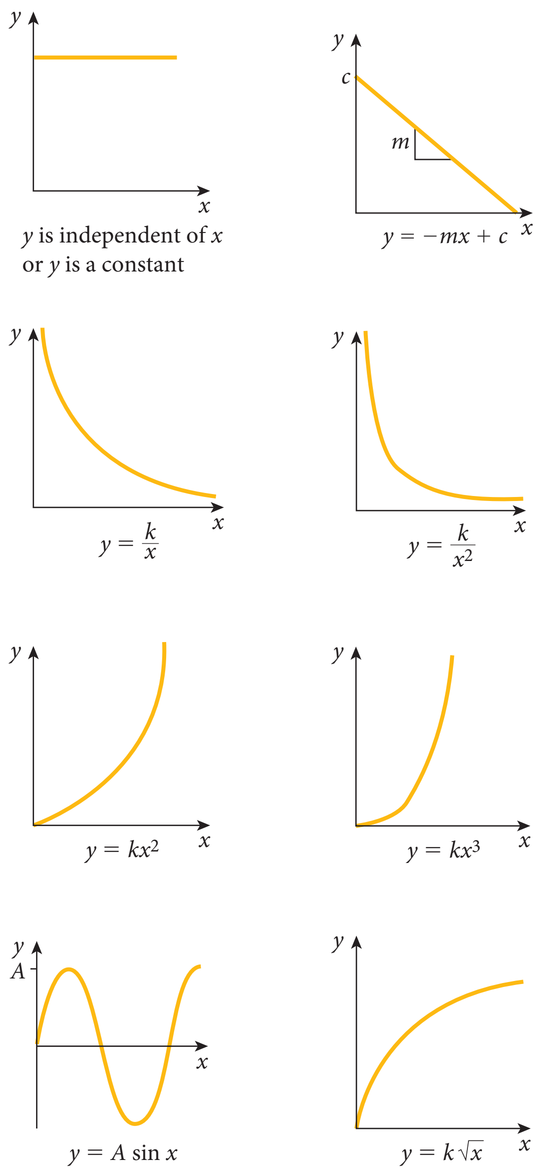


Figure A4.12 Graphs of important relationships

The gradient and the area under a graph may have physical meaning. The units of the gradient are the units of the quotient of the units of the vertical axis and units of the horizontal axis. The units of the area are the product of the units on the axes.

Interpolation and extrapolation

Interpolation is the estimation of a quantity between data points. Interpolation is only valid when there is good reason to believe that the graph is a smooth curve between the data points.

Extrapolation is the estimation of quantities beyond the domain and range of the data. This is less reliable than interpolation because the exact shape of the graph is unknown beyond the measured values.

In general, it is wise to be sceptical of the first and last data points on the graph. This means that, for simple relationships, you will need a minimum of three data points between the first and last points. The data points should also be spread out relatively evenly to ensure the data field is adequately sampled. There is little point in collecting three points near the first and a last point a long way away.

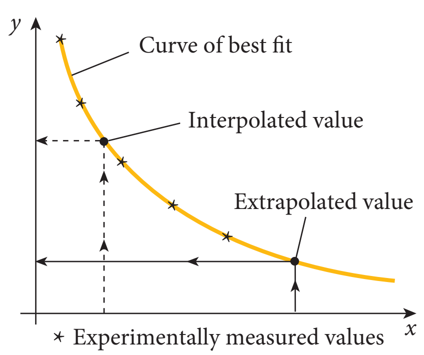


Figure A4.13 Interpolated values are between measured values. Extrapolated values are outside the range of measured values.

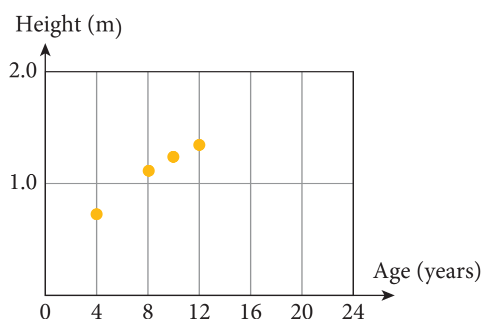


Figure A4.14 From this graph of the height of a student against age, predict height at birth, at 6 years of age and at 24 years of age. Comment on the accuracy of your predictions.

Numerical analysis

Another way to test for proportionality is to find the constant directly from the data coordinates. That is, you use the numbers directly to find whether the suspected relationship applies. If you refer back to Table A4.2 of P against V, you can see that, as P gets bigger and V gets smaller. This suggests an inverse relation. To check this numerically, we can use the following:

If: $P = \frac{c}{V}$, where c is a constant

then: $PV = c$

From Table A4.2 on page 358, the product PV is equal to 810, 795, 794 and 792. This is close enough to be considered a constant within the limits of experimental uncertainty. The equation becomes $PV = 798$ (the mean of the above values). This is close to 800, which is the value found graphically.

Graphical and numerical methods must both be evaluated against the data and the experimental uncertainty. Neither is superior.

If we had found that PV was not a constant, then we might have tried:

$$P = \frac{c}{V^2}$$

That is, $PV^2 = c$.

For linear relationships, $A = cB$, where c is a constant. Parabolic, square relationships, have $A = cB^2$, so: $A/B^2 = c$ etc.

Numerically measured values from the table can be substituted to check that the value of c is constant.

Power laws

Consider a series of cubes of sides 1.0, 2.0, 3.0 and 4.0 cm. If we calculate the perimeter of one side, the area of one side, the total surface area, the volume and the mass in terms of the
1.0 cm cube, we can find a relationship between the values.

The values are given in tabular form in Table A4.3.

1 Perimeter (P): There is a direct relationship between the perimeter and the length. If one is doubled, then the other is doubled, and so on.

$$P \propto l$$

2 Area (A): There is a square relationship between the area and the length of one side. If the length is doubled, then the area is increased fourfold (2^2), tripling the length increases the area ninefold (3^2), etc.

This can be seen to be true for both side and surface area, even though the values are different.

$$A \propto l^2$$

Table A4.3 Relationship between values and length of side

Length of side (cm)	Perimeter (cm)	Area of side (cm^2)	Total surface area (cm^2)	Volume (cm^3)	Mass units
1.0	4.0	1.0	6.0	1.0	1.0
2.0	8.0	4.0	24.0	8.0	8.0
3.0	12.0	9.0	54.0	27.0	27.0
4.0	16.0	16.0	96.0	64.0	64.0

3 Volume (*V*): There is a cubic relationship between the volume and the length of one side. If length is doubled, then the volume is increased eightfold (2^3); if it is increased by three, the volume is increased 27-fold (3^3) etc.

$$V \propto l^3$$

4 Mass (*m*): If we assume that the density remains constant, then the mass changes as the volume changes. That is, a cubic relationship exists between side length and mass. If the density changes, the value cannot be found unless other facts are known.

CURVE FITTING

For further information on curve fitting and data analysis, go to this website.

