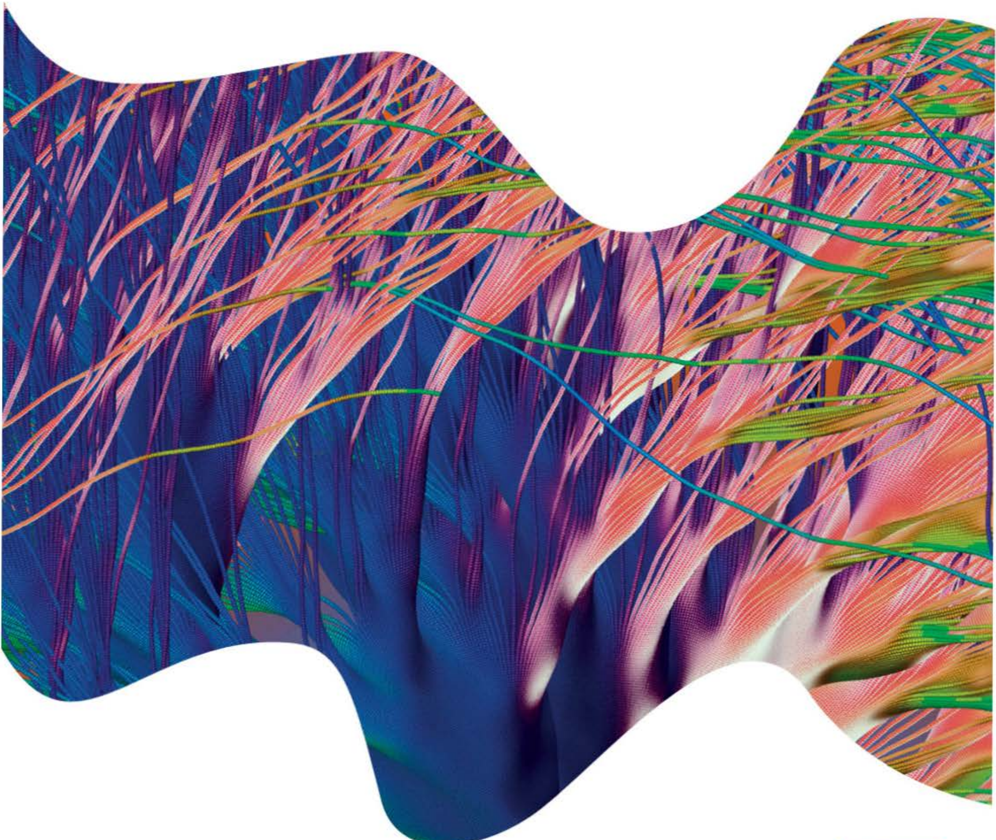
PEARSON

PHYSICS

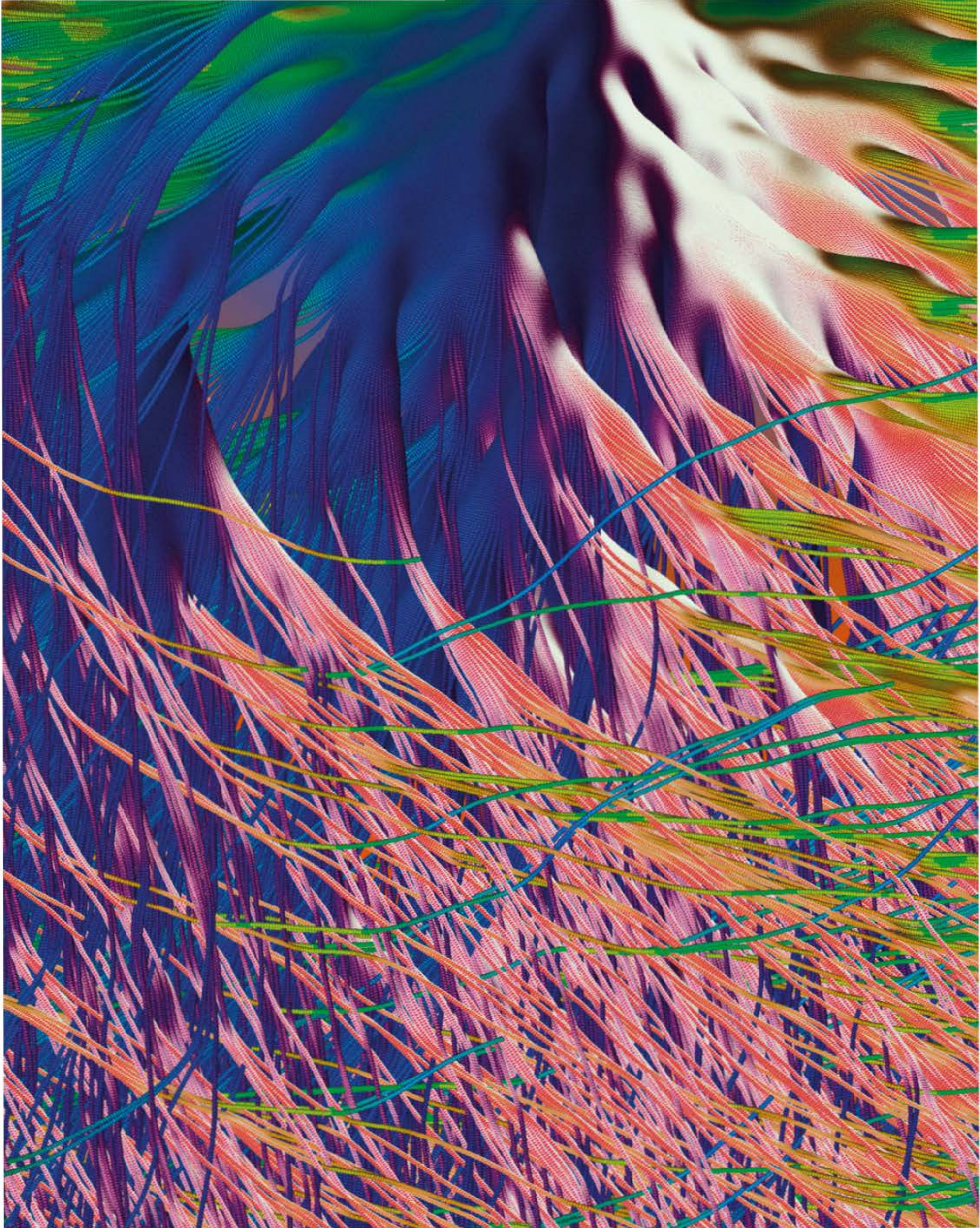
NEW SOUTH WALES

STUDENT BOOK

12



NSW
STAGE 6



Pearson Australia
(a division of Pearson Australia Group Pty Ltd)
707 Collins Street, Melbourne, Victoria 3008
PO Box 23360, Melbourne, Victoria 8012
www.pearson.com.au

First published 2019 by Pearson Australia
2022 2021 2020 2019
10 9 8 7 6 5 4 3 2 1

Lead Publishers: Misal Belvedere and Malcolm Parsons
Project Manager: Michelle Thomas
Production Editors: Laura Pietrobon, Virginia O'Brien
Lead Development Editor: Amy Sparkes
Content Developer: Bryonie Scott
Development Editors: Naomi Campanale, Haeyean Lee
Editor: Sam Trafford
Designer: Anne Donald
Rights & Permissions Editor: Samantha Russell-Tulip
Senior Publishing Services Analyst: Rob Curilli
Proofreader: Camha Pham
Indexer: Ann Philpott
Illustrator: DiacriTech
Printed in China by Golden Cup

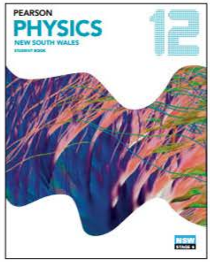
A catalogue record for this book is available from the National Library of Australia

ISBN 978 1 4886 1930 4
Pearson Australia Group Pty Ltd
ABN 40 004 245 943

Some of the images used in *Pearson Physics 12 New South Wales Student Book* might have associations with deceased Indigenous Australians. Please be aware that these images might cause sadness or distress in Aboriginal or Torres Strait Islander communities.

PEARSON
PHYSICS
NEW SOUTH WALES
12

Writing and development team



We are grateful to the following people for their time and expertise in contributing to the *Pearson Physics 12 New South Wales* project.

Bryonie Scott
Content Developer
Subject Lead

Doug Bail
Education Consultant
Contributing Author and Skills and Assessment Author

Amber Dommel
Teacher
Author

Norbert Dommel
Lecturer
Author

Tracey Fisher
Lecturer and Teacher
Skills and Assessment Author

Mark Hamilton
Teacher
Author

Kristen Hebden
Teacher
Author

Richard Hecker
Science Writer
Author

Brianna Hore
Teacher
Skills and Assessment Author

John Joosten
Educator
Skills and Assessment Author

David Madden
Teacher
Author

Svetlana Marchouba
Laboratory Technician
Safety Consultant

Jeff Stanger
Teacher
Author

Brett Stone
Principal Education Officer, NSW Department of Education
Author

Jim Sturgiss
Science Consultant
Author and Reviewer

Keith Burrows
Educator
Contributing Author

Rob Chapman
Educator
Contributing Author

Ann Conibear
Teacher
Contributing Author

Paul Cuthbert
Teacher
Contributing Author

Carmel Fry
Teacher
Contributing Author

Alistair Harkness
Teacher
Contributing Author

Jack Jurica
Teacher
Contributing Author

Greg Moran
Teacher
Contributing Author and Reviewer

Daniela Nardelli
Teacher
Contributing Author and Reviewer

John Nicholson
Teacher
Contributing Author

Craig Tilley
Science Writer
Contributing Author

Reuben Bolt
Director of the Nura Gili Indigenous Programs Unit, UNSW
Reviewer

Paul Looyen
Teacher
Reviewer

Michael O'Leary
Teacher
Reviewer

Trish Weekes
Science Literacy Consultant

Maria Woodbury
Teacher
Reviewer

George Howitt
Scientist
Answer Checker

Cameron Parsons
Scientist
Answer Checker

Gregory White
Scientist
Answer Checker

Adam Whittle
Scientist
Answer Checker

Contents

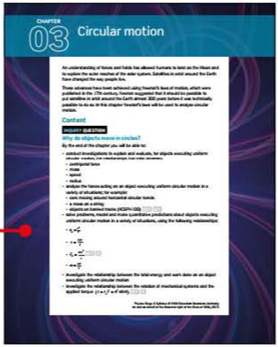
How to use this book

Pearson Physics 12 New South Wales

Pearson Physics 12 New South Wales has been written to be fully aligned with the new Stage 6 syllabus for New South Wales Physics. The book covers Modules 5 to 8 in an easy-to-use resource. Explore how to use this book below.

Chapter opener

The chapter opening page links the syllabus to the chapter content. Key content addressed in the chapter is clearly listed.



Section

Each chapter is clearly divided into manageable sections of work. Best-practice literacy and instructional design are combined with high-quality, relevant photos and illustrations to help students better understand the idea or concept being developed.

PhysicsFile

PhysicsFile boxes include a range of interesting and real-world examples to engage students.

Revision box

Revision boxes are used to remind students of vital concepts previously covered that are required for current learning.

Physics Inquiry

Physics Inquiry features are inquiry-based activities that assist students to discover concepts before learning about them. They encourage students to think about what happens in the world and how science can provide explanations.

Physics in Action

Physics in Action boxes place physics in an applied situation or a relevant context. They refer to the nature and practice of physics, its applications and associated issues, and the historical development of its concepts and ideas.

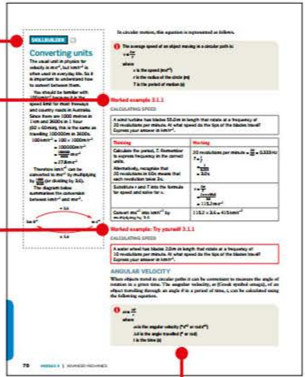
SkillBuilder

A SkillBuilder outlines a method or technique. They are instructive and self-contained. They step students through the skill to support science application.

Worked examples

Worked examples are set out in steps that show thinking and working. This format greatly enhances student understanding by clearly linking underlying logic to the relevant calculations.

Each Worked example is followed by a Try yourself activity. This mirror problem allows students to immediately test their understanding.



Converting units

Worked example 3.1.1

CALCULATING SPEED

Worked example: Try yourself 3.1.1

CALCULATING SPEED

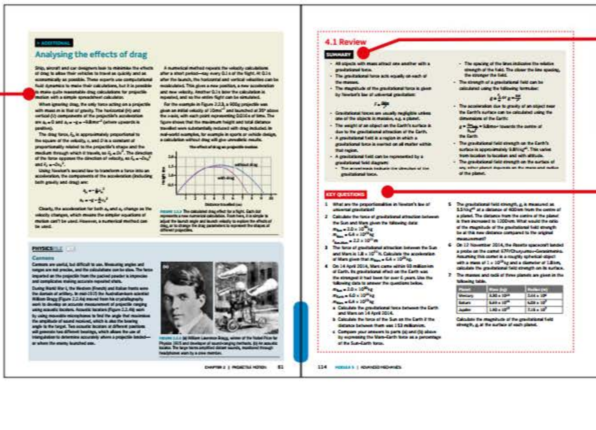
ANGULAR VELOCITY

Highlight box

Highlight boxes focus students' attention on important information such as key definitions, formulae and summary points.

Additional content

Additional content includes material that goes beyond the core content of the syllabus. They are intended for students who wish to expand their depth of understanding in a particular area.



Analysing the effects of drag

4.1 Review

SUMMARY

KEY QUESTIONS

Section summary

Each section has a summary to help students consolidate the key points and concepts.

Section review questions

Each section finishes with key questions to test students' understanding of and ability to recall the key concepts of the section.

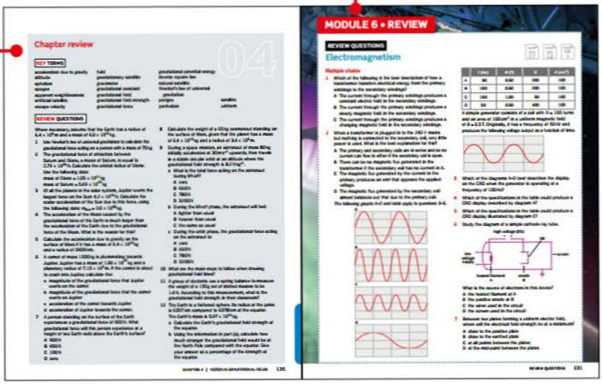
How to use this book

Chapter review

Each chapter finishes with a list of key terms covered in the chapter and a set of questions to test students' ability to apply the knowledge gained from the chapter.

Module review

Each module finishes with a set of questions, including multiple choice and short answer. These assist students in drawing together their knowledge and understanding, and applying it to these types of questions.



Icons

The NSW Stage 6 syllabus 'Learning across the curriculum' and 'General capabilities' content are addressed throughout the series and are identified using the following icons.

AHC A CC CCT DD EU ICT
IU L N PSC S WE

'Go to' icons are used to make important links to relevant content within the same Student Book.

GO TO ➤

This icon indicates when it is the best time to engage with a worksheet (WS), a practical activity (PA), a depth study (DS) or module review (MR) questions in *Pearson Physics 12 New South Wales Skills and Assessment* book.

WS 3.1

This icon indicates the best time to engage with a practical activity on *Pearson Physics 12 New South Wales* Reader+.

PA 3.1

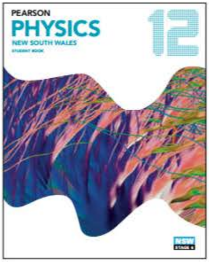
Glossary

Key terms are shown in **bold** in sections and listed at the end of each chapter. A comprehensive glossary at the end of the book includes and defines all the key terms.

Answers

Numerical answers and key short response answers are included at the back of the book. Comprehensive answers and fully worked solutions for all section review questions, Worked example: Try yourself features, chapter review questions and module review questions are provided on *Pearson Physics 12 New South Wales* Reader+.

Pearson Physics 12
New South Wales



Student Book

Pearson Physics 12 New South Wales has been written to fully align with the new Stage 6 syllabus for New South Wales. The Student Book includes the very latest developments in and applications of physics and incorporates best-practice literacy and instructional design to ensure the content and concepts are fully accessible to all students.

Skills and Assessment Book

Pearson Physics 12 New South Wales Skills and Assessment book gives students the edge in preparing for all forms of assessment. Key features include a toolkit, key knowledge summaries, worksheets, practical activities, suggested depth studies and module review questions. It provides guidance, assessment practice and opportunities for developing key skills.

Reader+ the next generation eBook

Pearson Reader+ lets you use your Student Book online or offline on any device. Pearson Reader+ retains the look and integrity of the printed book. Practical activities, interactives and videos are available on Pearson Reader+, along with fully worked solutions for the Student Book questions.

Teacher Support

Online teacher support for the series includes syllabus grids, a scope and sequence plan, and three practice exams per year level. Fully worked solutions to all Student Book questions are provided, as well as teacher notes for the chapter inquiry tasks. Skills and Assessment book resources include solutions to all worksheets, practical activities, depth studies and module review questions; teacher notes, safety notes, risk assessments and lab technician's checklists and recipes for all practical activities; and assessment rubrics and exemplar answers for the depth studies.

CHAPTER 01

Working scientifically

This chapter covers the skills needed to successfully plan and conduct primary and secondary-sourced investigations.

1.1 Questioning and predicting describes how to develop, propose and evaluate inquiry questions and hypotheses. When creating a hypothesis a consideration of the variables must be included.