Appendix 5: Periodic table of elements

Key

atomic number → 26 Fe (symbol of element) name of element → iron standard atomic weight → 55.85

Symbol of element:
- gas at room temperature (red)
- liquid at room temperature (green)
- solid at room temperature (black)
- synthetic (does not occur naturally) (blue)

- s block (orange)
- p block (blue)
- d block transition metals (yellow)
- d block lanthanoids and actinoids (pink)

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 H hydrogen [1.007, 1.009]																	2 He helium 4.003
3 Li lithium [6.938, 6.997]	4 Be beryllium 9.012											5 B boron [10.80, 10.83]	6 C carbon [12.00, 12.02]	7 N nitrogen [14.00, 14.01]	8 O oxygen [15.99, 16.00]	9 F fluorine 19.00	10 Ne neon 20.18
11 Na sodium 22.99	12 Mg magnesium [24.30, 24.31]											13 Al aluminium 26.98	14 Si silicon [28.08, 28.09]	15 P phosphorus 30.97	16 S sulfur [32.05, 32.08]	17 Cl chlorine [35.44, 35.46]	18 Ar argon 39.95
19 K potassium 39.10	20 Ca calcium 40.08	21 Sc scandium 44.96	22 Ti titanium 47.87	23 V vanadium 50.94	24 Cr chromium 52.00	25 Mn manganese 54.94	26 Fe iron 55.85	27 Co cobalt 58.93	28 Ni nickel 58.69	29 Cu copper 63.55	30 Zn zinc 65.38(2)	31 Ga gallium 69.72	32 Ge germanium 72.63	33 As arsenic 74.92	34 Se selenium 78.96(3)	35 Br bromine [79.90, 79.91]	36 Kr krypton 83.80
37 Rb rubidium 85.47	38 Sr strontium 87.62	39 Y yttrium 88.91	40 Zr zirconium 91.22	41 Nb niobium 92.91	42 Mo molybdenum 95.96(2)	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 lanthanoids	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 [204.3, 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		

57 La lanthanum 138.9	58 Ce cerium 140.1	59 Pr praseodymium 140.9	60 Nd neodymium 144.2	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
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

9780170242110

NUMERICAL ANSWERS

Chapter 1

Question set 1.1

1 g has units of $m\,s^{-2}$

6 $t = 6.1\,s$

7 $y = 16\,m$

8 $u = 10.4\,m\,s^{-1}$

Question set 1.2

4 a $\sin 0 = 0$, so $F_{net} = 0$

b $\sin 90° = 1$, so $F_{net} = w$

5 No

7 $F_{north} = 650\text{ N}\cos 38° = 512\text{ N}$

and

$F_{east} = 650\text{ N}\sin 38° = 400\text{ N}$

8 $F = 1600\,N$

Question set 1.3

5 $t = 2.6\,s$

6 $t = 2.34\,s$

7 $t = 35.1\,s$

$x = u_x t = (246\,m\,s^{-1})(35.1\,s) = 8635\,m$ or $8.6\,km$

8 $u_y = u\sin\theta$ so $u = \dfrac{u_y}{\sin\theta} = \dfrac{22.1\text{ m s}^{-1}}{\sin 60°} = 26\,m\,s^{-1}$

Question set 1.4

5 b Force increases by a factor of 4.

7 $F = 140\,kN$

Question set 1.5

2 $\Sigma F = 0, \Sigma\tau = 0$

6 $F = 714\,N$

7 $F = 50\,kN$

8 a $N = 14\,kN$

b The friction force is up the slope.

Chapter review questions

1 Horizontal acceleration of a projectile is zero

2 $\Sigma Fx = 0, \Sigma Fy = 0, \Sigma\tau = 0$

4 $u_x = u\cos\theta$

12 a $y = 63\,m$

b $t = 3.6\,s$

13 a $u_y = u\sin\theta = 45\,m\,s^{-1}\sin 15° = 12\,m\,s^{-1}$

14 $N = 747\,N$

(Hence the person feels as if they weigh $\dfrac{747\text{ N}}{9.8\text{ m s}^{-2}} = 76\,kg$.)

15 $mg = 18\,m\,s^{-1}$ or $66\,km\,h^{-1}$

16 a $F_{min} = 1.1\,kN$

b $F = 2.6\,kN$

17 a $u_y = u\sin\theta = 38\,m\,s^{-1}\sin 45° = 27\,m\,s^{-1}$

$u_x = u\cos\theta = 38\,m\,s^{-1}\cos 45° = 27\,m\,s^{-1}$

b Time taken $= 2t = 5.5\,s$

c $x = u_x t = (27\,m\,s^{-1})(5.5\,s) = 148.5\,m$

21 Twice the force is required.

Chapter 2

Question set 2.1

2 a i $g = \dfrac{F_g}{m} = \dfrac{GM}{r^2}$

ii $F_g = \dfrac{GM}{r^2}$

b The mass and the radius of the planet.

7 $F = 4000\,N$ or $4\,kN$

8 $a = 1.7\,m\,s^{-2}$

Question set 2.2

6 $g = 13\,m\,s^{-2}$

7 $N = 320\,N$

8 $M = 7.36 \times 10^{22}\,kg$

Question set 2.3

1 a $\Sigma\Delta E_p + \Sigma\Delta E_K = 0$ or $\Sigma\Delta E_p = -\Sigma\Delta E_K$

5 $W = 59\,kJ$

7 The gravitational field of an asteroid is small compared to Earth because of its much smaller mass.

8 a $\Delta s = 8.2\text{ m}$

b $v = 13\text{ m s}^{-1}$

c $v = 46\text{ m s}^{-1}$

Question set 2.4

5 $M = 5.64 \times 10^{26}\,kg$

6 $v = 9.4 \times 10^4\,m\,s^{-1}$

7 $T_{asteroid} = 2.8$ years

Chapter review questions

11 The force drops to 1% of its original value

13 $v = 889\ \text{m s}^{-1}$

14 a $g = 1.5\ \text{m s}^{-2}$

b $g = 1.5\ \text{m s}^{-2}$

c $\Delta PE = W = mg\Delta h = 5.0\ \text{kg} \times 1.5\ \text{m s}^{-2} \times 10\ \text{m} = 75\ \text{J}$

d Maximum KE gained is equal to the change in PE, 75 J.