1.2 Planning investigations will help you learn to identify risks in your investigation and to make sure all ethical concerns are considered. It is important to choose appropriate materials and technology to carry out your investigation. You will also need to confirm that your choice of variables allows for a reliable collection of data.

1.3 Conducting investigations is a guide to conducting scientific investigations. It describes methods for accurately collecting and recording data to reduce errors. Appropriate procedures need to be carried out when disposing of waste.

1.4 Processing data and information describes how to process your data appropriately. From an array of visual representations, you will learn how best to represent your information and how to identify trends and patterns in your data.

1.5 Analysing data and information explains how to analyse your results. It explains error and uncertainty and how to construct mathematical models to better understand the scientific principles of your research.

1.6 Problem solving will help you use critical thinking to demonstrate an understanding of the scientific principles underlying the solution to your inquiry question.

1.7 Communicating explains how to communicate an investigation clearly and accurately using appropriate scientific language, nomenclature and scientific notation.

Outcomes

By the end of this chapter you will be able to:

- develop and evaluate questions and hypotheses for scientific investigation PH12-1
- design and evaluate investigations in order to obtain primary and secondary data and information PH12-2
- conduct investigations to collect valid and reliable primary and secondary data and information PH12-3
- select and process appropriate qualitative and quantitative data and information using a range of appropriate media PH12-4
- analyse and evaluate primary and secondary data and information PH12-5
- solve scientific problems using primary and secondary data, critical thinking skills and scientific processes PH12-6
- communicate scientific understanding using suitable language and terminology for a specific audience or purpose PH12-7.

Content

By the end of this chapter you will be able to:

- develop and evaluate inquiry questions and hypotheses to identify a concept that can be investigated scientifically, involving primary and secondary data (ACSPH001, ACSPH061, ACSPH096) L
- modify questions and hypotheses to reflect new evidence CCT
- assess risks, consider ethical issues and select appropriate materials and technologies when designing and planning an investigation (ACSPH031, ACSPH097) EU PSC
- justify and evaluate the use of variables and experimental controls to ensure that a valid procedure is developed that allows for the reliable collection of data (ACSPH002)
- evaluate and modify an investigation in response to new evidence CCT
- employ and evaluate safe work practices and manage risks (ACSPH031) PSC WE
- use appropriate technologies to ensure and evaluate accuracy ICT N
- select and extract information from a wide range of reliable secondary sources and acknowledge them using an accepted referencing style L
- select qualitative and quantitative data and information and represent them using a range of formats, digital technologies and appropriate media (ACSPH004, ACSPH007, ACSPH064, ACSPH101) L N
- apply quantitative processes where appropriate N
- evaluate and improve the quality of data CCT N
- derive trends, patterns and relationships in data and information
- assess error, uncertainty and limitations in data (ACSPH004, ACSPH005, ACSPH033, ACSPH099) CCT
- assess the relevance, accuracy, validity and reliability of primary and secondary data and suggest improvements to investigations (ACSPH005) CCT N
- use modelling (including mathematical examples) to explain phenomena, make predictions and solve problems using evidence from primary and secondary sources (ACSPH006, ACSPH010) CCT
- use scientific evidence and critical thinking skills to solve problems CCT
- select and use suitable forms of digital, visual, written and/or oral communication L N
- select and apply appropriate scientific notations, nomenclature and scientific language to communicate in a variety of contexts (ACSPH008, ACSPH036, ACSPH067, ACSPH102) L N
- construct evidence-based arguments and engage in peer feedback to evaluate an argument or conclusion (ACSPH034, ACSPH036). CC DD

1.1 Questioning and predicting

Before starting your investigation, you need to understand the working scientifically skills essential to completing a meaningful scientific investigation. Working scientifically involves many dynamic and interrelated processes.

These are:

- Questioning and Predicting
- Planning Investigations
- Conducting Investigations
- Processing Data and Information
- Analysing Data and Information
- Problem Solving
- Communicating.

During this course you will choose and implement the processes appropriate to your investigation and use your knowledge and understanding of physics to draw and communicate conclusions and suggest areas for future research. This section is a guide to some of the key steps that should be taken when first developing your inquiry questions and hypotheses.

WHAT INITIATES AN INVESTIGATION?

There are many starting points for an investigation. Curiosity can be triggered through observation and advances in technology.

REVISION

In Year 11 you learnt that scientific investigations are broken down into primary investigations (such as experiments in a lab, field work or designing a **model**) and secondary-sourced investigations (such as a **literature review**).

Observation

Observation includes using all your senses and the instruments available to allow closer inspection of things that the human eye cannot see. Through careful observation, you can learn a lot about the forces acting on an object, including support forces and resistive forces such as friction and air resistance.

The idea for a primary investigation of a complex problem arises from prior learning and observations that raise further questions.

How observations are interpreted depends on past experiences and knowledge. But to enquiring minds, observations will usually provoke further questions, such as those given below.

- What velocity must a satellite be moving at to travel in a geosynchronous orbit?
- How does the motion of charged particles in an electric field compare to the movement of an object with mass in a gravitational field?
- What causes interference in signals travelling through a wire?
- What modifications could be made to improve the efficiency of a transformer?
- What applications are there for special relativity?
- What information does the spectrum from a star reveal?

Many of these questions cannot be answered by observation alone, but they can be answered through scientific investigations. Lots of great discoveries have been made when a scientist has been busy investigating another problem. Good scientists have acute powers of observation and enquiring minds, and they make the most of these chance opportunities.

ADVANCES IN TECHNOLOGY

Technology plays an important role in science. Atoms do not have sharply defined boundaries and so it is not possible to measure their radii directly. Developments in technology have allowed for the accumulation of evidence for scientific theories, laws and models, such as technology to measure the distance between nuclei of atoms in molecules. For example, in a hydrogen molecule (H_2) the two nuclei are 64 pm (picometres) apart. The radius of each hydrogen atom is assumed to be half of that distance, i.e. 32 pm.

The opposite is also true, with the development of new scientific theories, laws and models driving a need for new, improved technologies, such as the development of neutrino detectors, UV–visible spectroscopy, infrared spectroscopy and nuclear magnetic resonance spectroscopy.

REVISION

GO TO ➤ Year 11 Section 1.1

Inquiry question, hypothesis and purpose

The inquiry question, purpose (aim) and hypothesis are linked. Each of these can be refined during the planning of the investigation.

An **inquiry question** defines what is being investigated. It is important that you can interpret what an inquiry question is asking you to do. Compile a list of topic ideas and start a literature review to formulate your inquiry question. Evaluate and refine your inquiry question once you have decided on a topic. Remember to consider the resources available to you when deciding on your inquiry question.

A **hypothesis** is a prediction that is based on evidence and prior knowledge. A hypothesis often takes the form of a proposed relationship between two or more **variables** in a cause-and-effect relationship; in other words, 'If *x* is true and this is tested, then *y* will occur.'

A good hypothesis should be a statement that contains the independent and dependent variables, be measurable and be falsifiable. You may need to adjust your hypothesis as you conduct further research into your chosen topic. A **purpose** is a statement describing in detail what will be investigated. It is also known as the aim of your investigation. The purpose includes the key steps required to test the hypothesis. Each purpose should directly relate to the variables in the hypothesis, and describe how each will be measured. An experiment or investigation determines the relationship between variables and measures the results.

There are three categories of variables: **independent**, **dependent** and **controlled**. Variables are either **qualitative** (includes **nominal** and **ordinal** variables) or **quantitative** (includes **discrete** and **continuous** variables).

Reliable **primary** and **secondary sources** should be used when researching your topic and during your investigation.

PEER REVIEW

Scientists often publish their findings in peer-reviewed journals. Examples of peer-reviewed physics journals include:

- *New Journal of Physics*
- *Review of Modern Physics*
- *Journal of High Energy Physics*
- *Insights in Medical Physics*
- *Journal of Pure and Applied Physics*
- *Journal of Astrophysics and Aerospace technology.*

Peer-reviewed journals have an editorial board, comprising experts in a particular field, who read a draft article and ask questions of the author before agreeing to publish the article.

Your teacher may suggest that you partner with a student in your class, to provide each other with constructive feedback regarding your inquiry question and/or hypothesis. There are many benefits resulting from collaborating with others, including building on ideas and considering alternative perspectives. For example, the current understanding of light is the result of research conducted by several scientists in the 19th and 20th centuries.

PHYSICSFILE WE ICT

Detecting gravitational waves

In the early part of the twentieth century Einstein formulated the theory of general relativity. This theory predicted gravitational waves. Although hypothesised, scientists of the day were unsure how to detect them. Two American scientists, Kip Thorne from California Institute of Technology (Caltech) and Rainer Weiss from Massachusetts Institute of Technology (MIT), started to collaborate on detecting these waves.

Weiss was an experimental physicist, and had already started designs on a detector, and researched possible sources of disturbance that would interfere with the readings. Thorne was a theoretical physicist, and came to the collaboration with an understanding of the implications of gravitational waves, and the theoretical events that would cause them. In 1984 Caltech and MIT set up a joint project called LIGO (Laser Interferometer Gravitational-Wave Observatory).

LIGO has over 1000 scientists working in the team, led by Barry Barish, modifying and improving the detection equipment and the hypothesis, and interpreting the readings from the equipment. LIGO is so complicated and sensitive that the project was broken into different stages; it took three decades to set up the project and turn it on with the required sensitivity. The LIGO detector (Figure 1.1.1) uses long detector arms that are perpendicular, and a laser light that is sent down both arms, bouncing off mirrors at the end, and returning to the detector. Any change in length of the arms results in a changing interference pattern on the detector.

The consequences of detecting gravitational waves are wide reaching for science, and open up new areas of investigation as well as providing more evidence that supports the general theory of relativity. This was recognised when Weiss, Thorne and Barish were awarded the Nobel Prize in Physics in 2017.

FIGURE 1.1.1 LIGO is used in the search for gravitational waves.

When collaborating with others, consider the following:

- Is the inquiry question clear?
- Can the inquiry question be answered in the time available?
- Is the aim clear?
- Is a hypothesis written so that it can be disproved?
- Are the independent, dependent and controlled variables clearly defined?
- What are the strengths of the inquiry question?
- What questions do you have about the inquiry question?

1.1 Review

SUMMARY

- Before you begin your research, it is important to conduct a literature review. By utilising data from primary and/or secondary sources, you will better understand the context of your investigation to create an informed inquiry question.
- The purpose (aim) is a statement describing in detail what will be investigated.
- A hypothesis is a tentative explanation for an observation that is based on evidence and prior knowledge. A hypothesis must be testable and falsifiable and define a proposed relationship between two variables.
- Once an inquiry question has been chosen, stop to evaluate the question before progressing. The question may need further refinement or even further investigation before it is suitable as a basis for an achievable and worthwhile investigation. A major planning point is to not attempt something that is not possible to complete in the time available or with the resources on hand.
- There are three categories of variables: independent, dependent and controlled.

KEY QUESTIONS

1 Distinguish between the terms *inquiry question*, *hypothesis* and *purpose* of an investigation.

2 Which of the following describes an inquiry question?

A How does the diffraction pattern of light depend on the wavelength?

B The diffraction of light increases with an increase in wavelength.

C Diffraction and refraction have opposite trends with respect to wavelength.

D Diffraction is the bending of waves at a corner.

3 For each of the following hypotheses, select the independent and dependent variables.

a If water at 90°C is allowed to cool to room temperature in different shaped containers, the container with the largest surface area will reach room temperature in a shorter period of time.

b If the launch angle of a projectile is 45°, then the horizontal distance travelled will be a maximum.

c If you increase the thickness of foam bumpers attached to the front of a cart travelling at $1\,m\,s^{-1}$, then the force experienced during a head-on collision will be reduced.

d If the emf applied to a circuit consisting of ohmic resistors is increased, then the total current in the circuit will increase proportionally.

4 In an experiment about friction a student records the surface material that a ball is rolled over. What type of variable is the surface material?

5 Which of the following is the most specific inquiry question?

A Do helmets make riding a bike safer?

B Does the use of a foam helmet reduce the impact force when falling from a bike?

6 Which of the following inquiry questions is objective and specific?

A How does the angle of release affect the period of motion of a pendulum?

B Does the motion of a pendulum change when it is released from higher up?

C Pendulums make good metronomes.

7 Select the best hypothesis from the three options below. Give reasons for your choice.

A Hypothesis 1: If the mass on a spring is increased, then the length the spring is extended will increase proportionally.

B Hypothesis 2: Springs have a stiffness that is the force required to extended them 1 m.

C Hypothesis 3: Spring extension is measured in millimetres.

1.2 Planning investigations

After you have formulated your hypothesis, defined the purpose of your investigation and determined your variables, you will need to plan and design your investigation. Taking the time to carefully plan and design a practical investigation before beginning will help you to maintain a clear and concise focus throughout. Preparation is essential. This section is a guide to some of the key steps that should be taken when planning and designing a practical investigation.

CHOOSING AN APPROPRIATE TYPE OF INVESTIGATION

After you have drafted your inquiry question, purpose, hypothesis and variables, you will need to consider an appropriate type of investigation. Examples of various investigations are listed in Table 1.2.1. When selecting the type of investigation you will use, remember to consider how much time you have available, and whether you will need to work in a group or individually.

TABLE 1.2.1 Examples of different types of investigations.

Type of data/activity	Type of investigation	Example
Primary data	Design and conduct experiments	Measure and reliably compare the specific heat capacity for a range of metals.
	Test a claim	Heavier objects will go down a slide faster than light objects.
	Test a device	Test the efficiency of a solar panel.
Secondary data	Make a documentary or media report	Investigate the use of nuclear radiation in Australia.
	Conduct a literature review	Research applications of special relativity in everyday life.
	Develop an evidence-based argument	Create an argument to counter the claim that the Earth is flat.
	Write a journal article	Look at the style of a peer-reviewed journal and write a literature review in that particular style.
	Write an essay—historical or theoretical	Write about a Nobel Prize winner for Physics or the history of the development of the atomic model.
	Develop an environmental management plan	Create an environmental management plan for the safe handling of unstable nuclei based on your understanding of nuclear decay.
	Analyse a work of fiction or film for scientific relevance	Book: *The Martian*, Andy Weir (2011) and/or Film: *The Martian*, director Ridley Scott (2015)
	Create a visual presentation	Create a scientific poster.
	Investigate emerging technologies	Investigate perovskite solar cells.
Creating	Design and invent	Build your own hydro-generator.
	Create a working model	Create a working model of a door bell.
	Create a portfolio	Compile various resources to answer an inquiry question.
Field work	May be a starting point for an investigation	This could be initiated by the following stimuli: - an excursion - engagement with community experts.
Data analysis	Primary and secondary-sourced investigations	Use original data to construct and analyse graphs or tables. Analyse data from a variety of sources.

Organising information

It is important to be able to organise the information that you collect in your investigation. This is particularly important if your investigation is an in-depth literature review. Table 1.2.2 shows how you might be able to summarise information from primary and secondary sources.

TABLE 1.2.2 Example of categories that help you to keep track of information as you conduct a literature review.

Source of information	The effective temperatures of O-type stars from UV spectroscopy *Advances in Space Research Journal* Vol. 53, Issue 6, 15 March 2014, pp. 973–81
Author	Luciana Bianchi, Miriam Garcia
Country/region	International
Year of publication	2014
Sample size	3
Procedure	UV spectroscopy
Summary of conclusion	Use of UV spectroscopy to determine the parameters of the photosphere (surface) and stellar wind (ejected gases). Data fills in a gap when comparing effective temperature and spectral type. The downward trend was confirmed.
Key ideas	spectra, temperature, element formation

WRITING THE PROCEDURE

The procedure (also known as the method) of your investigation is a step-by-step description of how the hypothesis will be tested. Consider using a diagram of your equipment set-up such as the one shown in Figure 1.2.1.

When detailing the procedure, make sure it has the following elements so that it is a valid, reliable and accurate investigation.

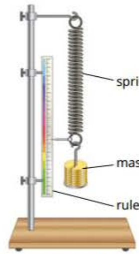


FIGURE 1.2.1 An equipment set-up for determining spring stiffness

REVISION

Methodology elements

- **Validity** refers to whether an experiment or investigation is in fact testing the set hypothesis and purpose. A valid investigation is designed so that only one variable is being changed at a time. To ensure validity, carefully determine the independent variable and how it will change, the dependent variable, and the controlled variables and how they will be maintained.
- **Reliability** refers to the notion that the experiment can be repeated many times and the average of the results from all the repeated experiments will be consistent. This can be maintained by defining the control and ensuring there is sufficient replication of the experiment.
- **Accuracy** is the ability to obtain the correct measurement. To obtain accurate results, you must minimise systematic errors.
- **Precision** is the ability to consistently obtain the same measurement. To obtain precise results, you must minimise **random errors**.

Build some testing into your investigation to confirm the accuracy and reliability of the equipment and your ability to read the information obtained. Reasonable steps to ensure the accuracy of the investigation include considering the unit in which the independent and dependent variables will be measured and the instruments that will be used to measure the variables.

Select and use appropriate equipment, materials and procedures. For example, select equipment that measures to smaller degrees to reduce uncertainty and repeat the measurements to confirm them.

Describe the materials and procedure in appropriate detail. This should ensure that every measurement can be repeated and the same result obtained within reasonable margins of experimental error (less than 5% is reasonable). **Percentage uncertainty**, sometimes referred to as percentage error, is a way to quantify how accurate a measurement is. This will be discussed in Section 1.4.

Recording numerical data

GO TO ➤ Section 1.4 page 17

When using measuring instruments, the number of significant figures (or digits) and decimal places you use is determined by how precise your measurements are. You will learn more about how to identify and use an appropriate number of significant figures and decimal places in Section 1.4.