15 $F = 356 \times 10^6\ \text{N}$

19 $v = 6.3 \times 10^3\ \text{m s}^{-1}$ or $6.3\ \text{km s}^{-1}$

Chapter 3

Question set 3.1

1 $E = \frac{F}{q}$, so $F = Eq$

5 $E = 4.5 \times 10^9\ \text{N C}^{-1}$

6 $a = 1.7 \times 10^{13}\ \text{m s}^{-2}$; downwards

7 $x = 14\ \text{m}$

Question set 3.2

2 $\frac{F_{q_1 \text{ on } q_2}}{F_{q_2 \text{ on } q_1}} = -1$

4 $F = 9.0 \times 10^6\ \text{N}$

5 a $r = 3.2 \times 10^{-5}\ \text{m}$

b $r \rightarrow \sqrt{2}r = 4.5 \times 10^{-5}\ \text{m}$

c $F \rightarrow \frac{F}{2^2} = \frac{F}{4} = 5.8 \times 10^{-20}\ \text{N}$

6 $Q = 1 \times 10^{-7}\ \text{C}$

Question set 3.3

3 $1\ \text{V} = 1\frac{\text{J}}{\text{C}} = 1\frac{\text{N m}}{\text{C}} = 1\frac{\text{kg m s}^{-2}\ \text{m}}{\text{A s}} = 1\ \text{kg m}^2\text{s}^{-3}\text{A}^{-1}$

5 $\Delta U = q\Delta V = (2.0\ \text{C})(1.5\ \text{V}) = 3.0\ \text{J}$

6 $\Delta U = \Delta q \Delta V = I\,\Delta t\,\Delta V = (1\ \text{A})(1\ \text{s})(12\ \text{V}) = 12\ \text{J}$

7 a $v = 1.9 \times 10^7\ \text{m s}^{-1}$

b $v = 4 \times 10^5\ \text{m s}^{-1}$

c $v = 3 \times 10^5\ \text{m s}^{-1}$

Chapter review questions

7 $|F_{\text{a on b}}| = |F_{\text{b on a}}| = |F_{\text{b on c}}| = |F_{\text{c on b}}| > |F_{\text{a on c}}| = |F_{\text{c on a}}|$

11 $F = Eq = (50\ \text{V m}^{-1})(1 \times 10^{-9}\ \text{C}) = 5.0 \times 10^{-8}\ \text{N}$

$a = \frac{F}{m} = \frac{5 \times 10^{-8}\ \text{N}}{1 \times 10^{-3}\ \text{kg}} = 5.0 \times 10^{-5}\ \text{m s}^{-2}$

12 a $F = Eq = (550\ \text{V m}^{-1})(1.6 \times 10^{-19}\ \text{C}) = 8.8 \times 10^{-17}\ \text{N}$

b $x = 3.2 \times 10^{-6}\ \text{m}$ or $3.2\ \mu\text{m}$

13 a $\left(\frac{1}{4\pi\varepsilon_0}\right)\frac{Q}{r^2} = 9.0 \times 10^9\ \text{N m}^2\ \text{C}^{-1} \times \frac{1.6 \times 10^{-19}\ \text{C}}{(3.8 \times 10^{-10}\ \text{m})^2} = 1.0 \times 10^{10}\ \text{N C}^{-1}$

b $F = Eq = (1.0 \times 10^{10}\ \text{N C}^{-1})(1.6 \times 10^{-19}\ \text{C}) = 1.6 \times 10^{-9}\ \text{N}$

15 a $F = \left(\frac{1}{4\pi\varepsilon_0}\right)\frac{Q^2_{\text{proton}}}{r^2} = 58\ \text{N}$

b $F = \frac{Gm_p^2}{r^2} = 4.7 \times 10^{-35}\ \text{N}$

17 a $KE_{\text{initial}} = \frac{1}{2}mv_i^2 = \frac{1}{2}(9.1 \times 10^{-31}\ \text{kg})(25 \times 10^3\ \text{m s}^{-1})^2 = 2.8 \times 10^{-22}\ \text{J}$

c $\Delta V = \frac{\Delta U}{q} = \frac{\Delta KE}{q} = \frac{2.8 \times 10^{-22}\ \text{J}}{1.6 \times 10^{-19}\ \text{C}} = 1.8 \times 10^{-3}\ \text{V}$ or $1.8\ \text{mV}$

Chapter 4

Question set 4.1

4 $F = ilB = (9\ \text{A})(1\ \text{m})(0.01\ \text{T}) = 0.09\ \text{N}$

5 a $F = ilB\sin\theta = (5.0\ \text{A})(2.0\ \text{m})(50 \times 10^{-6}\ \text{T})(\sin 30°) = 2.5 \times 10^{-4}\ \text{N}$

b $B = 3.5 \times 10^{-5}\ \text{T}$ or $35\ \mu\text{T}$

7 a $T = 5.2 \times 10^{-6}\ \text{s}$

b $f = 1.9 \times 10^5\ \text{Hz}$ or $190\ \text{kHz}$

c $B = 6.8 \times 10^{-6}\ \text{T}$ or $6.8\ \mu\text{T}$

Question set 4.2

7 $B = 2.0 \times 10^{-5}\ \text{T}$

8 $B = \frac{\mu_o I}{2\pi r}$

$r = 3.0\ \text{m}$

9 $B = \frac{\mu_o I}{2\pi r}$

$I = 2.5\ \text{A}$

Chapter review questions

9 $B = 12.5\,T$

10 $B = 2.7 \times 10^{-4}\,T$ or $0.3\,mT$

11 a $I = 7500\,A$

b $\frac{1}{2} \times 0.15\,T = 0.075\,T$

12 $B = 0.028\,T$ or $2.8 \times 10^{-2}\,T$ or $28\,mT$

Chapter 5

Question set 5.1

1 $T\,m^2$; also called Weber, Wb

6 $r = \sqrt{\frac{\phi}{B\pi}} = \sqrt{\frac{5.0 \times 10^{-3}\ T\ m^2}{0.25\ T \times \pi}} = 0.080\,m$ or $8.0\,cm$

7 $\frac{\Delta B}{\Delta t} = \frac{iR}{A} = \frac{0.050\ A \times 1.0\ W}{0.050\ m^2} = 1.0\,T\,s^{-1}$

8 a $4.5 \times 10^{-4}\,Wb$

Question set 5.2

4 a $I_{max} = 30\,A$, $I_{ave} = 0\,A$, $I_{rms} = \frac{30\ A}{\sqrt{2}} = 21\,A$

b $f = \frac{6\pi}{2\pi} = 3\,s^{-1}$

c $T = \frac{1}{f} = \frac{1}{3\ s^{-1}} = \frac{1}{3}$ s or $0.33\,s$

5 a $\varepsilon_{max} = 2\pi fnBdl = 2\pi(30\,Hz)(30)(1.0\,T)(0.24\,m)(0.4\,m) = 540\,V$

b $emf_{rms} = 380\,V$

6 $B = 0.1\,T$

7 a $f = 1200\,rpm \times \frac{1\ min}{60\ s} = 20\,Hz$

b $\varepsilon_{max} = 19\,V$ (rounded from 18.8 V)

c $\varepsilon_{rms} = \frac{emf_{max}}{\sqrt{2}} = \frac{18.8\ V}{\sqrt{2}} = 13\,V$