Data analysis

Data analysis is part of the procedure. Consider how the data will be presented and analysed, for example in tables, so that patterns can be seen. Graphs can show relationships and enable comparisons. Preparing an empty table showing the data that needs to be obtained will help in the planning of the investigation.

The nature of the data being collected, such as whether the variables are qualitative or quantitative, influences the type of method or tool that you can use to analyse the data. The aims and the hypothesis will also influence the choice of analysis tool.

It is a good idea to draft a table of results before you commence an experiment. A sample results table is shown in Table 1.2.3. Table 1.2.4 shows how you might analyse the different types of variables in an investigation.

TABLE 1.2.3 Results table for investigating the specific heat capacity of copper using a calorimeter.

	Voltage applied to calorimeter (V)	Current through calorimeter (A)	Starting temperature (°C)	Final temperature (°C)	Time taken (s)
Trial 1					
Trial 2					
Trial 3					
Uncertainty of equipment	±0.1	±0.1	±0.1	±0.1	±0.1

TABLE 1.2.4 Types of variables and the ways in which they could be measured and used within an investigation.

State the inquiry question	What is the specific heat capacity of copper?
List the independent variable. Is the variable quantitative or qualitative?	The independent variable is the electrical energy being supplied to the calorimeter (quantitative).
List the dependent variable(s). What equipment will you use to measure these?	The dependent variable is the time taken to increase the temperature of the calorimeter contents by a given amount.
List the variables that you will control. What will you do to control these variables?	All mineral waters will be stored at room temperature until analysed. The temperature of each will be measured with an ethanol-filled glass thermometer. The uncertainty of the thermometer is ±0.1°C.

Sourcing appropriate materials and technology

When designing your investigation, you will need to decide on the materials, technology and instrumentation that will be used to carry out your research. It is important to find the right balance between items that are easily accessible and those which will give you accurate results. You will also need to consider any costs or risks associated with using the technology, how familiar you are with using it, and any limitations the technology has that will impact your investigation. As you move on to conducting your investigation, it will be important to take note of the precision of your chosen instrumentation and how this affects the accuracy and validity of your results. This will be discussed in greater detail in Section 1.3.

Data logging

Ask your teacher whether data-logging equipment might be available. Some schools have access to equipment such as motion sensors (Figure 1.2.2), temperature sensors and force sensors.

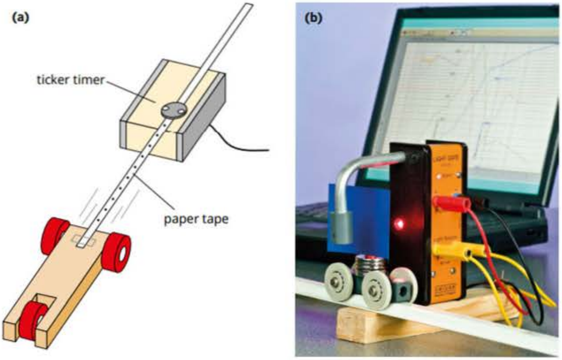


FIGURE 1.2.2 Determining the velocity of an object can be done with different levels of accuracy and precision. (a) Ticker tape can have an effect on the motion, changing the accuracy of the measurement, as well as the precision. (b) A motion sensor can measure the velocity with greater accuracy and precision.

Modifying the procedure

The procedure may need modifying as the investigation is carried out. The following actions will help to determine any issues in the procedure and how to modify them.

- Record everything.
- Be prepared to make changes to the approach.
- Note any difficulties encountered and the ways they were overcome. What were the failures and successes? Every test carried out can contribute to the understanding of the investigation as a whole, no matter how much of a disaster it may first appear.
- Do not panic. Go over the theory again, and talk to the teacher and other students. A different perspective can lead to a solution.

If the expected data is not obtained, don't worry. As long as it can be critically and objectively evaluated, the limitations of the investigation can be identified and further investigations proposed, the work is worthwhile.

COMPLYING WITH ETHICAL AND SAFETY GUIDELINES

Ethical considerations

Some investigations require an ethics approval; consult with the teacher. When you are planning an investigation, identify all parts of the procedure that might involve ethical considerations, and evaluate their necessity. You may be able to change the methodology to reduce or eliminate any possible ethical issues.

REVISION

Risk assessments

Risk assessments are performed to identify, assess and control hazards. A risk assessment should be performed for any situation, in the laboratory or outside in the field. Always identify the risks and control them to keep everyone as safe as possible.

Ways to reduce risk are, in order from most effective to least effective, elimination, substitution, engineering (modifying equipment), administrative control and **personal protective equipment**. Special care, such as wearing sunscreen, insect repellent, appropriate clothing and gloves as necessary, must be taken to minimise the risks that come with working outdoors. Someone with first-aid training should always be present, and any injuries or accidents must be immediately reported to your teacher or lab technician.

GO TO ➤ Year 11 Section 1.2

PHYSICSFILE EU

Ethics and radiation

There are many different types of radiation, including waves from the electromagnetic spectrum and alpha and beta particles released during the decay of radioactive nuclei. Some radiation occurs naturally, including cosmic radiation and the decay of unstable radioisotopes. Alpha particles are the nuclei of helium atoms, ${}^{4}_{2}He$, beta particles are electrons, and gamma rays are very-high-energy electromagnetic waves. Figure 1.2.3 shows that skin can stop the penetration of alpha particles. However, alpha particles are damaging to living tissues if ingested or inhaled, or if they enter a wound.

The decay of radioisotopes to produce different types of radiation can be represented by equations of balanced nuclear reactions. For example, the decay of naturally occurring uranium-238 produces thorium-234, an alpha particle and gamma radiation. This decay is represented by the following equation:

$${}^{238}_{92}U \rightarrow {}^{234}_{90}Th + {}^{4}_{2}He + \gamma$$

The Greek letter γ is used to represent gamma radiation. Notice that when the decay produces alpha or beta particles, a new element is also formed. Carbon-14 is a radioisotope that decays to produce beta particles, represented by the following equation:

$${}^{14}_{6}C \rightarrow {}^{14}_{7}N + {}^{0}_{-1}e$$

The following applications of radiation can be useful to society:

- the use of nuclear magnetic resonance, infrared, atomic absorption and UV-visible spectroscopy in chemical analysis
- radiotherapy as a cancer treatment
- use of X-rays to diagnose broken bones and dental decay
- the production of energy from nuclear fission reactions in power plants.

However, radiation can also be harmful to society:

- if nuclear weapons are used in war
- if nuclear waste contaminates water and food
- if medical practitioners and dentists are exposed to X-rays when imaging patients for medical purposes.

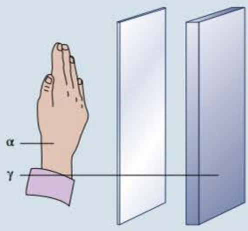


FIGURE 1.2.3 Skin can stop the penetration of alpha particles.

PEER REVIEW

Your teacher may suggest that you partner with a student in your class, to provide each other with constructive feedback regarding your planning.

Consider the following:

- Is the type of investigation appropriate and relevant to the inquiry question?
- Is the procedure written in easy-to-understand logical steps?
- Is the procedure valid and reliable?
- Has a thorough and complete risk assessment been undertaken before starting the investigation?
- What are the strengths of the type of investigation and procedure?
- What questions do you have about the type of investigation and procedure?

1.2 Review

SUMMARY

- The procedure of your investigation is a step-by-step plan. When detailing the procedure, ensure it complies as a valid, reliable and accurate investigation.
- It is also important to determine how many times the experiment needs to be replicated. Many scientific investigations lack sufficient repetition to ensure that the results can be considered reliable.
- Risk assessments must be carried out before conducting an investigation to make sure that when you carry out your procedure, you and others are kept safe. If you have elements of your investigation which are too high-risk you will need to re-evaluate your design.
- It is important to choose appropriate equipment for your experiment—not just personal protective equipment (PPE) that will help keep you safe, but also instrumentation that will give you accurate results.

KEY QUESTIONS

1 Give the correct term (accurate, reliable or valid) that describes an experiment with each of the following conditions.

- **a** The experiment addresses the hypothesis and aims.
- **b** The experiment is repeated and consistent results are obtained.
- **c** Appropriate equipment is chosen for the desired measurements.

2 A student wants to measure the range of a projectile on the school oval. They intend to be there for an hour in summer. Which of the following PPE should be included in the risk assessment for the activity on the oval? More than one response may be correct.

- **A** use of a fume cupboard
- **B** sunhat and sunscreen
- **C** gloves
- **D** eye goggles

3 A journal article reported the materials and procedure used to conduct an experiment. The experiment was repeated three times, and all values were reported in the results section of the article.

Repeating an experiment and reporting results supports:

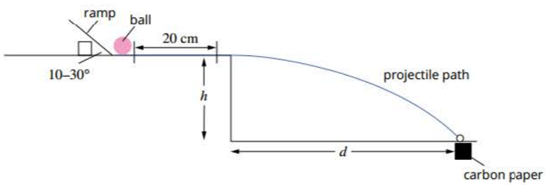
- **A** precision
- **B** reliability
- **C** accuracy
- **D** systematic errors

4 You are conducting an experiment to determine the effect of bumper foam density on the impact force during a collision.

Identify:

- **a** the independent variable
- **b** the dependent variable
- **c** at least one controlled variable.

5 A student included a diagram of the experimental set-up in her procedure. Use the diagram to write clear instructions for the procedure.



1.3 Conducting investigations

Once the planning and design of a practical investigation is complete, the next step is to undertake the investigation and record the results. As with the planning stages, there are key steps and skills to keep in mind to maintain high standards and minimise potential errors throughout the investigation (Figure 1.3.1).

This section will focus on the best methods of conducting a practical investigation, and of systematically generating, recording and processing data.

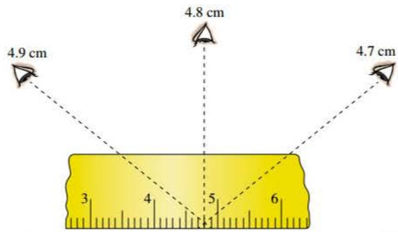


FIGURE 1.3.1 When reading analog devices, it is important to observe the meter or ruler straight on to avoid parallax errors.

REVISION

Collecting and recording data

For an investigation to be scientific, it must be objective and systematic. Always record all quantitative and qualitative data collected and the methods used to collect the data in your logbook. Also record any incident, feature or unexpected event that might have affected the quality or validity of the data. The recorded data is known as **raw data**. Usually this data needs to be processed in some manner before it can be presented.

Safe work practices

Remember to always employ safe work practices while conducting your experiment. See Section 1.2 for how to conduct risk assessments.

You will also need to keep in mind safe procedures to follow when disposing of waste—consult your teacher or education and government websites for information.

GO TO ➤ Year 11 Section 1.2

REVISION

GO TO ➤ Year 11 Section 1.3

Identifying errors

Errors can occur for a variety of reasons. Being aware of potential errors helps you to avoid or minimise them. As shown in Figure 1.3.2, there are three types of errors that can occur in an experiment.

Mistakes are avoidable errors. A measurement that involves a mistake must be rejected and not included in any calculations, or averaged with other measurements of the same quantity.

A **systematic error** is an error that is consistent and will occur again if the investigation is repeated in the same way. Systematic errors are usually a result of instruments that are not calibrated correctly or methods that are flawed. Whatever the cause, resulting errors are in the same direction for every measurement and the average will be either too high or too low as a result. These lead to bias. Examples of bias are shown in Figure 1.3.3.

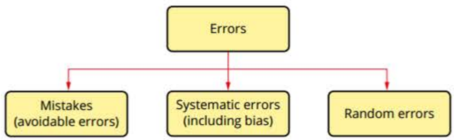


FIGURE 1.3.2 Types of errors that can be made in an experiment.

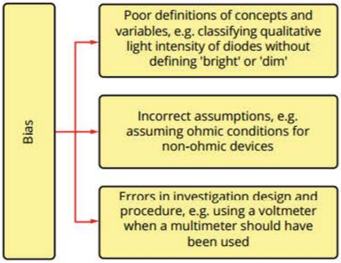


FIGURE 1.3.3 Types and examples of bias in an electric circuit analysis.

Random errors occur in an unpredictable manner and therefore follow no regular pattern. The measurement is sometimes too large and sometimes too small. The effects of random errors can be reduced by taking multiple measurements of the same quantity, then calculating an average.

Techniques to reduce error

Use appropriate equipment

Use the equipment that is best suited to the data that needs to be collected to validate the hypothesis. Determining the units of the data being collected and at what scale will help to select the correct equipment. Using the right units and scale will ensure that measurements are more accurate and precise (with smaller systematic errors).

To minimise errors, check the precision of the equipment that you intend to use. Voltmeters and ammeters often have different scales that can be used depending on the expected reading (Figure 1.3.4). It is always best to start with the largest scale, and adjust it until the measurement is in the smallest range that contains it.

Use calibrated equipment

Before carrying out the investigation, make sure the instruments or measuring devices are properly calibrated and are, in general, functioning correctly.

Use equipment correctly

Use the equipment properly. Ensure any necessary training has been done to use the equipment and that you have had an opportunity to practise using the equipment before beginning the investigation. Improper use of equipment can result in inaccurate, imprecise data with large errors, and the validity of the data can be compromised.

Incorrect reading of measurements is a common mistake. Make sure all of the equipment needed in the investigation can be used correctly and record the instructions in detail so they can be referred back to if the data doesn't appear correct.

Increase sampling size

In general, the larger the sample taken for analysis, the more precise the measured values will be. However, there are practical limits to this. For example, you will be limited by the size of the container and number of containers you can transport back to school, if collecting samples in the field.

Repeat the investigation

As discussed in Section 1.2, reliability is ensured by repeating your experiment. Modifications to your procedure may be needed before repeating the investigation to ensure all variables are being tested under the same conditions.

Reference secondary-sourced information

As you conduct your investigation, it is important to make note of any secondary-sourced information that you use. This will then be included in your written report. This is discussed in further detail in Section 1.7.

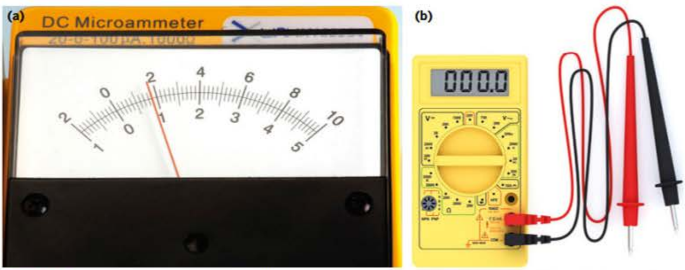


FIGURE 1.3.4 a) An analog ammeter with two different scales, and b) a digital multimeter with multiple scales to record current and voltage.

PEER REVIEW

Your teacher may suggest that you partner with a student in your class, to provide each other with constructive feedback regarding your data collection.

Consider the following:

- Is the data presented clearly in tables and/or graphs?
- Are all entries in the logbook dated?
- Are appropriate headings and units included?
- If anything unexpected occurred during data collection, was this recorded?
- What are the strengths of the data collected?
- What questions do you have about the data collected?

1.3 Review

SUMMARY

- It is essential that during the investigation you record the following in your logbook:
  - all quantitative and qualitative data collected
  - the methods used to collect the data
  - any incident, feature or unexpected event that might have affected the quality or validity of the data.
- Mistakes are avoidable errors, and measurements affected by mistakes should be discarded.
- A systematic error is an error that is consistent and will occur again if the investigation is repeated in the same way. Systematic errors are usually a result of instruments that are not calibrated correctly or methods that are flawed.
- Random errors occur in an unpredictable manner and are generally small. A random error could be, for example, the result of a researcher reading the same result correctly one time and incorrectly another time.

KEY QUESTIONS

1 Where should observations and measurements be recorded?

2 Identify the type of error that is described in each scenario below, and how it could be avoided.
  - **a** the electronic scale was not zeroed/tared before samples were weighed
  - **b** 50 g masses were used rather than 5 g masses
  - **c** one result is significantly lower than all the rest
  - **d** one student judged a water sample as 'cloudy', while another student recorded it as 'very cloudy'

3 The velocity of a cart after rolling down a ramp is being measured by timing the cart as it travels between two markers known to be 20 cm apart. A student doing this investigation recorded the following times with the same initial conditions: 0.25 s, 0.28 s, 0.02 s, 0.20 s, 0.22 s, 0.72 s. Identify the measurements that were a mistake. Find the average time.

4 Identify whether each error is a mistake, a systematic error or a random error.
  - **a** A weigh scale that should have measured a mass of 25.00 ± 0.03 g actually measured 24.92 ± 0.03 g.
  - **b** A student misread the value on the slotted mass for a Hooke's law investigation.
  - **c** A ball was weighed three times with the following results: 147.93 g, 147.94 g and 147.92 g.

1.4 Processing data and information

Once you have conducted your investigation and collected data, you will need to find the best way of collating this. This section is a guide to the different forms of representation that will help you to better understand your data.

RECORDING AND PRESENTING QUANTITATIVE DATA

Raw data is unlikely to be used directly to test the hypothesis. However, raw data is essential to the investigation and plans for collecting the raw data should be made carefully. Consider the formulae or graphs that will be used to analyse the data at the end of the investigation. This will help to determine the type of raw data that needs to be collected in order to test the hypothesis.

Once you have determined the data that needs to be collected, prepare a table in which to record the data.

REVISION

Significant figures

Significant figures are the numbers that convey meaning and precision. The number of significant figures used depends on the scale of the instrument. A significant figure is an integer or a zero that follows an integer. A zero that comes before any integers is not significant.