d $\varepsilon(t) = \varepsilon_{max} \cos(2\pi ft) = 19\,V \cos(40\pi t)$

9 $n = 127$

Question set 5.3

5 $V_S = V_P \times \frac{N_S}{N_P} = 240\ V \times \frac{200}{1000} = 48\,V$

6 a Step up transformer

b $\frac{N_P}{N_S} = \frac{I_S}{I_P} = \frac{10\ A}{30\ A} = \frac{1}{3}$

7 a $V_s = 60\,kV$

b $P = VI = (120\,V)(8.0\,A) = 960\,W$

c Same as the input, 960 W

d $I_S = \frac{P_S}{V_S} = \frac{960\ W}{60000\ V} = 0.016\,A$

9 a $V = IR = (0.6\,A)(10\,\Omega) = 6\,V$

b $V = IR = (1.0\,A)(10\,\Omega) = 10\,V$

c i $N_S = V_S \times \frac{N_P}{V_P} = 6\ V \times \frac{1200}{240\ V} = 30$ turns

ii $N_S = V_S \times \frac{N_P}{V_P} = 10\ V \times \frac{1200}{240\ V} = 50$ turns

d $VPR = 6\,V + 10\,V = 16\,V$

e $I = \frac{V}{R} = \frac{16\ V}{10\ \Omega} = 1.6\,A$

Question set 5.4

6 a $F_{AB} = 0.18\,N$; downwards

c 0.18 N upwards

d $\Sigma\tau = 7.2 \times 10^{-3}\,Nm$, out of the page

e Anticlockwise

7 $I_2 = I_1\left(\frac{\tau_2}{\tau_1}\right) = 1.5\ A \times \frac{0.015\ N\ m}{7.2 \times 10^{-3}\ N\ m} = 3.1\,A$

Question set 5.5

1 James Clerk Maxwell

6 a $f = 4.7 \times 10^{14}\,Hz$

b $t = 2.6\,s$

8 a $\lambda = 550\,nm$

Chapter review questions

6 $Wb = 1\,T\,m^2 = 1\,kg\,s^{-2}A^{-1}m^2 = 1\,kg\,m^2 s^2 A^1$

13 a $I_p = 3.0\,A \sin(6.2\,s^{-1}\,t)$

b $I_{p-p} = 0.64\,A$

c $I_{ave} = 0$

d $I_{rms} = 0.22\,A$

14 b $F = nilB = 7.7 \times 10^{-2}\,N$ or $77\,mN$

c $\tau_{net} = 6.2\,m\,N\,m$

16 $B = 0.24\,T$

18 a $V_S = V_P \times \frac{N_S}{N_P} = 240\ V \times \frac{250}{500} = 120\,V$

c $I_S = I_P \times \frac{V_P}{V_S} = 2.4\ A \times \frac{240\ V}{120\ V} = 4.8\,A$

19 b $n = 108$ turns

20 a $\lambda = 0.17\,m$ or $17\,cm$

b $t = 0.0011\,s$ or $1.1\,ms$

Chapter 6

Question set 6.2

2 $x = x' + v\Delta t$

$y = y' + v\Delta t$

$\vec{v}_{AB} = \vec{v}_{AC} + \vec{v}_{CB}$

5 a i $v_{\text{person to bus}} = 2.0\,\text{m s}^{-1}$

ii $v_{\text{person to outside}} = v_{\text{person to bus}} + v_{\text{bus to outside}} = 14\,\text{m s}^{-1}$

b i $x = v\Delta t = (2.0\,\text{m s}^{-1})(5.0\,\text{s}) = 10\,\text{m}$

ii $x = v\Delta t = (14.0\,\text{m s}^{-1})(5.0\,\text{s}) = 70\,\text{m}$

Question set 6.4

6 $L' = 40\,\text{m}$

8 $t_{\text{pion}} = 17\,\text{ns}$

Question set 6.5

1 a $m = \gamma m_o$

b $p = \gamma p_o$

c $E = m_o c^2$

d $\Delta E = \Delta m_o c^2$

4 $E = (\gamma - 1)m_p c^2 = 1.0 \times 10^{-10}\,\text{J}$

5 $v = c\sqrt{1 - \frac{1}{\gamma^2}} = 0.89c$

6 $E = 7.46\,\text{MeV}$

Chapter review questions

5 a $v_p = v_{\text{p to train}} + v_{\text{train}} = 13\,\text{m s}^{-1}$

b $s = vt = 13\,\text{m s}^{-1} \times 5\,\text{s} = 65\,\text{m}$

8 a i $c = 3.00 \times 10^8\,\text{m s}^{-1}$

ii $c = 3.00 \times 10^8\,\text{m s}^{-1}$

b i 7.5 s

ii 2.5 s

9 $v = 0.10c$

10 35 years

11 23 ns

12 447 m

14 16 years

15 a Yes

b 5 years

17 a $0.75c$

b 56 m

18 $1.637 \times 10^{-13}\,\text{J}$

19 a i 660 m

b 500 m

c Yes

21 a $9.65 \times 10^{-21}\,\text{kg m s}^{-1}$

b 0.74 m

Chapter 7

Question set 7.1

6 a 560 nm

b 9.0 cm

Question set 7.2

5 $\lambda_{max} = \frac{b}{T} = \frac{2.898 \times 10^{-3}\ \text{m K}}{450\ \text{K}}$

$= 6.4 \times 10^{-6}$ m or 6.4 μm

6 a $4.04 \times 10^{-21}\,\text{J}$

b $2.0 \times 10^{12}\,\text{Hz}$

7 a 350 nm (from graph)

b 8000 K

8 a 1000 nm

Question set 7.3

5 a Copper

b Lithium

6 $E \leq hf - W$

7 a $6.7 \times 10^{-34}\,\text{J s}$

b 3.8 eV

Chapter review questions

7 Platinum, Pt

10 a 0

b 0.0393 m or 3.93 cm

c 0.0785 m or 7.85 cm

11 3000 nm; red

12 1.1 mm

13 a $8.8 \times 10^{-19}\,\text{J}$

b $2.3 \times 10^{-7}\,\text{m}$

c $1.3 \times 10^{15}\,\text{Hz}$

14 a 5.05×10^{-7} m or 505 nm; green

b $2.69 \times 10^{-19}\,\text{J}$

15 Aluminium or possibly lead

16 1.54×10^{-3} m or 1.54 mm

17 a $1.63 \times 10^{-24}\,\text{J}$

$1.02 \times 10^{-5}\,\text{eV}$

Chapter 8

Question set 8.1

5 3

6 4.6×10^{14} Hz

6.5×10^{-7} m = 650 nm

7 9.1×10^{-8} m or 91 nm

9 2.2×10^{6} m s^{-1}

Question set 8.2

5 a 1.1×10^{-35} m

6 a 0

b 3.3×10^{-3} m = 3.3 mm

c 6.7×10^{-3} m = 6.7 mm

7 a 1.2×10^{-23} kg m s^{-1}

b 3.1×10^{-11} m

c Less than

d Before: $\lambda = \frac{h}{p} = \frac{6.626 \times 10^{-34}\ \text{J s}}{1.2 \times 10^{-23}\ \text{kg m s}^{-1}}$

$= 5.5 \times 10^{-11}$ m

After: $\lambda = \frac{h}{p} = \frac{6.626 \times 10^{-34}\ \text{J s}}{0.9 \times 10^{-23}\ \text{kg m s}^{-1}}$

$= 7.4 \times 10^{-11}$ m

Chapter review questions

11 91×10^{-9} m or 91 nm, ultraviolet

12 $x = 10\ \text{eV} \pm 2.76\ \text{eV} = 12.6\ \text{eV}$ or 7.24 eV

13 a 3.45×10^{-11} m

b 1250 eV

14 a 6

b $n = 4$ to $n = 1$

c $n = 4$ to $n = 3$

15 $y_0 = 0$

$y_1 = 2.5 \times 10^{-5}$ m

$y_2 = 5.0 \times 10^{-5}$ m

19 a $F = \frac{k_e q_1 q_2}{r^2} = \frac{k_e e^2}{r^2}$

b $\frac{1}{2}mv^2 = \frac{k_e e^2}{2r} = \text{KE}$

c Total energy = KE + U = $\frac{k_e e^2}{2r} - \frac{k_e q_1 q_2}{r} = -\frac{k_e e^2}{2r}$

d $r = 5.2 \times 10^{-11}$ m

20 a $\lambda = 4.84 \times 10^{-6}$ m

b $y_0 = 0$

$y_1 = 0.65$ m or 65 cm

21 $V = 4.5$ V

Chapter 9

Question set 9.1

6 $V = 50\ \text{GV} = 50 \times 10^{9}$ V

7 a 6.4×10^{-7} J

b 0.7 m s^{-1}

8 $m = 0.512$ MeV/c^2

9 $B = 0.016$ T

Question set 9.4

5 $\Sigma^{+} \rightarrow \overline{n} + \pi^{+}$

6 $\tau^{+} \rightarrow e^{+} + \overline{\nu_{\tau}} + \nu_e$

A τ^{+} is likely to decay in this way.

Chapter review questions

3 a B = 0

b B = 0

c B = 0

d B = +1

4 a $L_e = 0$

b $L_e = +1$

c $L_e = 0$

d $L_e = +1$

11 b i τ: 1784 MeV c^{-2} $\times 1.78 \times 10^{-30}$ kg/(MeV c^{-2})

$= 3.17 \times 10^{-27}$ kg

ii Ω: 1672 MeV c^{-2} $\times 1.78 \times 10^{-30}$ kg/(MeV c^{-2})

$= 2.97 \times 10^{-27}$ kg

12 a i $V = 92.2 \times 10^{3}$ V or 922 kV

ii $B = 1.9 \times 10^{-6}$ T

13 $B = 5.6 \times 10^{-6}$ T

21 a $V = 1.03 \times 10^{7}$ V or 10 MV

b $B = 8.8 \times 10^{-6}$ T

22 $f = 1.1 \times 10^{4}$ Hz

Chapter 10

Question set 10.2

1 Strong, weak and electromagnetic forces

5 $F_G = 2 \times 10^{-34}$ N

$F_E = \frac{k_e q_1 q_2}{r^2} = 0$

6 $F_G = 2 \times 10^{-34}$ N

$F_E = 230$ N

Question set 10.3

6 $v = 0.097\ \text{m s}^{-1}$, about 10 cm/s

7 6.1×10^9 ly or 6 billion light years away

8 a $\lambda = 9.66 \times 10^{-7}$ m or approx. 1 μm (infrared)

b $\lambda = 9.66 \times 10^{-4}$ m or approx. 1 mm (microwave)

9 a $\lambda = 4.97 \times 10^{-7}$ m or 497 nm

b $v = 9.9 \times 10^7\ \text{m s}^{-1}$

c $R = 4.5 \times 10^9$ ly or 4.5 billion light years away

Chapter review questions

17 a $R = 2.0 \times 10^9$ ly or 2 billion light years

b $\lambda_{obs} = 763$ nm

21 $R = 7.9 \times 10^9$ ly or 8 billion light years away

GLOSSARY

absorption spectrum the wavelengths (or frequencies or energies) of radiation absorbed by a material

accurate the degree to which a measurement result approaches the 'true value'

action-at-a-distance non-contact force; one object experiences a force due to the presence of another object that is not touching it

aether a medium through which light waves were thought to propagate

albedo the ratio of light reflected by a surface to light incident on it; a surface with an albedo of 1 is perfectly reflective, and an albedo of 0 is perfectly absorbing

alternating current (AC) a current that varies with time between positive and negative values, usually sinusoidally

analogue a device or scale that gives a continuous measurement; the scale is continuous and may show any value in a range

angular momentum, *L* momentum associated with rotational or orbital motion, $L = mvr$

antimatter matter composed of antiparticles, such as positrons, antiprotons and antineutrons

antiparticle each particle has an antiparticle that has the same mass but opposite charge and magnetic moment to the particle

apparent weight subjective experience of weight, which is generally measured as the normal force exerted on an object that is not in equilibrium

apparent weightlessness the experience of having no normal force exerted on you; this occurs during free fall

armature the frame of the rotating part of a motor or generator, holding one or more coils

aurora australis the southern lights; light produced by the collision of high-energy charged particles with air molecules close to the South Pole

aurora borealis the northern lights; light produced by the collision of high-energy charged particles with air molecules close to the North Pole

back emf an induced emf that opposes the flow of current in a circuit, particularly in a coil of a motor

baryon a heavy particle, with baryon number $B = +1$

baryon number, *B* quantum number associated with baryons; baryons have $B = +1$, antibaryons have $B = -1$

best estimate value chosen to represent the indication value or measurement result of a measurand

Big Bang theory the theory that the universe began with a massive explosion of matter from a single point

black body an object with a perfectly absorbing surface, which emits radiation with a spectrum that is characteristic of the temperature of the object

black body radiation the electromagnetic radiation emitted by a black body, with a spectrum characteristic of the temperature of the body

boson particle with integer spin, $s = 0, 1$. These particles do not obey the exclusion principle. Examples include the exchange particles

centre of mass the average position of the mass in an object or group of objects. It is the point at which the gravitational force can be modelled as acting when the object is in a gravitational field

centripetal centre-seeking; applied to force and acceleration on and of objects moving with uniform circular motion

charge-reversal symmetry if all particles in an allowed reaction are replaced with their antiparticles (which have opposite charge), the new reaction is also allowed under known conservation laws

cloud chamber a chamber containing a supersaturated vapour through which high-energy particles pass, causing condensation of the vapour, and thus allowing their tracks to be visualised

collider a particle accelerator in which two particles are accelerated in opposite directions and collide

commutator a device for reversing the direction of current flow in a motor or generator

component the part of a vector pointing in a particular direction (usually horizontal or vertical)

concentric having the same centre

contact force a force applied to an object by another object having physical contact

continuous able to take any value, sometimes within a fixed range, as distinct from discrete or quantised

continuous variable a variable that is able to take any value, sometimes within a fixed range

continuous spectrum a spectrum containing radiation of all wavelengths, for example a rainbow

controlled variable the variable that is controlled by the experimenter, so that its values are chosen; also called the independent variable

cosmic microwave background radiation the observed radiation coming from all points in space corresponding to radiation from a black body at 3 K; it is believed to come from an earlier, much hotter stage of the universe's evolution

cosmological principle the assumption that matter is uniformly distributed throughout the universe on all scales; it is one of the assumptions underlying the Big Bang theory

coulomb the unit of charge, named after Charles Augustin de Coulomb

Coulomb constant the constant of proportionality for electric fields, $k_e = \frac{1}{4\pi\epsilon_0} = 9 \times 10^9\,N\,m^2\,C^{-2}$