When you record raw data and report processed data, use the number of significant figures available from your equipment or observation. Using either a greater or smaller number of significant figures can be misleading. For example, Table 1.4.1 shows measurements of five samples taken using an electronic balance accurate to two decimal places. The data was entered into a spreadsheet to calculate the mean, which was displayed with four decimal places. You would record the mean as 5.69 g, not 5.6940 g, because two decimal places is the precision limit of the instrument. Recording 5.6940 g would be an example of false precision.

TABLE 1.4.1 An example of false precision in a data calculation.

Sample	1	2	3	4	5	Mean
Mass (g)	5.67	5.75	5.62	5.71	5.72	5.6940

Multiplying or dividing

When multiplying or dividing measurements, report the final value to the least number of significant figures used within that calculation. For example, to calculate the resistance, in ohms, of a component using the potential difference and current, the formula is $V = IR$. If $V = 5.62$ V (three significant figures), and the current (I) is 1.465 A (four significant figures), the resistance, in ohms, should be reported to the least number of significant figures, i.e. to three significant figures ($V = IR$; $5.62 = 1.465 \times R$; $R = 3.84\,\Omega$ to three significant figures).

Although digital scales can measure to many more than two figures and calculators can give 12 figures, be sensible and follow the significant figure rules.

Adding or subtracting

When adding or subtracting measurements, report the calculated value to the least number of decimal places.

For example, when calculating the net force on an object, the following measurements were recorded.

Weight of tennis ball: 0.5733 N

Air resistance (frictional force): 0.93 N

Net force = 0.93 − 0.5733 = 0.3567 = 0.36 N. The original answer of 0.3567 N is an example of false precision as one value in the data was not measured to five decimal places. The final answer must be stated to the smallest number of decimal places seen in the data, i.e. 0.36 N.

Decimal places

As with significant figures, you must be careful to record your measurements to the precision of the equipment used. If a weighing balance can report a mass measurement to two decimal places, you should record your value to two decimal places.

When adding or subtracting measurements, report the calculated value to the least number of decimal places used in the calculation. For example, in an investigation the following measurements were recorded:

Sample A 4.93

Sample B 5.54

Sample C 4.82

The average of these results is

$$\frac{4.93 + 5.54 + 4.82}{3} = 5.0966666667$$

As the sample results are all to two decimal places, this average also needs to be rounded to two decimal places, i.e. 5.10. It would be misleading to report the average to a greater number of decimal places.

ANALYSING AND PRESENTING DATA

The raw data that has been obtained needs to be presented in a way that is clear, concise and accurate.

There are a number of ways of presenting data, including tables, graphs, flow-charts and diagrams. The best way of visualising the data depends on its nature. Try several formats before making a final decision, to create the best possible presentation.

REVISION

GO TO ➤ Year 11 Section 1.4

Presenting raw and processed data in tables

Tables organise data into rows and columns and can vary in complexity according to the nature of the data. Tables can be used to organise raw data and processed data or to summarise results.

The simplest form of a table is a two-column format. In a two-column table, the first column should contain the independent variable and the second column should contain the dependent variable.

Tables should have a descriptive title, column headings (including the unit) and numbers aligned to the decimal point.

A table of processed data usually presents the average values of data, the **mean**. However, the mean on its own does not provide an accurate picture of the results.

To report processed data more accurately, the **uncertainty** should be presented as well.

Uncertainty

When there is a range of measurements of a particular value, the mean must be accompanied by the uncertainty for your results to be presented as a mean in an accurate way.

Uncertainty is calculated as:

uncertainty = ± maximum variance from the mean

Percentage uncertainty

Percentage uncertainty (also known as percentage error) is a way to quantify how accurate a measurement is. To calculate the percentage uncertainty of your data you take the uncertainty and divide it by the measurement, then multiply by 100.

Other descriptive statistical measures

Other statistical measures that can be used, depending on the data obtained, are:

- **mode**: the mode is the value that appears most often in a data set. This measure is useful to describe qualitative or discrete data.
- **median**: the median is the 'middle' value of an ordered list of values. The median is used when the data range is spread, for example, due to the presence of unusual results, making the mean unreliable.

Table 1.4.2 is a guide to choosing the most appropriate measure of central tendency for a data set.

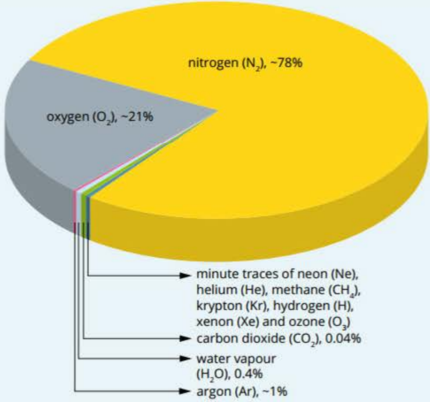
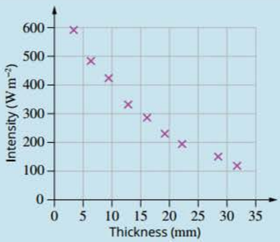
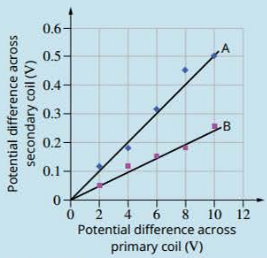
TABLE 1.4.2 The most appropriate measure of central tendency to use in descriptive statistics depends on the type of data.

Type of data	Mode	Median	Mean
nominal (qualitative)	✓	✗	✗
ordinal (qualitative)	✓	✓	maybe
discrete or continuous (quantitative)	✓	✓	✓

Graphs

It is important that you choose an appropriate graph type to suit the data that you have collected. Table 1.4.3 summarises suitable graphs for discrete or continuous quantitative data.

TABLE 1.4.3 Suitable graph types for quantitative data.

Type of data	Appropriate type of graph	Examples
discrete	**bar graph**, histogram or **pie diagram**	 **FIGURE 1.4.1** Pie chart showing distribution of elements from a sample analysis.
continuous	line graph or **scatter plot**, including a **trend line** if appropriate	 **FIGURE 1.4.2** Scatter plot of the radiation intensity through different thickness material.  **FIGURE 1.4.3** Graph, including trend line, showing results of an investigation into two different transformers to see which is better to use with a model motor.

In general, tables provide more detailed data than graphs, but it is easier to observe trends and patterns in data in graphical form than in tabular form.

Graphs are used when two variables are being considered and one variable is dependent on the other. The graph shows the relationship between the variables.

There are several types of graphs that can be used, including line graphs, bar graphs and pie diagrams. The best one to use will depend on the nature of the data.

General rules to follow when making a graph (Figure 1.4.4) include the following:

- Keep the graph simple and uncluttered.
- Use a descriptive title.
- Represent the independent variable on the *x*-axis and the dependent variable on the *y*-axis.
- Make axes proportionate to the data.
- Clearly label axes with both the variable and the unit in which it is measured.

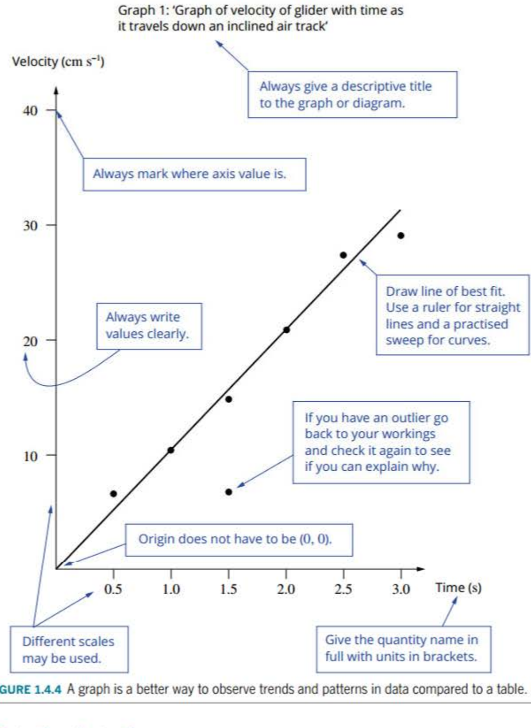


FIGURE 1.4.4 A graph is a better way to observe trends and patterns in data compared to a table.

REVISION

Outliers

Sometimes when the data is collected, there may be one point that does not fit with the **trend** and is clearly an error. This is called an **outlier**. An outlier is often caused by a mistake made in measuring or recording data, or from a random error in the measuring equipment. If there is an outlier, include it on the graph, but ignore it when adding a trend line.

Distorting the truth

Poorly constructed graphs can alter your perception of the data. For example, in Figure 1.4.5 you can see two graphs that show the same data—the measurement of the period of oscillation of a pendulum with different masses.

One graph exaggerates the difference in the period of the pendulum by using only a scale of 1.49 to 1.59 s on the *y*-axis. It is important to make sure the graphs you create do not misrepresent your data in this way. You should also be wary of distorted data when interpreting graphs in other publications.

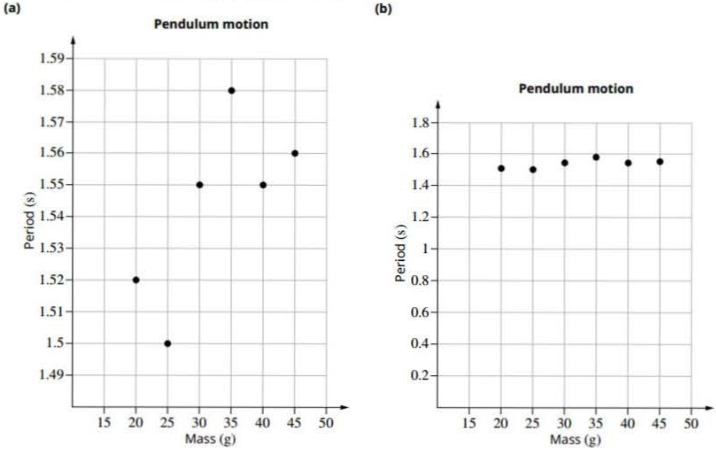


FIGURE 1.4.5 (a) The graph shows the difference in oscillation period times with different masses, with a scale on the *y*-axis that highlights the difference in measurements. (b) The graph shows the same measurements using a different *y*-axis scale.

PEER REVIEW

Your teacher may suggest that you partner with a student in your class, to provide each other with constructive feedback regarding the processing of your data.

Consider the following:

- Is the processed data presented clearly in tables and/or graphs?
- Have any outliers been identified?
- Has the appropriate type of graph been selected to display the processed data?
- Are appropriate headings and units included?
- What are the strengths of the processed data collected?
- What questions do you have about the processed data collected?

1.4 Review

SUMMARY

- The number of significant figures and/or decimal places used depends on the scale of the instrument used. It is important to record data to the number of significant figures or decimal places available from the equipment or observation.
- Consider how the data will be presented and analysed. A wide range of analysis tools could be used. For example, tables organise data so that patterns can be established, and graphs can show relationships and comparisons.
- The simplest form of a table is a two-column format in which the first column contains the independent variable and the second column contains the dependent variable.
- When there is a range of measurements of a particular value, the mean must be accompanied by the uncertainty for your results to be presented as a mean in an accurate way.
- General rules to follow when making a graph include the following:
  - Keep the graph simple and uncluttered.
  - Use a descriptive title.
  - Represent the independent variable on the *x*-axis and the dependent variable on the *y*-axis.
  - Make the axes proportionate to the data.
  - Clearly label axes with both the variable and the unit in which it is measured.

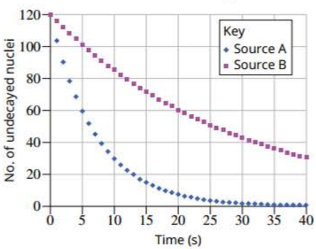
KEY QUESTIONS

1 For the data set 21, 27, 19, 21, 24, 26, 22, determine:
a the mean
b the mode
c the median.

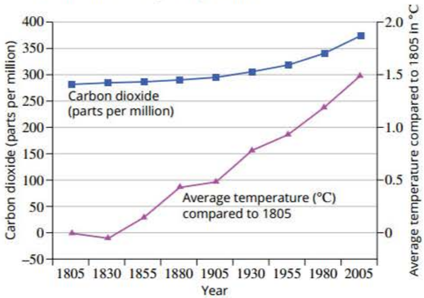
2 **a** How many significant figures are in the value 22.06 mL?
b When multiplying or dividing, how many significant figures should be reported for the calculation?
c When adding or subtracting, how many decimal places should be reported for the calculation?

3 Which axis should be used to represent the:
a dependent variable
b independent variable?

4 A graph of the radioactive decay of two different elements is shown. How many un-decayed nuclei are there in each source after 20 s?



5 The following graph describes both the amount of carbon dioxide present in the atmosphere over time as well as the average temperature.



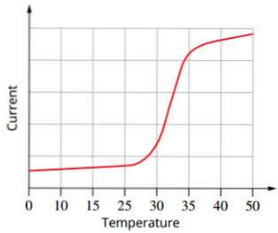
a Describe the trend of carbon dioxide in the atmosphere, and the average global temperature from 1805 to 2005.
b What conclusions can be drawn from this data?

6 An investigation was done into the temperature increase caused by a known energy increase. The following data was obtained.

Energy added (kJ)	Increase in temperature (°C)
0	0
250	4
500	9
750	15
1000	20
1250	30
1500	31
1750	35
2000	39
2250	46
2500	51
unknown	26

a Plot the data on a scatter plot, using graph paper or a spreadsheet program.
b Define the term outlier.
c Identify any outliers in this set of data.
d Draw a trend line.
e Use the graph to determine the unknown amount of energy that resulted in the 26°C temperature increase.

7 Describe at least four ways the graph below could be improved.



8 You are measuring the range of a projectile on the school oval. Discuss the accuracy and precision of your results if you are:
a using a large stride that you initially test against a metre ruler
b using a trundle wheel
c using a long tape measure that anchors to the ground where the projectile launched.

FIGURE 1.5.1 To discuss and conclude your investigation, analyse the raw and processed data.

1.5 Analysing data and information

Now that the chosen topic has been thoroughly researched, the investigation has been conducted and data collected, it is time to draw it all together. You will now need to analyse your results (Figure 1.5.1) to better understand the physical processes behind them.

FACTORS THAT CAN AFFECT THE INTERPRETATION, ANALYSIS AND UNDERSTANDING OF DATA

Correlation and causation

You need to be careful to distinguish between a correlation between two variables and a cause-and-effect relationship. For example, Figure 1.5.2 shows a direct correlation between the amount of cheese consumed in the USA per person and the number of people who died in the USA by becoming tangled in their bedsheets. However, we cannot assume that consumption of cheese increases the risk of death from entanglement in bedsheets. In other words, two sets of data might be correlated but have no relationship to each other (there is no causation).

Per capita cheese consumption
correlates with
number of people who died by becoming tangled in their bedsheets

Correlation: 94.71% ($r = 0.947091$)

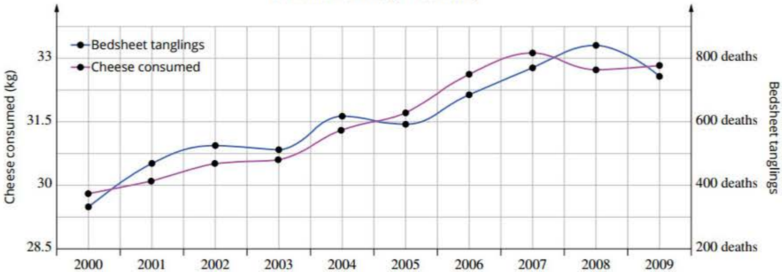


FIGURE 1.5.2 Correlation does not always mean causation.

It is important to consider whether the relationship between data can be due to other factors. For example, reports of sales of both ice-cream and sunscreen being greater in Sydney in February compared to in July could be a consequence of the hotter weather, instead of any relationship between the sales of each product.

Evaluating the data

Some useful questions to consider when interpreting data include:

- Has the original question been answered?
- Do the results meet expectations? Do they make sense?
- What are the main conclusions? Are there other interpretations?
- Is the supporting data of sufficient quality? How current is it? How was it collected?
- Can the results be supported statistically, i.e. are they statistically significant?

REVISION

Explaining results in the discussion

The key sections of the discussion are:

- analysing and evaluating data
- evaluating the investigative method
- explaining the link between the investigation findings and the relevant physics concepts.

Consider the message to be conveyed to the audience when writing the discussion. At the conclusion of the discussion, the audience must have a clear idea of the context, results and implications of the investigation.

GO TO ➤ Year 11 Section 1.5

Analysing and evaluating data

In the discussion, the findings of the investigation need to be analysed and interpreted.

- State whether a pattern, trend or relationship was observed between the independent and dependent variables. Describe what kind of pattern it was and specify under what conditions it was observed.
- Were there discrepancies, deviations or anomalies in the data? If so, these should be acknowledged and explained.
- Identify any limitations in the data you have collected. Perhaps a larger sample or further variations in the independent variable would lead to a stronger conclusion.

Remember that the results may be unexpected. This does not make the investigation a failure. However, the findings must be discussed and linked back to the hypothesis, aims and method.

Describing more complex trends

You may find that you need to describe a more complex trend in data than a simple line or exponential relationship. Figure 1.5.3 shows a current–temperature curve of a thermistor. This trend shows a slow increase in current when the temperature increases up to 35°C. Then a rapid increase in current occurs when the temperature is increased to 50°C. After 50°C, the current slowly increases as the temperature increases.

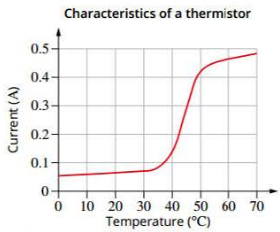


FIGURE 1.5.3 Current–temperature graph showing how current flowing through a thermistor describes a more complex relationship to temperature than simply linear or exponential.

Reading information from a graph

An important skill is being able to extract information from a graph. Figure 1.5.4 shows the current–voltage relationships at different light levels (lux) for an electrical component called a photodiode.

Photodiodes react to the environment, allowing different levels of current through depending on the potential difference across them, and the light intensity landing on them. These graphs can be used to design a circuit, or to gain information about the light landing on the diode. Worked example 1.5.1 shows the steps to use this information to find the current in a circuit.

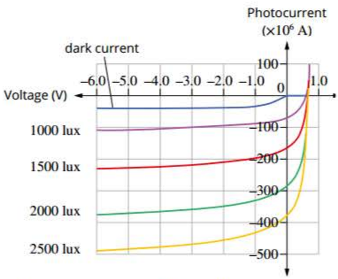


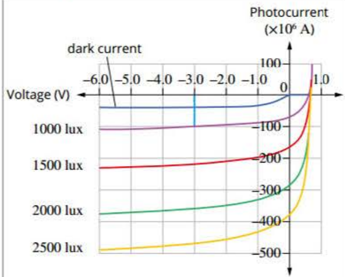
FIGURE 1.5.4 Line graph showing continuous quantitative data.

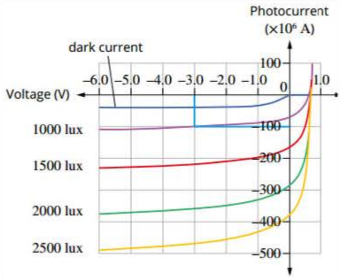
Worked example 1.5.1

READING INFORMATION FROM A GRAPH

Using the graph in Figure 1.5.4, determine the current at 1000 lux in a circuit with a potential difference of –3.0 V.

<table>
<tr><th>Thinking</th><th>Working</th></tr>
<tr><td>Locate the appropriate lux line.

Draw a vertical line from the voltage on the x-axis to intersect with the curve of the appropriate lux.</td><td>
</td></tr>
<tr><td>Draw a horizontal line from the intersection point of the curve to the y-axis.

The point on the y-axis represents the current after the photodiode.</td><td>
</td></tr>
<tr><td>Read the current, taking note of the units on the axis.</td><td>The horizontal line intersects the y-axis at -100×10^{-6} A. Therefore -1×10^{-4} A will flow through the wire after the photodiode.</td></tr>
</table>

Worked example: Try yourself 1.5.1

READING INFORMATION FROM A GRAPH

Using the graph in Figure 1.5.4, determine the current at 2000 lux in a circuit with a potential difference of –2.0 V.

REVISION

GO TO ➤ Year 11 Section 1.5

Mathematical models

After analysing your data using tables and graphs, it might be possible to find a mathematical relationship to describe your results. For example, your graph may produce a straight line so that there is some sort of linear relationship between the two variables.

Graphing a linear relationship

When analysing data from a linear relationship, it is first necessary to obtain a graph of the data and an equation for the line that best fits the data. The entire process can be done on paper but most people will use a computer spreadsheet, a scientific or a graphics calculator, or some other computer-based process. In what follows, it is not assumed that you are using any particular technology.

If you are plotting your graph on paper, then proceed as follows:

1. Plot each data point on clearly labelled, unbroken axes.
2. Identify and label, but otherwise ignore, any suspect data points.
3. Draw, by eye, the 'line of best fit' for the points. The points should be evenly scattered either side of the line.
4. Locate the vertical axis intercept and record its value as 'c'.
5. Choose two points on the line of best fit to calculate the gradient. Do not use two of the original data points as this will not give you the gradient of the line of best fit.
6. Write the equation for a straight line $y = mx + c$, replacing x and y with appropriate symbols, and use this equation for any further analysis.

If you are using a computer or a graphics calculator, then proceed as follows:

1. Plot each data point on clearly labelled, unbroken axes.
2. Identify suspect data points and create another data table without the suspect data.
3. Plot a new graph without the suspect data. Keep both graphs as you don't actually discard the suspect data but you do eliminate it from the analysis.
4. Plot the line of best fit. How you do this depends on the model of calculator or the software being used.
5. Compute the equation of the line of best fit that will give you values for gradient, m, and vertical axis intercept, c.
6. Write $y = mx + c$, replacing x and y with appropriate symbols, and use this equation for any further analysis.

Manipulating non-linear data

Suppose you were examining the relationship between two quantities, B and d, and had good reason to believe that the relationship between them is:

$$B = \frac{k}{d}$$

where k is some constant value. Clearly, this relationship is non-linear and a graph of B against d will not be a straight line. By thinking about the relationship, it can be seen that in 'linear form':

$$\begin{array}{l} B = k\,\frac{1}{d} \\ \uparrow \quad \uparrow \uparrow \\ y = m\;x + c \end{array}$$

A graph of B (on the vertical axis) against $\frac{1}{d}$ (on the horizontal axis) will be linear. The gradient of the line will be k and the vertical intercept, c, will be zero. The line of best fit would be expected to go through the origin because, in this case, there is no constant added and so c is zero.

In the above example, a graph of the raw data would just show that B is larger as d is smaller. It would be impossible to determine the mathematical relationship just by looking at a graph of the raw data.

By choosing a possible relationship based on the shape of the initial graph and your knowledge of various mathematical forms, it is possible to work out how the data must be manipulated to give a linear graph. Then, by following the same steps above for graphing a linear relationship, it is possible to find a mathematical model for non-linear data. It may be necessary to try several mathematical forms to find one that seems to fit the data.

REVISION

GO TO ➤ Year 11 Section 1.5

Evaluating the method

It is important to discuss the limitations of the investigation method. Evaluate the method and identify any issues that could have affected the validity, accuracy, precision or reliability of the data. Sources of errors and uncertainty must also be stated in the discussion.

Once any limitations or problems in the procedure have been identified, recommend improvements on how the investigation could be conducted if repeated.

Bias

Bias may occur in any part of the investigation method, including sampling and measurements. Bias is a form of systematic error resulting from the researcher's personal preferences or motivations. There are many types of bias, including poor definitions of both concepts and variables, incorrect assumptions and errors in the investigation design and procedure. Some biases cannot be eliminated, but should at least be addressed in the discussion.

Accuracy

In the discussion, evaluate the degree of accuracy and precision of the measurements for each variable of the hypothesis. Comment on the uncertainties obtained.

When relevant, compare the chosen method with any other methods that might have been selected, evaluating the advantages and disadvantages of the selected method and the effect on the results.

Validity

Validity refers to whether an experiment or investigation actually tests the set hypothesis and aims. Factors influencing validity include:

1 whether your experiment measures what it claims to measure (i.e. your experiment should test your hypothesis)
2 whether the independent variable influenced the dependent variable in the way you thought it would (i.e. the certainty that something observed in your experiment was the result of your experimental conditions, and not some other cause that you did not consider)
3 the degree to which your findings can be generalised.

Reliability

When discussing the results, indicate the range of the data obtained from replicates. Explain how the sample size was selected. Larger samples are usually more reliable, but time and resources might have been scarce. Discuss whether the results of the investigation have been limited by the sample size.

The control group is important to the reliability of the investigation, as it helps determine if a variable that should have been controlled has been overlooked and may explain any unexpected results.

Error

Discuss any source of systematic or random error. When limitations of the method and results have been identified, suggest ways of improving the investigation.

CRITICALLY EVALUATING RESOURCES

Not all sources are **credible**. It is essential to critically evaluate the content and its origin. Questions you should always ask when evaluating a source include:

- Who created this message? What are the qualifications, **expertise**, **reputation** and **affiliation** of the authors?
- Why was it written?
- Where was the information published?
- When was the information published?
- How often is the information referred to by other researchers?
- Are the conclusions supported by data or evidence?
- What is implied?
- What is omitted?
- Are any opinions or values being presented in the piece?
- Is the writing objectively and accurately describing a scientific concept or **phenomenon**?
- How might other people understand or interpret this message differently from me?

When evaluating the validity or bias of websites consider its domain extension:

- .gov stands for government
- .edu stands for education
- .org stands for non-profit organisations
- .com stands for commercial/business.

When conducting a literature review, it can be useful to summarise your findings in a table, such as that shown in Table 1.5.1.

TABLE 1.5.1 A table can be used to help summarise information while researching. This example centres around the type of insulation used to increase temperature stability in domestic housing.

Type of insulation	Notes	Reference	Reliability of source
Wool batts	Traditional insulation around the world. Wool and wool–hemp have varying properties due to the differing densities. They absorb water which also affects their performance. A renewable source that has been used for decades.	Ye, Z., Wells, C. M., Carrington, C. G. and Hewitt, N. J. (2006). Thermal conductivity of wool and wool–hemp insulation. *Int. J. Energy Res.*, 30: 37–49. doi:10.1002/er.1123	Credible primary source of data in peer-reviewed journal—research sponsored by New Wool Products.
Reeds	Thermal conductivity of 0.045–0.056 $\mathrm{W\,m\,K^{-1}}$, hollow centres create air gaps that are great insulators. Great acoustic insulation, although potentially a fire hazard without proper treatment.	Francesco Asdrubali, Francesco D'Alessandro, Samuele Schiavoni, (2015). A review of unconventional sustainable building insulation materials, *Sustainable Materials and Technologies*, Vol. 4, pp. 1–17.	Credible—peer-reviewed journal with primary sources.
Cattails	Mould resistant, can deal with water without rotting. Very stable with soft sponge structure that creates low thermal conductivity. This is a high-yield plant that has a very low carbon footprint.	Ayre, J. (2013). 'Cattails as Building Insulation—New Research Brings to Light the Great Advantages of the Material'. https://cleantechnica.com/2013/06/14/cattails-as-building-insulation-new-research-brings-to-light-the-great-advantages-of-the-material/	Journal article, secondary source. Opinion piece in Clean Technica magazine.
Recycled denim	Although marketed in a mainstream way, the testing does not back up claims that this product is as thermally efficient as standard glass-fibre batts. Additionally, cotton is a high-resource crop, and insulation being sold as recycled are only partially recycled and partially pre-consumer.	Simi Hoque (2010). Net Zero Energy Homes: An Evaluation of Two Homes in the Northeastern United States. *Journal of Green Building:* Spring 2010, Vol. 5, No. 2, pp. 79–90.	Credible. Peer-reviewed journal with primary data.

PEER REVIEW

Your teacher may suggest that you partner with a student in your class, to provide each other with constructive feedback regarding the analysis of your data.

Consider the following:

- Is the data accurately discussed?
- Have any limitations of the method or data collection been identified?
- Have any recommendations been made to improve the investigation if time permitted repeating the study?
- What are the strengths of the data analysis?
- What questions do you have about the data analysis?

FIGURE 1.5.5 Honest evaluation and reflection play important roles in analysing the procedure.

1.5 Review

SUMMARY

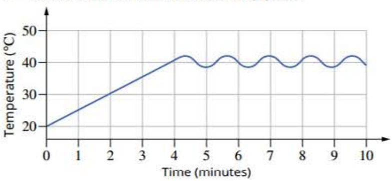
- After completing your investigation you will need to analyse and interpret your data. A discussion of your results is required where the findings of the investigation need to be analysed and interpreted.
  - State whether a pattern, trend or relationship was observed between the independent and dependent variables. Describe what kind of pattern it was and specify under what conditions it was observed.
- If possible, create a mathematical model to describe your data.
  - Were there discrepancies, deviations or anomalies in the data? If so, these should be acknowledged and explained.
  - Identify any limitations in the data collected. Perhaps a larger sample or further variations in the independent variable would lead to a stronger conclusion.
- It is important to discuss the limitations of the investigation method. Evaluate the method and identify any issues that could have affected the validity, accuracy, precision or reliability of the data. Sources of errors and uncertainty must also be stated in the discussion.
- When discussing the results, indicate the range of the data obtained from replicates. Explain how the sample size was selected. Larger samples are usually more reliable, but time and resources are likely to have been scarce. Discuss whether the results of the investigation have been limited by the sample size.

KEY QUESTIONS

1 What types of graphs would be suitable for displaying discrete data?

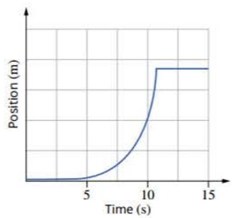
2 List five important features to put on a line graph when displaying results.

3 Describe the trend in the following graph:

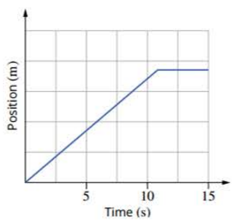


4 Which of the following graphs shows a constant (non-zero) acceleration for the first 10 s and then the object continuing at a constant (non-zero) velocity?

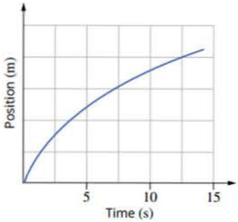
A



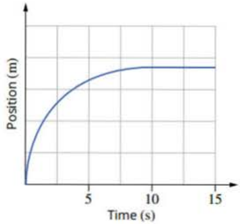
B



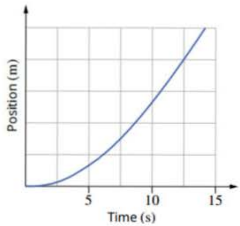
C



D



E



5 List features of a credible primary or secondary source.

1.6 Problem solving

Having analysed your results, you can then apply them to physics concepts in order to evaluate your conclusions. In this section you will learn how analysing your investigation leads to a better understanding of the underlying scientific principles of your research.

MODELLING

Table 1.6.1 lists examples of investigations that could be used to explain phenomena, make predictions and solve problems.

TABLE 1.6.1 Examples of models that can be used in investigations.

Type of investigation	Application	Example
using mathematical modelling to make predictions	using formulae	Calculate the period of a pendulum swinging with a string length of 10 cm, moving through an angle of 30°.
using computer modelling	computer simulations used in numerical analysis	Calculate the distance travelled in a lift from the accelerometer data on your phone.
critiquing claims made by the media	compare emotive fear-based articles with evidence-based claims	Evaluate the effect of human CO_2 levels on the climate of the Earth using thermodynamics.

LITERATURE REVIEW

Table 1.6.2 shows an example of how summarising information in a table can help you identify patterns in a literature review.

TABLE 1.6.2 An in-progress summary of information obtained through a literature review regarding the thermal equilibrium of the Earth.

Source	Author	Country/ region	Year of publication	Type	Key ideas	Credibility
NASA feature article, World of Change: Global Temperatures https://earthobservatory.nasa.gov/Features/WorldOfChange/decadaltemp.php	M. Carlowicz	USA	2010	review	Global temperature change is different to local weather. Temperature has been increasing faster since 1980 than previously.	author is a qualified scientist, credible science organisation
Climate Science Needs Effective Images. *Journal of Earth Science and Climate Change,* 2017, 8: p. 410. DOI: 10.4172/2157-7617.1000410	J. E. Thompson		2017	article	Using effective images to simplify and communicate science could increase acceptance of the science.	recent article, no critical data
Changes in the Flow of Energy through the Earth's Climate System, American Physical Society, https://www.aps.org/units/fps/newsletters/200904/trenberth.cfm	Kevin E. Trenberth			article		qualified author, not peer reviewed
Latitudinal temperature gradients and climate change. DOI: 10.1029/97JD03649	D. Rind		1998	article	Gradient of temperature from equator to pole becoming steeper.	qualified authors, peer reviewed
http://www.bom.gov.au/climate				data	Australian/global climate data. Includes trends and mapping of extreme weather events.	qualified authors, credible science organisation

DISCUSSING RELEVANT PHYSICS CONCEPTS

To make the investigation more meaningful, it should be explained within the right context, meaning the related physics ideas, concepts, theories and models. Within this context, explain the basis for the hypothesis.

Relating your findings to a physics concept

During the analysing stage of your investigation (Section 1.5), you were able to find trends, patterns and mathematical models of your results. This is the framework needed in which to discuss whether the data supported or refuted the hypothesis. Ask questions such as:

- Was the hypothesis supported?
- Has the hypothesis been fully answered? (If not, give an explanation of why this is so and suggest what could be done to either improve or complement the investigation.)
- Do the results contradict the hypothesis? If so, why? (The explanation must be plausible and must be based on the results and previous evidence.)

Providing a theoretical context also enables comparison of the results with existing relevant research and knowledge. After identifying the major findings of the investigation, ask questions such as:

- How does the data fit with the literature?
- Does the data contradict the literature?
- Do the findings fill a gap in the literature?
- Do the findings lead to further questions?
- Can the findings be extended to another situation?

Be sure to discuss the broader implications of the findings. Implications are the bigger picture. Outlining them for the audience is an important part of the investigation. Ask questions such as:

- Do the findings contribute to or impact on the existing literature and knowledge of the topic?
- Are there any practical applications for the findings?

DRAWING EVIDENCE-BASED CONCLUSIONS

A conclusion is usually a paragraph that links the collected evidence to the hypothesis and provides a justified response to the research question. Indicate whether the hypothesis was supported or refuted and the evidence on which this is based (that is, the results). Do not provide irrelevant information. Only refer to the specifics of the hypothesis and the research question and do not make generalisations.

What type of evidence is needed to draw valid conclusions?

A valid conclusion can be drawn if:

- the method was designed to obtain data to answer the aim and hypothesis
- only one independent variable was changed
- the dependent variable was measured or observed
- all other variables were controlled
- the experiment was replicated
- the method could be followed by another person
- the data obtained was accurate and precise
- there were no significant limitations with the method and data obtained
- any links were made between the data obtained and chemical theory.