Coulomb's law the relationship between force, charge and distance for two point charges

couple moment the torque or moment resulting from two equal but opposite forces acting on an object

critical test an experiment that, having been rigorously conducted, shows that one or other of competing theories is false

crossing symmetry if a particle in an allowed reaction is crossed to the other side of the reaction and replaced with its antiparticle, the new reaction is also allowed under known conservation principles as long as enough energy is available

cut-off frequency, f_0 the minimum frequency of light needed to eject an electron from a metal surface discrete able to take only specific values, not continuous. A line spectrum is a discrete spectrum

dark energy energy that is predicted from the increasing rate of expansion of the universe, but which is not identifiable as any currently known energy form

dark matter matter that is postulated to explain gravitational effects but which is not observable by the emission or reflection of light

dependent variable the variable that changes as a result of changes to the independent or controlled variable

derived data data that is deduced from raw data by mathematical manipulation, such as graphs, algebraic equations and geometric constructions

design brief the document that specifies the requirements for a design, including performance of the final product

deterministic predictable, able to be determined if enough information is available

digital able to measure only a limited number of possible values, usually within a fixed range

diode a circuit component that allows current to flow in only one direction

dipole a closely spaced pair of positive and negative charges (electric dipole) or north and south poles (magnetic dipole)

direct current (DC) a current that flows in a single direction

discrete variable a variable that is able to take only specific values, not continuous; for example, a line spectrum is a discrete spectrum

dispersion relation the relationship between the frequency of a wave and its wavelength

eddy current a circular current induced in a conductor due to a changing magnetic field

electric field the field due to electric charges, which applies a force to electric charges

electric potential potential energy per unit charge in an electric field

electric potential energy the potential energy stored in an electric field; the potential energy of a charge in an electric field

electrical permittivity a physical property of a medium associated with electricity

electromagnetic field model a combination of the electric field model and the magnetic field model, including the interaction between the two fields

electromagnetic force the combination of electric and magnetic forces acting on a charge, due to an electric and a magnetic field

electromagnetic induction the production of an emf due to a time-varying magnetic field

electromagnetic wave coupled oscillating electric and magnetic fields, which in vacuum form a transverse wave that propagates at speed *c*

electron charge charge of an electron, $-1.6 \times 10^{-19}\,C$

electron diffraction the interference of electrons due to the interaction of their probability waves

electron volt, eV a unit of energy equal to $1.6 \times 10^{-19}\,J$

electrostatic field model the model that assigns an electric field to stationary charges; it is this field that exerts forces on other charges

electrostatic force force due to and acting on stationary charges

electroweak theory theory combining electromagnetic and weak interactions

elementary particle a particle that does not have internal structure and cannot be broken into constituent particles. Leptons are considered to be elementary particles; mesons and baryons are not

emission spectrum the spectrum of radiation emitted by an object, for example, black body radiation or atomic spectra from a discharge tube

energy levels the allowed energies of a nucleus–electron system; often referred to as electron energy levels, even though they are characteristic of the atom, not of individual electrons

epicycles circles on circles, as used to describe the orbits of the planets

equilibrium having a constant state of motion; acted on by no unbalanced forces or torques

exchange particles or field particles particles that mediate interactions, such as photons, which are the field particle of the electromagnetic field; also called gauge bosons

exclusion principle the principle that no two electrons in an atom may have identical sets of quantum numbers; it was first stated by Wolfgang Pauli in 1925

explanation generalised account of why a body of data occurs

extrapolation extension beyond the measured range of data to read or construct new data that has not been measured

falsifiability principle used to determine the experimental data that would disprove a model, law or theory; data from a critical test

falsifiable able to be disproved

fermion particle with half-integer spin, $s = \frac{1}{2}, \frac{3}{2}, \ldots$, that obeys the exclusion principle; fermions include protons, neutrons and electrons

ferromagnetic having magnetic properties like iron; able to be magnetised and retain the magnetisation so that the material is magnetic

field the means by which action-at-a-distance forces are exerted

field particles *see* exchange particles

field theory the theory that describes forces as being mediated by fields and potential energy as being stored in fields

flavours (quarks) the six classifications of quark types: up, down, strange, charm, top and bottom

free fall falling with the acceleration *g*, the local gravitational field strength

friction the component of the contact force between surfaces which is parallel to the surfaces, and resists relative motion of the surfaces

Galilean transformation equations relating coordinates in one inertial frame to those in another inertial frame; classical relativistic transformations

gauge bosons the field or exchange particles predicted by gauge theory

generator a device that converts kinetic energy to electric potential energy and produces a current

geostationary orbit of period 24 h approx. directly above a single point on Earth's equator

geosynchronous orbit of period 24 h approx. above a great circle on Earth

grand unified theory (GUT) a theory that unites all four fundamental forces in a single model and explains the symmetry-breaking mechanism that caused them to separate into the four distinct forces we know of. There is as yet no widely accepted GUT

gravitas Aristotelian idea about the 'heaviness' of objects made of earth that allowed them to fall in straight lines towards Earth

gravitational field the field that mediates the gravitational force between all objects with mass; the field surrounding all objects with mass, $g = \frac{GM}{r^2}$

gravitational potential energy the potential energy associated with the interaction of objects via the gravitational force. The potential energy is stored in the gravitational field

ground level the lowest possible energy level of a nucleus–electron system

hadrons particles with large mass; the two types of hadrons are mesons and baryons

hypothesis a tentative prediction, usually based on an existing model or theory; also a tentative explanation of an observation based on an existing model or theory

independent variable a variable on which another variable depends; the controlled variable

indication value a single result of a measurement; the indication value gives a hint as to the 'true value'

induced current a current created by a changing magnetic field

induced emf an emf created by a changing magnetic field

induction motor an AC motor in which the torque is due to a current induced by the changing magnetic field created by an AC current in a coil

inertial frame of reference a non-accelerating frame of reference

interpolation to read or construct a new data point that has not been measured but is within the range of measured data

invariant same in all frames of reference

inverse-square law describes a relationship in which the dependent variable is proportional to the square of the inverse of the independent variable