Read the examples of weak and strong conclusions in tables 1.6.3 and 1.6.4 for the hypothesis and research question shown.

TABLE 1.6.3 Examples of weak and strong conclusions to the hypothesis.

Hypothesis: If the launch angle of a projectile is increased, then the flight range will also increase.	
Weak conclusion	**Strong conclusion**
The increase in launch angle increased range, and then the range decreased.	An increase in launch angle from 10° to 46° resulted in an increase in projectile range from 30 cm to 150 cm. Increasing the launch angle from 46° to 70° decreased the range from 150 cm to 35 cm.

TABLE 1.6.4 Examples of strong and weak conclusions in response to the inquiry question.

Inquiry question: Does the temperature of a wire affect its electrical resistance?	
Weak conclusion	**Strong conclusion**
The results show that temperature does affect the resistance.	Analysis of the results on the effect of an increase in the temperature of wire from 5°C to 40°C showed an inverse relationship in which the electrical resistance of 2 mm diameter copper wire decreased from $1.7 \times 10^{-6}\,\Omega$ to $1.1 \times 10^{-6}\,\Omega$. These results support the current knowledge that an increase in the temperature of copper results in a decrease in its resistance.

REVISION

GO TO ➤ Year 11 Section 1.6

Interpreting scientific and media texts

Sometimes you may be required to investigate claims and conclusions made by other sources, such as scientific and media texts. As discussed in Section 1.4, some sources are more credible than others. Once you have analysed the validity of the primary or secondary text, you will be able to follow the same steps described above in evaluating their conclusions in order to solve scientific problems.

PEER REVIEW

Your teacher may suggest that you partner with a student in your class, to provide each other with constructive feedback regarding the analysis of your data.

Consider the following:

- Are the findings accurately discussed in relation to relevant physics concepts?
- Is the conclusion strong?
- Is the conclusion directed to the aim, hypothesis and inquiry question?
- What are the strengths of the conclusion?
- What questions do you have about the conclusion?

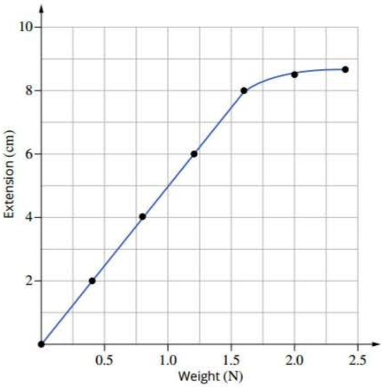
1.6 Review

SUMMARY

- To make the investigation more meaningful, it should be explained within the right context, outlining the related physics ideas, concepts, theories and models. Within this context, explain the basis for the hypothesis.
- Indicate whether the hypothesis was supported or refuted and on what evidence this is based (that is, the results). Do not provide irrelevant information. Only refer to the specifics of the hypothesis and the research question and do not make generalisations.

KEY QUESTIONS

1 The following results were obtained in an experiment testing a spring to determine the spring constant.



$$k = \frac{\text{rise}}{\text{run}} = \frac{8}{1.6} = 5\ \text{cm}\,\text{N}^{-1}$$

What sentence is the best concluding statement? Give a reason why the other options are not appropriate.

A All springs have a stiffness of $5\,\text{cm}\,\text{N}^{-1}$ when a force is applied between 0 N and 1.6 N.

B The spring stiffness is always measured in $\text{cm}\,\text{N}^{-1}$.

C The stiffness of the spring tested was $5\,\text{cm}\,\text{N}^{-1}$ for forces between 0 N and 2.5 N.

D For forces between 0 N and 1.6 N, the spring tested had a stiffness of $5\,\text{cm}\,\text{N}^{-1}$.

2 A student determined that the experimental relationship between the range (in m) of a projectile and the initial velocity (in $\text{m}\,\text{s}^{-1}$) was $R = 7.2 \times 10^{-2} \times v^2$. They found this by measuring the distance travelled at a range of velocities between $1.3\,\text{m}\,\text{s}^{-1}$ and $5.8\,\text{m}\,\text{s}^{-1}$.

a Is it appropriate to use this equation to make a prediction of the range for a velocity of $4.0\,\text{m}\,\text{s}^{-1}$ for the same ball and launch mechanism? If it is, what does the equation predict?

b Is it appropriate to use this equation to make a prediction of the range for a velocity of $14.0\,\text{m}\,\text{s}^{-1}$ for the same ball and launch mechanism? Give a reason for your answer. If it is, what does the equation predict?

c Is it appropriate to use this equation to make a prediction of the range for a velocity of $4.0\,\text{m}\,\text{s}^{-1}$ for a heavier ball using the same launch mechanism? If it is, what does the equation predict?

3 A useful mathematical relationship used by physicists is $\vec{F}_{net} = m\vec{a}$, where $\vec{F}_{net}$ is the net force on an object (in N), m is the mass of an object (in kg) and $\vec{a}$ is the acceleration of the object (in $\text{m}\,\text{s}^{-2}$). Transpose this equation to make:

a $\vec{a}$ the subject

b m the subject.

4 A procedure was repeated five times. How should the following statement be rewritten?

We conducted many repeats of the procedure.

1.7 Communicating

The way you approach communicating your results will depend on the audience you want to reach. If you are communicating with a general audience, you may want to discuss your investigation in the style of a news article or blog post. These types of communication don't use too much scientific language as you need to assume that your audience does not have a science background.

Throughout this course you will need to present your research using appropriate scientific language and notation. There are many different presentation formats that are used, such as posters, oral presentations and reports. This section covers the main characteristics of effective science communication and report writing, including objectivity, clarity, conciseness and coherence.

PRESENTING YOUR WORK

Your teacher will specify how you should present your work. Two common formats are a scientific report and a scientific poster.

First identify your audience. This might be your teacher, your peers or the general public.

Second, consider how you could use figures such as labelled diagrams, graphs or flowcharts to aid communication; for example, to show how equipment was set up, or to summarise a process.

Scientific report

Scientific reports can be written to describe the findings of an experiment or a literature review.

A scientific report usually has the following components or sub-headings:

- Name (and lab-partner name)
- Date
- Title
- Introduction
  - summary of relevant background information
  - purpose
  - hypothesis
  - definition of independent, dependent and controlled variables
- Materials
  - equipment
- Procedure
  - step-by-step instructions that the reader can follow to replicate the experiment if required
  - can include two-dimensional scientific diagrams
- Results
  - summary data presented in tables and/or figures
  - all tables and figures numbered, with an appropriate title
- Discussion
  - brief explanation of the significance of the results
  - direct reference to data in tables and/or figures
  - links to relevant physics theory
  - limitations of the research and proposed avenues for further research
- Conclusion
  - summary of the research findings using evidence
  - relates to the aim and the hypothesis
- References
  - formatted in a consistent style; for example, Amercian Psychological Association (APA) referencing style.

Scientific poster

Scientists often present their research at conferences in the form of a scientific poster. This is a summary of their research that includes visual support such as tables, diagrams, graphs and flowcharts (Figure 1.7.1).

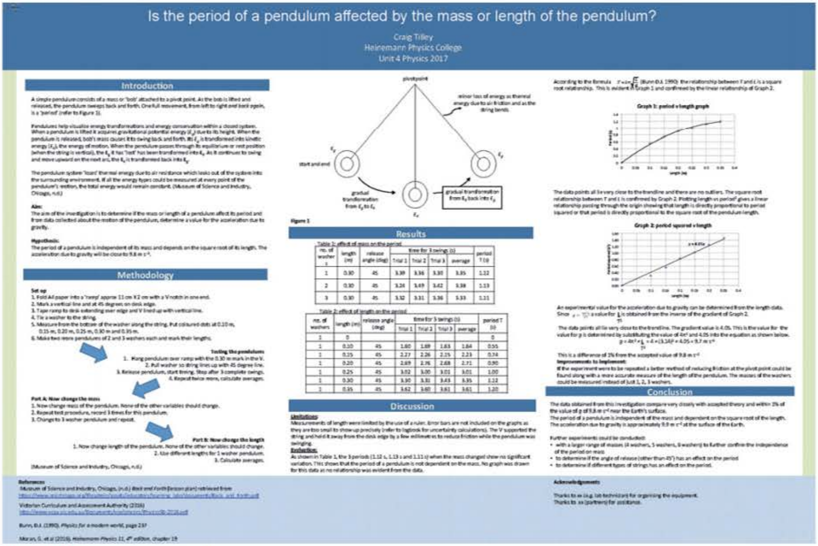


FIGURE 1.7.1 An example of a scientific poster.

Regardless of the style of your presentation, it is a good idea to help the reader navigate the work by numbering tables and figures and referring to these in the body of the report.

Historical or theoretical essay

An essay contains the following elements/characteristics:

- It has a formal structure, including introduction, body paragraphs and conclusion.
- The introduction states the focus of the essay.
- Each paragraph makes a new point supported by evidence.
- Each paragraph has a link back to the previous paragraph.
- The concluding paragraph draws all of the information together but does not introduce any new information.
- Any visuals are included at the end of the essay in an appendix.

Oral presentation

Consider the following elements when you prepare an oral presentation.

- Try to make your presentation engaging.
- Use cue cards but do not read directly from them.
- Look at the audience as you speak.
- Smile and appear confident (Figure 1.7.2).
- Try not to fidget.

FIGURE 1.7.2 When giving an oral presentation use your body language to engage the audience. Make eye contact, look around the audience and smile.

REVISION

Structuring a report

Your report should have a clear, logical structure.

Introduction

- The first paragraph should introduce your research topic and define key terms.

Body paragraphs

- Each paragraph should cover one main idea.
- The first sentence of each paragraph should summarise the content of the paragraph.
- Use evidence to support statements.
- Avoid very long or very short paragraphs.

Conclusion

- The final section should summarise the main findings.
- It should relate to the title of the investigation.
- It should include a discussion of the limitations of the investigation.
- It should discuss implications and applications of the research and potential future research.

GO TO ➤ Year 11 Section 1.7

Other types of presentation

There are many other ways to present your findings, such as:

- making a documentary or media report
- developing an evidence-based argument
- developing an environmental management plan
- analysing a work of fiction or film for scientific relevance
- creating a visual presentation.

Analysing relevant information

Scientific research should always be objective and neutral. Any premise presented must be supported with facts and evidence to allow the audience to make its own decision. Identify the evidence supporting or contradicting each point you want to make. It is also important to explain connections between ideas, concepts, theories and models. Figure 1.7.3 shows the questions you need to consider when writing your investigation report.

Once you have analysed your sources, annotate your outline, indicating where you will use evidence and what the source of that evidence is. Try to introduce only one idea per sentence and one theme per paragraph.

For example, for a report on 'Experimental research into biodegradability of plastics', the third paragraph might contain information from:

- Selke et al. (2015), who reported no significant degradation
- Chiellini et al. (2007), who reported a significant degradation.

FIGURE 1.7.3 Discuss relevant information, ideas, concepts and implications, and make sure your discussion is relevant to the question under investigation.

A report should include an analysis and synthesis of your sources. The information from different sources needs to be explicitly connected, and it should be clear where sources agree and disagree. In this example, the final sentences could be:

> Selke et al. (2015) reported that tests of plastic polymers treated with biodegradation additives resulted in no significant biodegradation after three years. This finding contrasts with that of Chiellini et al. (2007), who reported significant biodegradability of additive-treated polymers.
>
> The different results can be explained by differences in the studies. The 2007 study tested degradation in natural river water, whereas the 2015 study tested degradation under ultraviolet light, aerobic soil burial and anaerobic aqueous conditions (Chiellini et al. 2007; Selke et al. 2015). As well as using different additives and different experimental techniques, Selke et al. (2015) used additive rates of 1–5% and tested polyethylene terephthalate (PET) as well as polyethylene, whereas Chiellini et al. (2007) used additive rates of 10–15% and tested only polyethylene.
>
> Both studies were conducted under laboratory conditions, so they may not reflect what happens in the natural environment.

REVISION

GO TO ➤ Year 11 Section 1.7

Writing for science

Scientific reports are usually written in an objective or unbiased style. This is in contrast with English writing that most often uses the subjective techniques of **rhetoric** or **persuasion**.

Scientific writing can be written either in first-person or in third-person narrative. Your teacher may advise you on which to select. In either case, ensure that you keep the narrative point of view consistent.

Be careful of words that are absolute, such as *always*, *never*, *shall*, *will* and *proven*. Sometimes it may be more accurate and appropriate to use qualifying words, such as *may*, *might*, *possible*, *probably*, *likely*, *suggests*, *indicates*, *appears*, *tends*, *can* and *could*.

It is important to write concisely, particularly if you want to engage and maintain the interest of your audience. Use shorter sentences that are less verbose (i.e. do not contain too many words).

Identify concepts that can be explained using visual models and information that can be presented in graphs or diagrams. This will not only reduce the word count of your work but will also make it more accessible for your audience.

EDITING YOUR REPORT

Editing your report is an important part of the process. After editing your report, save new drafts with a different file name and always back up your files in another location.

Pretend you are reading your report for the first time when editing. Once you have completed a draft, it is a good idea to put it aside and return to it with 'fresh eyes' a day later. This will help you find areas that need further work and give you the opportunity to improve them. Look for content that is:

- ambiguous or unclear
- repetitive
- awkwardly phrased
- too lengthy
- not relevant to your research question
- poorly structured
- lacking evidence
- lacking a reference (if it is another researcher's work).

Use a spellchecker tool to help you identify typographical errors, but first check that your spellchecker is set to Australian English. Also be wary of words that are commonly misused, for example:

- where/were
- their/they're/there
- affect/effect
- its/it's
- which/that.

REVISION

GO TO ➤ Year 11 Section 1.7

References and acknowledgements

All the quotations, documents, publications and ideas used in the investigation need to be acknowledged in the 'references and acknowledgements' section in order to avoid plagiarism and to ensure authors are credited for their work. References and acknowledgements also give credibility to the study and allow the audience to locate information sources should they wish to study them further. The standard referencing style used is the American Psychological Association (APA) academic referencing style.

For example, a book would be referenced as: Rickard, G., et al. (2016), *Pearson Science 9 Student Book* (2nd ed.), Pearson Education, Melbourne, Australia. An example of a website reference is: Adamson, D. H. (2017). 'Plastics' in *World Book Advanced*, accessed 26 February 2018, from http://www.worldbookonline.com/advanced/article?id=ar434080.

In-text citations

Each time you write about the findings of other people or organisations, you need to provide an in-text citation and provide full details of the source in a reference list. In the APA style, in-text citations include the first author's last name and date in brackets (author, date). List the full details in your list of references.

The following examples show the use of in-text citation.

> It was proposed that the Los Alamos model of fission be modified to separate the neutron contribution for binary fission (Madland & Kahler 2017).

Or

> Madland & Kahler (2017) proposed a modification to the Los Alamos model to better align with current experimental results, including a modification to binary fission neutron contributions.

The bibliographic details of the example above would be:

> Madland, D.G. & Kahler, A.C. (2017). Refinements in the Los Alamos model of the prompt fission neutron spectrum. *Nuclear Physics A*, Vol. 957, pp. 289–311, ISSN 0375-9474.

There are many online guides to help you format your reference list and in-text citations. Some websites and journals suggest how to quote a source in a particular referencing style, such as APA.

PEER REVIEW

Your teacher may suggest that you show your work to a peer in order to seek constructive feedback. Part of the skill of a working scientist is to be open to receiving feedback, and to give constructive feedback to others.

Consider the following:

- Is the purpose clear?
- Is the format organised logically?
- Could you repeat the investigation without any additional information?
- Are tables and figures numbered?
- Is the reader guided when to refer to tables and figures from the body of the text?
- Are in-text citations used to acknowledge all sourced images and content?
- Are chemical conventions/nomenclature consistently followed?
- What are the strengths of the work?
- What questions do you have about the work?

MEASUREMENT AND UNITS

In every area of physics, scientists attempt to quantify the phenomena that were studied. In practical demonstrations and investigations, measurements are made, and those measurements are processed so that conclusions can be drawn. Scientists have a number of conventional ways of interpreting and analysing data from their investigations. There are also conventional ways of writing numerical measurements and their units.

REVISION

GO TO ➤ Year 11 Section 1.7

Prefixes and conversion factors

Conversion factors should be used carefully. You should be familiar with the prefixes and conversion factors in Table 1.7.1. The most common mistake made with conversion factors is multiplying rather than dividing. Note that the table gives all conversions as a multiplying factor.

TABLE 1.7.1 Prefixes and conversion factors.

Multiplying factor		Prefix	Symbol
1 000 000 000 000	10^{12}	tera	T
1 000 000 000	10^{9}	giga	G
1 000 000	10^{6}	mega	M
1 000	10^{3}	kilo	k
0.01	10^{-2}	centi	c
0.001	10^{-3}	milli	m
0.000 001	10^{-6}	micro	μ
0.000 000 001	10^{-9}	nano	n
0.000 000 000 001	10^{-12}	pico	p

Do not put spaces between prefixes and unit symbols; so, for example, for kilometres write km not k m.

It is important to give the symbol the correct case (upper or lower case). There is a big difference between 1 mm and 1 Mm.

REVISION

Correct use of unit symbols

The correct use of unit symbols removes ambiguity, as symbols are recognised internationally. The symbols for units are not abbreviations and should not be followed by a full stop unless they are at the end of a sentence.

The product of a number of units is shown by separating the symbol for each unit with a dot or a space. The division or ratio of two or more units can be shown in fraction form, using a slash (for example, m/s), or using negative indices (for example, m s^{-1}). Prefixes should not be separated by a space.

Numbers and symbols should not be mixed with words for units and numbers. For example, thirty grams and 30 g are correct, while 30 grams and thirty g are incorrect.

Scientific notation

To overcome confusion or ambiguity, measurements are often written in scientific notation. Quantities are written as a number between one and ten and then multiplied by an appropriate power of ten.

You should be routinely using scientific notation to express numbers.

GO TO ➤ Year 11 Section 1.7

PEER REVIEW

Your teacher may suggest that you partner with a student in your class, to provide each other with constructive feedback regarding the presentation of your investigation.

Consider the following:

- Are the findings accurately communicated using physics-specific language?
- Is the presentation written for the appropriate audience?
- Are tables, images, diagrams, graphs and flowcharts referenced in the text?
- Are appropriate in-text citations included for theoretical concepts and images not created by the student?
- Are all in-text citations and references listed in a consistent style, such as APA referencing style?
- What are the strengths of the presentation?
- What questions do you have about the presentation?

1.7 Review

SUMMARY

- A scientific report must include an introduction, body paragraphs and conclusion.
- The conclusion should include a summary of the main findings, a conclusion related to the issue being investigated, limitations of the research, implications and applications of the research, and potential future research.
- Scientific writing uses unbiased, objective, accurate, formal language. Scientific writing should also be concise and qualified.
- Visual support can assist in conveying scientific concepts and processes efficiently.
- Ensure you edit your final report.
- Scientific notation needs to be used when communicating your results.

KEY QUESTIONS

1 Which of the following statements is written in scientific style?

A The results were fantastic...
B The data in Table 3 indicates...
C The researchers felt...
D The smell was awful...

2 Which of the following statements is written in third-person narrative? (More than one response may be correct.)

A The researchers reported...
B Samples were analysed using...
C The experiment was repeated three times...
D I reported...

3 A student wishes to confirm the theoretical relationship for a pendulum, $T = 2\pi\sqrt{\frac{L}{g}}$, where T is the period and L is the length of the pendulum.

a Find the following:
  i the independent variable
  ii the dependent variable
  iii the variables that must be held constant.

b Draw an experimental set-up, and write an accompanying method that would allow this relationship to be tested.

4 **a** Write 255 000 in scientific notation.
b Write 0.000 000 432 in scientific notation.
c Explain why scientific notation is used.

Chapter review

01

KEY TERMS

accuracy
affiliation
bar graph
bias
continuous variable
controlled variable
credible
data analysis
dependent variable
discrete variable
expertise
hypothesis
independent variable
inquiry question
literature review
mean
median
mistake
mode
model
nominal variable
ordinal variable
outlier
percentage uncertainty
personal protective equipment (PPE)
persuasion
phenomenon
pie diagram
precision
primary source
purpose
qualitative
quantitative
random error
raw data
reliability
reputation
rhetoric
scatter plot
secondary source
significant figures
systematic error
trend
trend line
uncertainty
validity
variable

REVIEW QUESTIONS

1 Create an appropriate graph for the following data:

Voltage difference (V)	Current (mA)
2	65
4	121
6	200
8	268
10	330
12	401

2 Which one of the following would not support a strong conclusion to a report?

A The concluding paragraphs are relevant and provide supporting evidence.

B The concluding paragraphs are written in emotive language.

C The concluding paragraphs include a reference to limitations of the research.

D The concluding paragraphs include suggestions for further avenues of research.

3 Which of the following consists only of secondary sources of information?

A a periodic table, an article published in a science magazine, a science documentary, a practical report written by a Year 12 student

B an article published in a peer-reviewed science journal, an article published in a science magazine, a science documentary

C a periodic table, an article published in a science magazine, a science documentary, this Year 12 textbook

D a science article published in a newspaper, an article published in a science magazine, a science documentary, a practical report written by a Year 12 student

4 What is the correct way to cite in text the following source in APA style?

Selke, S., Auras, R., Nguyen, T. A., Aguirre, E. C., Cheruvathur, R., & Liu, Y. (2015). Evaluation of biodegradation—promoting additives for plastics. *Environmental Science & Technology, 49*(6), 3796–3777.

A However, Selke et al. (2015) did not find any significant difference in biodegradeability.

B However, Selke et al. did not find any significant difference in biodegradeability[1].

C However, Selke et al. did not find any significant difference in biodegradeability (Selke, S., Auras, R., Nguyen, T. A., Aguirre, E. C., Cheruvathur, R., & Liu, Y. (2015). Evaluation of biodegradation—promoting additives for plastics. *Environmental Science & Technology, 49*(6), 3769–3777).

D However, Selke et al. (2015) did not find any significant difference in biodegradeability (Evaluation of biodegredation-promoting additives for plastics. *Environmental Science & Technology*).

5 Explain the meaning of the terms *aim*, *hypothesis* and *variables* in an investigation.

6 Consider the following research question:

'Is the radial acceleration experienced on the Luna Park roller-coaster proportional to the radius of curvature of the track?'

Which of the following is the independent, dependent and controlled variables?

a radial acceleration **b** track cart

c radius of track curvature

7 Consider the following hypothesis:
'The use of higher density foam in hockey goalie shin pads reduces the force on the ball as it comes to a stop.'
Name the independent, dependent and controlled variables for an experiment with this hypothesis.

8 List the independent and dependent variables in the following investigations.
  a To determine the spring coefficient for springs with a variety of material thicknesses.
  b To compare the thermal expansion of different metals with a temperature increase of 50°C.
  c To investigate the effect of mass on the motion of a toboggan, and explain the variation in terms of friction.

9 Convert $2\,\mathrm{W\,mm^{-2}\,K^{-1}}$ into $\mathrm{W\,m^{-2}\,K^{-1}}$.

10 Convert 30.00 mL into L.

11 Define the following terms:
  a mean
  b mode
  c median.

12 The measurements of mass (in g) of a metal were 7.02, 6.47, 6.92, 7.21, 6.53, 6.53. What is the uncertainty of the average of these values?

13 Which of the statistical measurements of mean, mode and median is most affected by an outlier?

14 Identify whether the following are mistakes, systematic errors or random errors.
  a A student spills some water from the container while measuring the temperature increase.
  b The reported measurements are above and below the true value.
  c A 25.00 g mass measures 25.5 g on the scale.

15 A student is confused about rules for significant figures. Summarise the following for the student.
  a Give examples of values with two, three, four and five significant figures.
  b Summarise the rules for reporting a calculated value that involved multiplication or division.
  c Summarise the rules for reporting a calculated value that involved addition or subtraction.

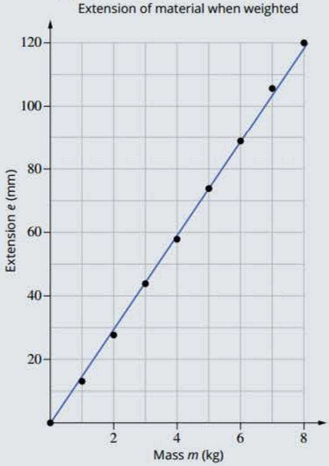
16 a Explain the terms 'accuracy' and 'precision'.
  b When might an investigation be invalid?
  c All investigations have limitations. Use an example to explain the meaning of 'limitations' of an investigation method.

17 A scientist designed and completed an experiment to test the following hypothesis: 'Increasing the temperature of the wire would result in a decrease in the resistance of the wire.'
The discussion section of the scientist's report included comments on the reliability, validity, accuracy and precision of the investigation.
Determine which of the following sentences comment on the reliability, validity, accuracy or precision.
  a Five different wires of the same length from the same source were tested at each temperature. The resistance was measured three times and averaged.
  b The temperature and resistance of the wire were recorded using data-logging equipment. The resistance of some of the wires was measured using an analog voltmeter and ammeter.
  c The data-logging equipment was calibrated for temperature before use. The equipment was calibrated before measurements were taken.
  d The voltage/current sensor (data logger) measured voltage to the nearest 0.1 V and the current to the nearest 0.1 mA. An analog voltmeter measured the voltage to the nearest 1 V, and the ammeter to the nearest 10 mA.

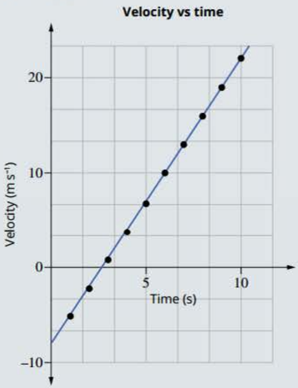
18 What is the purpose of referencing and acknowledging documents, ideas, images and quotations in your investigation?

19 A scientist designed and completed an experiment to test the following hypothesis:
'Increasing the temperature of a wire would result in a decrease in resistance.'
  a Write a possible aim for this scientist's experiment.
  b What would be the independent, dependent and controlled variables in this investigation?
  c What kind of data would be collected? Would it be qualitative or quantitative?
  d List the equipment that could be used and the type of precision expected for each item.
  e What would you expect the graph of the results to look like if the scientist's hypothesis was correct?

20 Determine the experimental equation from the following graph.

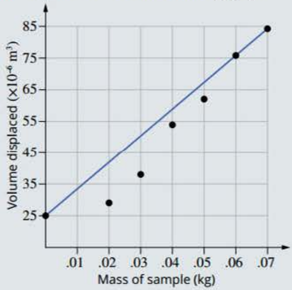


21 a Determine the experimental equation from the following graph.



b Interpret the gradient and intercept in a physics context.

22 Identify the errors in the following graph.



23 a Draw a graph for the following data obtained from a spring-extension practical.

Force (N)	Length (mm)
0	250
10	305
20	375
30	435
40	500
50	560

b Find the experimental equation, and write a sentence interpreting the gradient and *y*-intercept.

24 A student wishes to design an experiment to determine the time it takes for a cup of coffee to cool to room temperature from different starting temperatures.

a What are the independent, dependent and controlled variables?

b What is the inquiry question and the hypothesis?

c Design a method that could be followed by the student to get repeatable results.

d Create a table that could be used to collect data from the experiment.

25 Compare the crash-test data for three different cars using the Australasian New Car Assessment Program (ANCAP) ratings.

a Which car had the highest minimum frontal offset score?

b Which car had the highest side impact score?

c Which car had the highest total score?

d Present your findings in a table.

26 Research the use of international collaboration in particle accelerators, and the major advances that have been made possible. Ensure that you include a list of references and acknowledgements for your sources. Present your research in digital form.

MODULE 5 Advanced mechanics

Motion in one dimension at constant velocity or constant acceleration can be explained and analysed relatively simply. However, motion is frequently more complicated because the net force on objects can vary in size and direction, causing objects to move in two or three dimensions.

In this module, you will develop an understanding that all forms of complex motion can be understood by analysing the forces acting on a system, including the energy transformations taking place within and around the system. By applying new mathematical techniques, you will model and predict the motion of objects within systems. You will examine two-dimensional motion, including projectile motion and uniform circular motion, along with the orbital motion of planets and satellites, which are modelled as an approximation to uniform circular motion.

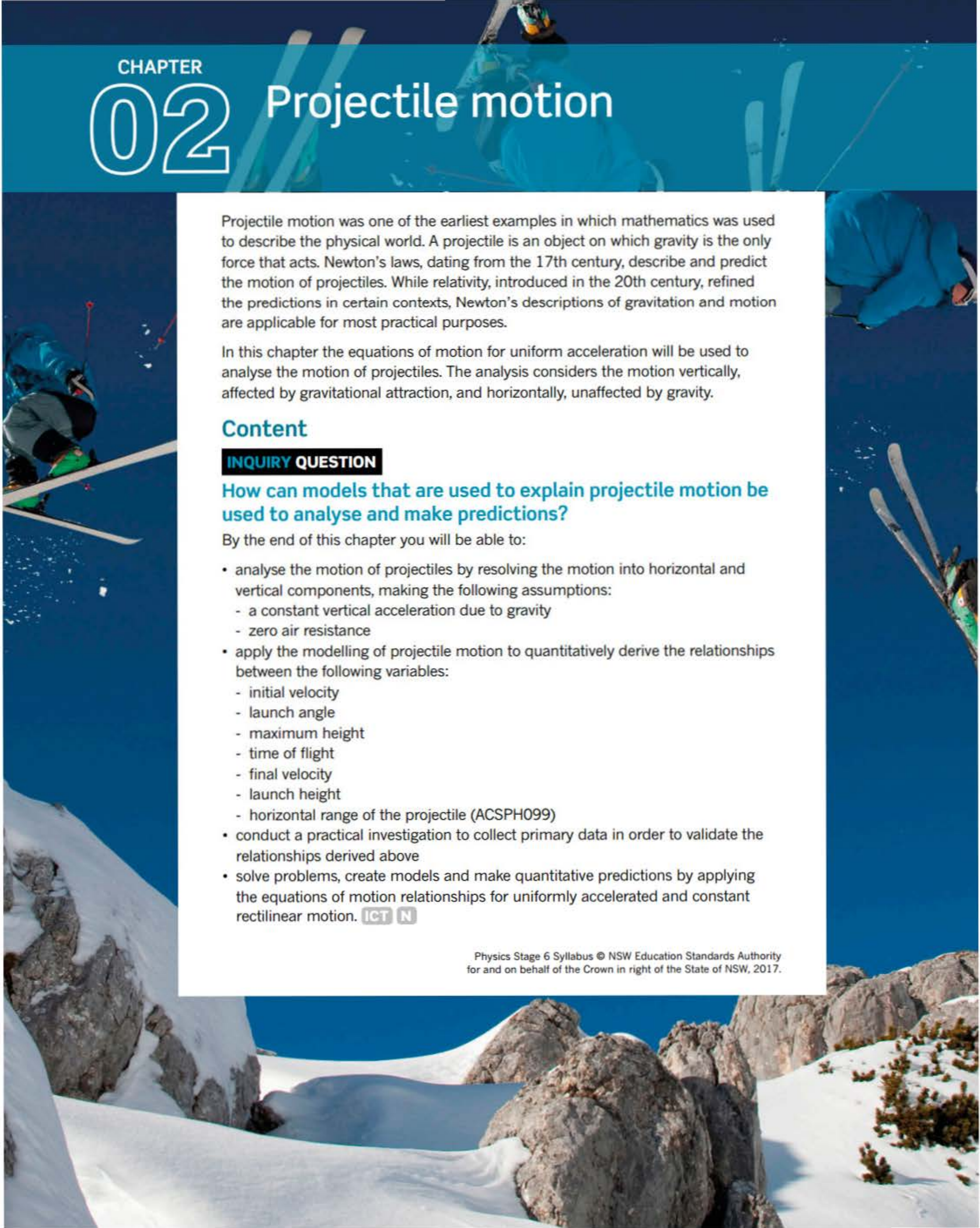
Outcomes

By the end of this module you will be able to:

- select and process appropriate qualitative and quantitative data and information using a range of appropriate media PH12-4
- analyse and evaluate primary and secondary data and information PH12-5
- solve scientific problems using primary and secondary data, critical thinking skills and scientific processes PH12-6
- communicate scientific understanding using suitable language and terminology for a specific audience or purpose PH12-7
- describe and analyse qualitatively and quantitatively circular motion and motion in a gravitational field; in particular, the projectile motion of particles PH12-12

Throughout this book vectors will be represented either in italics, for example *F*, or with an arrow above the variable, for example $\vec{F}$. Regardless of the way a vector is represented, it is important for you to understand that the same rules apply when completing any calculation involving vectors. A vector is a variable which has both a magnitude and a direction. Being able to understand when a variable is described by a vector is an important skill for a physicist to develop. See the revision box on page 50 for further detail regarding the treatment of vectors in the study of motion.

It is recommended that you check the online formulae sheet for the most up-to-date notation used for the New South Wales Stage 6 Physics course.



CHAPTER 02

Projectile motion

Projectile motion was one of the earliest examples in which mathematics was used to describe the physical world. A projectile is an object on which gravity is the only force that acts. Newton's laws, dating from the 17th century, describe and predict the motion of projectiles. While relativity, introduced in the 20th century, refined the predictions in certain contexts, Newton's descriptions of gravitation and motion are applicable for most practical purposes.

In this chapter the equations of motion for uniform acceleration will be used to analyse the motion of projectiles. The analysis considers the motion vertically, affected by gravitational attraction, and horizontally, unaffected by gravity.

Content

INQUIRY QUESTION

How can models that are used to explain projectile motion be used to analyse and make predictions?

By the end of this chapter you will be able to:

- analyse the motion of projectiles by resolving the motion into horizontal and vertical components, making the following assumptions:
  - a constant vertical acceleration due to gravity
  - zero air resistance
- apply the modelling of projectile motion to quantitatively derive the relationships between the following variables:
  - initial velocity
  - launch angle
  - maximum height
  - time of flight
  - final velocity
  - launch height
  - horizontal range of the projectile (ACSPH099)
- conduct a practical investigation to collect primary data in order to validate the relationships derived above
- solve problems, create models and make quantitative predictions by applying the equations of motion relationships for uniformly accelerated and constant rectilinear motion. ICT N

2.1 Projectiles launched horizontally

REVISION

Kinematics

A scalar quantity requires a **magnitude** and a unit—3 m, 5 s, $9.80\,m\,s^{-2}$.

A vector quantity is a scalar quantity with a direction. The magnitude (or size) of a vector describes the scalar equivalent. For example, the magnitude of an object's velocity will be equal to its speed.

Vectors can be represented using a variety of different notations. Throughout this book you will see some vectors written with an arrow above the variable, i.e. $\vec{F}$ or $\vec{p}$. This is known as **vector notation**. There are different sorts of vector notation, for example the same vector can be represented as $\tilde{s}$, **s** or *s*. Table 2.1.1 on page 51 outlines the different scalars and vectors related to kinematics and dynamics that appear throughout this book and the symbol used to represent each variable.