length contraction length appears shorter in a reference frame that is moving relative to a stationary frame

lepton number, *L* the quantum number associated with leptons, there is one for each lepton type: $L_e = +1$ for electrons and electron neutrinos, $L_\mu = +1$ for muons and muon neutrinos and $L_\tau = +1$ for the tau and tau neutrino; the corresponding antiparticles have lepton number −1

leptons a family of elementary particles: electrons, taus and muons and their neutrinos and all their antiparticles

lever arm the distance between the point of application of a force and a pivot point

limit of reading the minimum uncertainty in a measurement due to the precision with which the scale can be read

line of best fit the line that most accurately fits the data, usually calculated using linear regression

line spectrum an emission or absorption spectrum consisting of discrete lines, characteristic of the energy levels of a particular atom or molecule

linear accelerator or linac a device in which electric and magnetic fields are used to accelerate charged particles to high speeds in a straight line

linearise to make linear; to convert into a form that can be described by a straight line

logbook the record of an experiment or investigation kept by the scientist performing the experiments; it is a legal record of the experiments and their results

Lorentz factor $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$

low Earth orbit an orbit between 250 km and 1000 km above Earth

magnetic braking braking due to the interaction of eddy currents and an external magnetic field

magnetic field the field created by moving charges, including charges in magnetic materials, which exerts a force on moving charges and magnetic materials

magnetic flux the magnetic field passing through a given area, $\Phi = BA\cos\theta$

magnetic flux density the magnitude of the magnetic field, measured in T

magnetic force the force that a magnetic field exerts on a moving charge or current

magnetic moment, μ also called the dipole moment, it is a vector of magnitude proportional to the magnetic field produced by a current loop, $\mu = IA$. The magnetic moment points in the direction of the field at the centre of the current loop

magnetic permeability physical property of a medium associated with magnetism

magnetic pole magnetic north or south pole, a point from which field lines come out or go in

magnetite an iron oxide, Fe_3O_4, that is a natural magnetic material

mass defect difference between energy before and after a nuclear decay or reaction

mass spectrometer a device that uses a magnetic field to characterise materials by the atoms and molecules they contain

mean lifetime statistical measure of the time after which a decaying sample is effectively extinguished; mean lifetime = 1.44 × half-life

measurand quantity being measured

measurement result best estimate of a 'true value'; numerical value based on judgements about one or more attempts to measure the 'true value'

meson a particle, generally (but not always) with mass between that of a lepton and a baryon

model a representation of a system or phenomenon that explains the system or phenomenon; a model may be mathematical equations, a computer simulation, a physical object, words or other form

modes of vibration characteristic patterns of oscillation, usually with a discrete set of allowed frequencies

moment *see* torque

motor a device that converts electric potential energy into kinetic energy, usually rotational kinetic energy

muon exotic particle formed by cosmic rays in the upper atmosphere

negligible any value or variation in a value that is too small to be taken into account

net force the resultant of all the forces acting on an object

normal force a reaction force perpendicular to the surface; a component of the contact force

outlier a data point that does not fit the pattern shown by other measured data points

particle accelerator a device in which electric and magnetic fields are used to accelerate beams of particles to high speeds

percentage error proportional error expressed as a percentage; not defined by BIPM

percentage uncertainty proportional uncertainty, expressed as a percentage

permeability of free space, μ0 the physical constant that determines the strength of the magnetic field produced by a current in vacuum. It has the value $4\pi \times 10^{-7}\,\mathrm{T\,m\,A^{-1}}$

permittivity of free space, ϵ_0 the physical constant that determines how large an electric field is produced by a charge in vacuum. It has the value $8.9 \times 10^{-12}\,C^2N^{-1}m^{-2}$

photocurrent the current formed by electrons ejected from a surface by incident photons

photoelectric effect the ejection of electrons from a surface by incident photons of sufficient energy

photoelectron an electron ejected from a metal surface following absorption of a photon of sufficient energy

photon a particle or quanta of light, having energy $E = hf$

plagiarism presenting someone else's work, including their words or ideas, as your own

Planck constant the constant of proportionality between energy and frequency for photons: $h = 6.626 \times 10^{-34}\,J\,s$

polar having regions of positive and negative charge

positron the antiparticle of the electron, with charge $+e$ and mass m_e

potential difference the difference in potential between two points in an electric field

potential energy energy stored in a system due to the interaction of components in the system via forces; energy stored in a field. Potential energy gives a system the ability to do work

precise the degree to which individual measurements cluster around the mean

precision the variation in repeated measurements, or the uncertainty of a measuring device

primary data data that you have measured or collected yourself

probabilistic not deterministic, unable to be predicted regardless of how much information is known

proper length length measured in an inertial frame of reference in which the object is stationary

proper time time measured between two events occurring at the same location in an inertial stationary frame of reference

proportional error difference between a measurement result and an accepted value, expressed as a fraction of the accepted value; not defined by BIPM

proportional uncertainty relative uncertainty

qualitative non-numerical data; descriptive information

qualitative measurements measurements with descriptive or non-numerical results

quantised existing in discrete amounts, not able to be divided into arbitrarily small amounts

quantitative numerical data; specific amount

quantitative measurements measurements with numerical values

quantum a discrete unit or amount of some physical property, such as energy, charge, mass or angular momentum

quark type of elementary particle (along with leptons and gauge bosons). Quarks come in six flavours and are fermions

radius of curvature the radius of a curve that is an arc (part of a circle in shape)

random error a variation that affects a measurement in a random way so that the measurement is as likely to change in any one direction as in any other

raw data original data taken directly from a measurement system

reaction diagram diagram showing the interaction of particles. Particles are represented as lines with arrows, and interactions are represented as a circle; time is shown on the horizontal axis

rectifier a circuit, typically consisting of diodes, that converts AC to DC

redshift the observed shift to longer wavelength of spectral lines in distant stars

reference the source of a specific piece of information or quotation; to state the source of information

relative uncertainty ratio of uncertainty to value

relativistic kinetic energy $\Delta E_k = (\Delta m)c^2 \Rightarrow \Delta E_k = (\gamma - 1)\,m_0c^2$

relativistic mass, relativistically corrected mass mass measured in a reference frame that is moving relative to the reference frame in which the rest mass was measured

relativity principle the laws of physics are the same in all inertial frames of reference

reliable highly likely to be true; a trustworthy source of information or reproducible data

representation model of reality

reproducible giving the same result, within uncertainty, when repeated measurements are made

research question the specific question that a particular experiment or investigation is designed to answer

resolution the limit of reading of a measuring device

rest energy $E = m_0c^2$

rest mass mass measured in an inertial reference frame in which the object is at rest

root mean square (rms) the average AC potential difference or current that produces the same power in a load as a DC potential difference or current of the same magnitude

rotor the rotating part, typically coils, of a motor or generator

scatter graph a graph or plot showing data points, without a line joining the points, and used to demonstrate or determine a mathematical relationship between variables; the axes are defined by the variables

secondary data data or information that has been collected by someone else

significant figure digit reported in a measurement result; the number of significant figures is the number of meaningful digits in a measurement result

simultaneity idea that the same event will be seen to occur at the same time in different reference frames

singularity a point in space and time at which the density of matter/energy is infinite and the volume is infinitesimally small

solenoid a coil of current-carrying wire that creates a large uniform field within the coil

spectroscope a device that disperses radiation by energy (or wavelength or frequency) so that a spectrum may be observed and measured

spectrum the distributed components of light or another wave arranged by frequency (or wavelength)

spin in quantum theory, a property of particles, including electrons, that results in them having their own magnetic moment and hence magnetic field

Standard Model the current most widely accepted model of particle physics that uses quarks and leptons to explain the nature of matter and exchange particles to explain the origins of three of the four fundamental forces: strong force, weak force and electromagnetic force; does not include a theory for gravity

stator the stationary or non-rotating coils of a motor or generator

stellar aberration position of star due to relative motion of Earth

step-down transformer a transformer with a lower output potential difference than the input potential difference

step-up transformer a transformer with a higher output potential difference than the input potential difference

stopping voltage the reverse bias voltage required to stop the flow of photoelectrons in a photoelectric effect experiment

strong nuclear force or nuclear force the force that acts between nucleons (protons and neutrons) to hold the nucleus together; it is mediated by pions

symmetry the invariance of physical laws under transformations such as translation, reflection and rotation in time and space

symmetry breaking a change in the behaviour of a physical system or the laws of physics that govern its behaviour when a symmetry operation such as a translation, reflection or rotation in time or space takes place

synchrotron a machine that uses electric and magnetic fields to accelerate charged particles to large velocities while containing them in rings, to produce high-energy light

systematic error an error that acts to give a consistent offset in data; for example, a zero error

tension the pulling force applied to an object by a string or cable

tesla the unit of magnetic field, $1\,\text{T} = 1\,\text{kg}\,\text{s}^{-1}\,\text{C}^{-1}$; it is named after Nikolai Tesla

theory a collection of models and concepts that explain specific systems or phenomena; scientific theories allow predictions to be made and hence are falsifiable

time dilation a longer time is measured

time-reversal symmetry when an allowed reaction is written such that it runs in the opposite direction in time; the new reaction is also allowed in that it does not break any of the known conservation laws

torque the turning effect of a force; the product of force and distance from the axis of rotation or pivot to the point of application of the force; also called moment

transformer a device for changing the magnitude of an AC potential difference

true value the exact value of a measurand; the 'true value' is an ideal that can never be known with certainty

unbalanced force when the resultant of two or more forces acting is not zero, i.e. $\Sigma F \neq 0$

uncertainty estimate of the range of values within which the 'true value' of a measurement or derived quantity lies; the extent to which the result of an experiment is unknown or unpredictable

uncertainty bars bars drawn above and below and/or to left and right of a data point on a graph to indicate the size of the uncertainty in that point

unified theory any theory that demonstrates how fundamental forces can be united, and explains the mechanism by which they become distinct, for example, the electroweak theory

uniform circular motion circular motion with constant speed

uniform field a field with a constant magnitude and direction over some region in space

valid results that are affected only by a single independent variable and hence are reproducible

van der Waals forces electrostatic forces due to charged regions of molecules, resulting in weak chemical bonds including hydrogen bonds

variable something that can change or be changed, as distinct from a constant, which cannot

vector a quantity that has magnitude and direction

vector cross product the vector cross product, $\vec{C} = \vec{A} \times \vec{B}$, gives a vector perpendicular to both $\vec{A}$ and $\vec{B}$ with magnitude $C = AB\sin\theta$, where θ is the angle between $\vec{A}$ and $\vec{B}$. The right-hand rule gives the direction of $\vec{C}$

vector dot product the scalar product of two vectors, $C = \vec{A} \cdot \vec{B} = AB\cos\theta$

volt, V unit of potential difference, $1\,\text{V} = 1\,\text{J}\,\text{C}^{-1}$

voltage more correctly called the potential difference, it is measured in volts

wave equation a differential equation that describes wave behaviour; its solutions are wave functions that are typically sinusoids

wave-particle duality the dual nature of matter and energy, requiring both wave and particle models to completely explain all observed behaviour of matter and energy

weak force the force necessary for β decay of nuclei. It is mediated by W and Z bosons. At very high energies it is unified with the electromagnetic force as part of the electroweak force

weight the gravitational force that acts on an object, $w = mg = \frac{GmM}{r^2}$

work energy transferred due to the action of a force, $W = Fs$

work function the energy required to eject an electron from a metal surface; effectively, it is the ionisation energy for the bulk material

zero error scale is not zero when measurements are taken; also called a calibration error

INDEX

nelsonnet

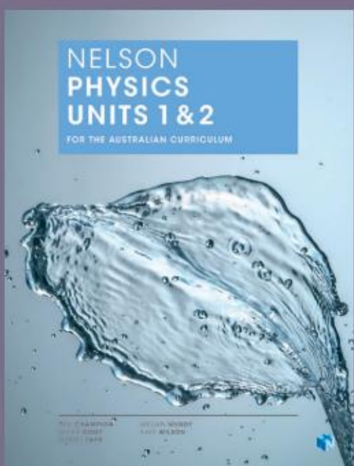
Register online at:
www.nelsonnet.com.au
to access digital resources to support improved learning & teaching*

NELSON
PHYSICS
UNITS 3 & 4

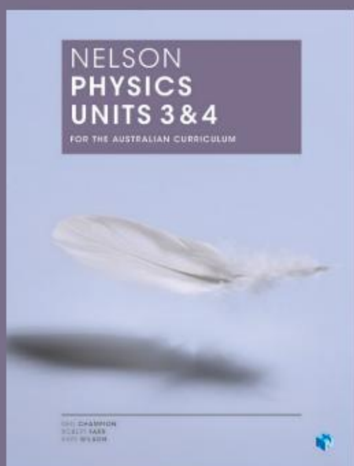
FOR THE AUSTRALIAN CURRICULUM

Nelson Physics Units 3 & 4 for the Australian Curriculum has been specifically designed to cover the three strands of the ACARA Australian Curriculum – Senior Physics.

Written to the highest achievement standard, this engaging series brings Physics to life with clear language and relevant examples. Case studies and Scientific literacy boxes help students connect the study of Physics to the real world.



NELSON PHYSICS
UNITS 1 & 2
ISBN 9780170242103



NELSON PHYSICS
UNITS 3 & 4
ISBN 9780170242110

NELSON
CENGAGE Learning
For learning solutions, visit **cengage.com.au**

*Please note that complimentary access to NelsonNet and the NelsonNetBook is only available to teachers who use the accompanying student textbook as a core educational resource in their classroom. Contact your sales representative for information about access codes and conditions.

ISBN: 978-0170242110
9 780170 242110