If a variable is a vector then you need to apply a direction convention. For example, a swimmer is completing laps of a pool. They first swim 25 m east before turning back and swimming 10 m west. To calculate their displacement from their initial position, using the direction conventions outlined in Figure 2.1.1 on page 51, the displacement will be:

$$s = s_1 + s_2 = 25 + -10$$
$$= +15$$
$$= 15 \text{ m east}$$

The equations of motion can be used in situations where there is a constant acceleration *a* (in $m\,s^{-2}$). These allow you to model the motion of objects and predict values for the initial velocity *u* (in $m\,s^{-1}$), final velocity *v* (in $m\,s^{-1}$), displacement *s* (in m) and time *t* (in s). A direction convention is always required when using the equations of motion. The equations of motion for uniform acceleration are:

$$v = u + at$$
$$s = \tfrac{1}{2}(u+v)t$$
$$s = ut + \tfrac{1}{2}at^2$$
$$v^2 = u^2 + 2as$$

WS 5.1

A **projectile** is any object that is thrown or projected into the air and is moving freely—that is, it has no power source (such as a rocket engine or propeller) driving it. A netball as it is passed, a cricket ball that is hit for six and an aerial skier flying through the air are all examples of projectiles. This type of flight is known as **ballistic** flight. People have long argued about the path that projectiles follow; some thought they were circular or had straight sections. It is now known that projectiles move in smooth parabolic paths, like that shown in Figure 2.1.2 on page 51. This section analyses projectile motion using Newton's laws and the equations of motion.

A projectile:

- is macroscopic (that is, not microscopic)
- does not travel too fast (that is, not at a substantial fraction of the speed of light).

Newton's treatment of projectile motion is fine for balls, rockets and even planets.

Particles smaller than about 1 mm diameter, or moving so fast that relativity plays an effect, need more sophisticated theories.

FREE-FALLING PROJECTILES

It is a very common misconception that when a projectile travels forwards through the air, it has a forward force acting on it, but this is incorrect. There might have been some forward force acting as the projectile is launched, but once the projectile is released, this forward force is no longer acting. This causes the projectile to continually deviate from a straight-line path to follow a parabolic path.

The only force acting on a projectile during its flight is its **weight**, which is the force due to gravity, $F_g = mg$. This force is constant and always directed vertically downwards. This motion is known as **free fall**. Given that the only force acting on a projectile is the force due to gravity, it follows that a projectile near the surface of the Earth must have a vertical acceleration of $9.8\,m\,s^{-2}$ downwards throughout its motion. The mass of the projectile does not affect its acceleration due to gravity.

Worked example 2.1.1

PROJECTILE UNDER FREE FALL

A ball is allowed to free fall from a height of 40 m. Take the acceleration due to gravity to be $g = 9.8\,\text{m}\,\text{s}^{-2}$ and ignore air resistance.

a Calculate the time that the ball takes to land.

Thinking	Working
Let the downward direction be positive. Write out the information relevant to the vertical component of the motion. Note that at the instant the ball is released it is not moving, so its initial vertical velocity is zero.	Down is positive. Vertically: $s = +40\,\text{m}$ $u = 0\,\text{m}\,\text{s}^{-1}$ $a = +9.8\,\text{m}\,\text{s}^{-2}$ $t = ?$
In the vertical direction, the ball has constant acceleration, so use equations for uniform acceleration. Select the equation that best fits the information you have.	$s = ut + \frac{1}{2}at^2$
Substitute values, rearrange and solve for t.	$40 = 0 + \frac{1}{2} \times 9.8 \times t^2$ $t^2 = \frac{40}{4.9}$ $t = \sqrt{\frac{40}{4.9}}$ $= 2.857\,\text{s}$ $= 2.9\,\text{s}$ (to two significant figures)

b Calculate the velocity of the ball at the point of impact.

Thinking	Working
Let the downward direction be positive. Write out the information relevant to the vertical component of the motion.	Vertically, with down as positive: $s = +40\,\text{m}$ $u = 0\,\text{m}\,\text{s}^{-1}$ $v = ?$ $a = +9.8\,\text{m}\,\text{s}^{-2}$ $t = 2.9\,\text{s}$
To find the final vertical speed, use the equation for uniform acceleration that best fits the information you have.	$v = u + at$
Substitute values, rearrange and solve for v.	Vertically: $v = 0 + 9.8 \times 2.9$ $= 28\,\text{m}\,\text{s}^{-1}$ down
Give the velocity with a magnitude and a direction.	The final velocity of the ball is $28\,\text{m}\,\text{s}^{-1}$ down.

Worked example: Try yourself 2.1.1

PROJECTILE UNDER FREE FALL

A ball is allowed to free fall from a height of 30 m on Mars. Take the acceleration due to gravity to be $g = 3.8\,\text{m}\,\text{s}^{-2}$ and ignore air resistance.

a Calculate the time that the ball takes to land.

b Calculate the velocity of the ball at the point of impact.

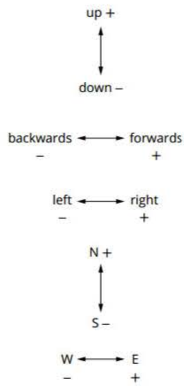


FIGURE 2.1.1 Common sign and direction conventions

TABLE 2.1.1 The scalars and vectors used in kinematics and dynamics

Scalar	Vector
distance travelled d	displacement s
speed v	velocity v
acceleration a	acceleration a
time t	

FIGURE 2.1.2 A multi-flash photograph of a golf ball bouncing on a hard surface and moving to the right. The parabolic shape of the path is clearly seen. Each peak is a little lower as some energy is lost after each bounce.

Projectiles can be launched at any angle. The launch velocity needs to be resolved into vertical and horizontal components using trigonometry in order to complete most problems. For projectiles launched horizontally, the horizontal velocity is constant and is equal to the horizontal launch velocity. Assuming there is no air resistance, the only force that is acting on a projectile is gravity. Gravity acts as a vertical acceleration so it has no effect on the horizontal motion of a projectile.

Tips for solving projectile motion problems

In solving a projectile motion problem, construct a diagram showing the projectile's motion to set the problem out clearly. Write out the information supplied for the horizontal and vertical components separately.

- In the horizontal direction, the velocity v_H of the projectile is constant, so the following formula can be applied: $v_{av} = \frac{s}{t}$.
- In the vertical direction, the projectile is affected by a constant acceleration due to gravity, $g = 9.80\,m\,s^{-2}$ downwards, so the equations of motion for uniform acceleration are used. It is important to clearly specify whether up or down is the positive or negative direction. Either choice will work just as effectively, as long as the same convention is used consistently throughout each problem.
- The speed of the projectile at any point can be calculated using Pythagoras' theorem by considering the horizontal and vertical components. If the velocity of the projectile is required, it is necessary to provide a direction with respect to the horizontal plane as well as the speed of the projectile.

SKILLBUILDER N

Understanding vector components

When vectors lie in one dimension, it is relatively simple to understand their motion. However, in projectile motion, the object moves in a two-dimensional plane, so it is important to understand how to break vectors down into their components.

These types of vector problems in physics (and other applications) often involve using trigonometric relationships to find the unknown side of a right-angled triangle.

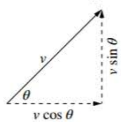
You will recall that:

$$\sin\theta = \frac{\text{opposite}}{\text{hypotenuse}} \quad \cos\theta = \frac{\text{adjacent}}{\text{hypotenuse}} \quad \tan\theta = \frac{\text{opposite}}{\text{adjacent}}$$

In the triangle shown below, the vector v represents the velocity of a projectile launched at some angle θ from the ground. The velocity can be broken down into its horizontal (v_H) and vertical (v_V) components using the trigonometric relationships outlined above.

$$\sin\theta = \frac{\text{opposite}}{\text{hypotenuse}} = \frac{v_V}{v} \qquad \cos\theta = \frac{\text{adjacent}}{\text{hypotenuse}} = \frac{v_H}{v}$$

$$v_V = v\sin\theta \qquad v_H = v\cos\theta$$

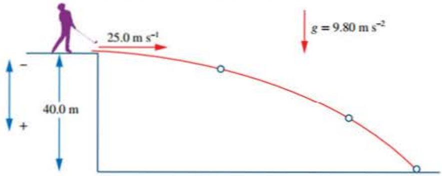


Breaking a vector into its horizontal and vertical (orthogonal) components is a commonly used tool in physics. As the components are perpendicular they form a right-angled triangle. The original vector can then be found from the horizontal and vertical components using Pythagoras' theorem, so that $v^2 = v_H{}^2 + v_V{}^2$.

Worked example 2.1.2

PROJECTILE LAUNCHED HORIZONTALLY

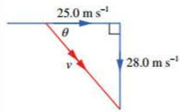
A ball is hit horizontally from the top of a 40.0 m high cliff with a velocity of 25.0 m s^{-1} to the right. Using an acceleration due to gravity of $g = 9.80$ m s^{-2} and ignoring air resistance, calculate the following values:



a the time that the ball takes to land	
Thinking	**Working**
Let the downward direction be positive. Write out the information relevant to the vertical component of the motion. Note that the instant the ball is hit, it is travelling only horizontally, so its initial vertical velocity is zero.	Down is positive. Vertically: $s = +40.0$ m $u = 0$ m s^{-1} $a = +9.80$ m s^{-2} $t = ?$
In the vertical direction, the ball has constant acceleration, so use equations for uniform acceleration. Select the equation that best fits the information you have.	$s = ut + \frac{1}{2}at^2$
Substitute values, rearrange and solve for t.	$40.0 = 0 + \frac{1}{2} \times 9.80 \times t^2$ $t = \sqrt{\frac{40.0}{4.9}}$ $= 2.86$ s (to three significant figures)

b the distance that the ball travels from the base of the cliff; that is, the total horizontal distance travelled	
Thinking	**Working**
Write out the information relevant to the horizontal component of the motion. As the ball is hit horizontally, the initial speed gives the horizontal component of the velocity throughout the flight. Never round off a value until the final calculation, otherwise you will introduce a rounding error. Use the calculated value of t.	Horizontally: $u = 25.0$ m s^{-1} $t = 2.86$ s from part (a) $s = ?$
Select the equation that best fits the information you have.	As horizontal speed is constant, you can use $v_{av} = \frac{s}{t}$
The distance travelled will be equal to the magnitude of the displacement. Substitute values, rearrange and solve for s. The distance travelled is a scalar quantity so no direction is required.	$25.0 = \frac{s}{2.86}$ $s = 25.0 \times 2.86$ $= 71.4$ m

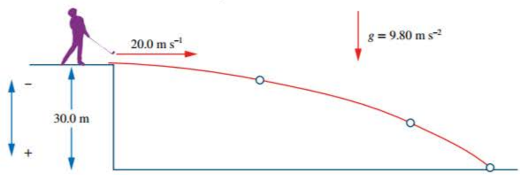
c the velocity of the ball as it lands.

Thinking	Working
Find the horizontal and vertical components of the ball's speed as it lands. Write out the information relevant to both the vertical and horizontal components.	Horizontally: $u = v_H = 25.0\,\text{m s}^{-1}$ Vertically, with down as positive: $u = 0\,\text{m s}^{-1}$ $a = +9.80\,\text{m s}^{-2}$ $s = +40.0\,\text{m}$ $t = 2.86\,\text{s}$ $v = ?$
To find the final vertical speed, use the equation for uniform acceleration that best fits the information you have.	$v = u + at$
Substitute values, rearrange and solve for the variable you are looking for, in this case v.	Vertically: $v_V = u + at$ $= 0 + 9.80 \times 2.86$ $= 28.0\,\text{m s}^{-1}$ down
Add the components as vectors.	
Use Pythagoras' theorem to work out the actual speed, v, of the ball.	$v = \sqrt{{v_H}^2 + {v_V}^2}$ $= \sqrt{25.0^2 + 28.0^2}$ $= \sqrt{1409}$ $= 37.5\,\text{m s}^{-1}$
Use trigonometry to solve for the angle, θ.	$\theta = \tan^{-1}\left(\frac{28.0}{25.0}\right)$ $= 48.2°$
Indicate the velocity with a magnitude and direction relative to the horizontal.	The final velocity of the ball is $37.5\,\text{m s}^{-1}$ at 48.2° below the horizontal.

Worked example: Try yourself 2.1.2

PROJECTILE LAUNCHED HORIZONTALLY

A ball of mass 100 g is hit horizontally from the top of a 30.0 m high cliff with a speed of $20.0\,\text{m s}^{-1}$. Using acceleration due to gravity of $g = 9.80\,\text{m s}^{-2}$ and ignoring air resistance, calculate the following values:



a the time that the ball takes to land

b the distance that the ball travels from the base of the cliff, i.e. the range of the ball

c the velocity of the ball as it lands.

PA 5.1

PHYSICS IN ACTION ICT

The effect of drag

The interaction between a projectile and the air can have a significant effect on the motion of the projectile (Figure 2.1.3), particularly if it has a large surface area and relatively low mass. If you try to throw an inflated party balloon, it will not travel very far compared to throwing a marble at the same initial speed.

The size of the **air resistance** (or drag force) that acts on an object as it moves depends on several factors.

- The speed, *v*, of the object. The faster an object moves, the greater the drag force becomes.
- The cross-sectional area of the object in its direction of motion. Greater area means greater drag.
- The aerodynamic shape of the object. A more streamlined shape experiences less drag.
- The density of the air. Higher air density means greater drag.

Drag forces are proportional to the square of the velocity, and there will be a characteristic vertical velocity, called the **terminal velocity**, in which the drag forces balance the gravitational forces. With no net force acting on the projectile, it no longer accelerates. The projectile has reached its maximum rate of descent.

If there is less air, there is less drag due to air resistance and the body may free fall at a much faster rate. In October 2012 Felix Baumgartner undertook a free fall from a weather balloon located 39 km above the surface of the Earth (nearly at the top of the stratosphere and over four times higher than Mt Everest) and reached a top speed of over 1300 km h^{-1} during the fall (Figure 2.1.4). As he descended closer to the surface of the Earth, the atmosphere thickened, drag forces increased, and his velocity slowed. A parachute for the last kilometre of fall meant he survived the stunt.

FIGURE 2.1.4 Felix Baumgartner begins his 39 km fall.

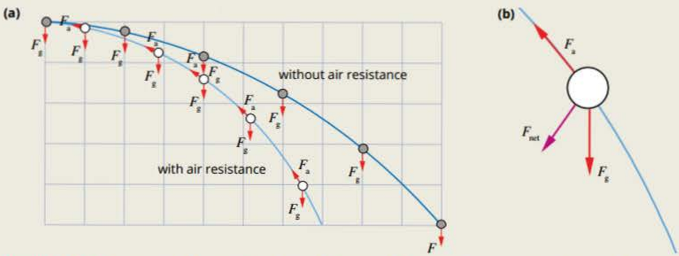


FIGURE 2.1.3 (a) The two paths of a projectile with and without air resistance (F_a). When air resistance is included, the horizontal distance travelled is shorter. (b) When air resistance is acting, the net force acting on the object is no longer vertically down.

2.1 Review

SUMMARY

- If air resistance is ignored, the only force acting on a projectile is its weight, i.e. the force of gravity, F_g.
- As a projectile is acting only under the force of gravity, the equations of motion for uniform acceleration can be used:

$$v = u + at$$
$$s = \tfrac{1}{2}(u+v)t$$
$$s = ut + \tfrac{1}{2}at^2$$
$$v^2 = u^2 + 2as$$

- Projectiles move in parabolic paths that can be analysed by considering the horizontal and vertical components of the motion.
- The horizontal velocity of a projectile remains constant throughout its flight if air resistance is ignored. Therefore, the following equation for average velocity can be used for this component of the motion:

$$v_{av} = \tfrac{s}{t}$$

KEY QUESTIONS

For the following questions use an acceleration due to gravity of $g = 9.8\,\text{m}\,\text{s}^{-2}$ and ignore air resistance.

1 A 300 g projectile is dropped from 200 m. How long does it take to touch the ground?

2 A projectile is launched straight upwards at $5.0\,\text{m}\,\text{s}^{-1}$ from 50 m above the ground. What is its velocity as it hits the ground?

3 A projectile is thrown upwards at $20\,\text{m}\,\text{s}^{-1}$. What is its maximum height?

4 In 1971, astronaut Alan Shepard took a golf club to the Moon and hit a couple of balls. Which one or more of the following is correct?

A The balls travelled in straight lines because there is no gravity.

B The balls travelled in parabolic arcs.

C The balls travelled much further than if they had been hit in an identical manner on Earth.

D The balls went into orbit.

5 A golfer practising on a range with an elevated tee 4.9 m above the ground strikes a ball so that it leaves the club with a horizontal velocity of $20\,\text{m}\,\text{s}^{-1}$.

a How long after the ball leaves the club will it land on the ground?

b What horizontal distance will the ball travel before landing on the ground?

c What is the acceleration of the ball 0.50 s after being hit with the golf club?

d Calculate the speed of the ball 0.80 s after it leaves the golf club.

e With what speed will the ball hit the ground?

2.2 Projectiles launched obliquely

PHYSICS INQUIRY CCT N

Projectile variables

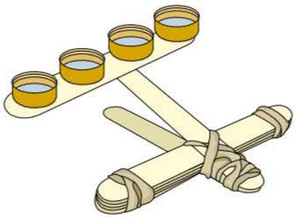
How can models that are used to explain projectile motion be used to analyse motion and make predictions?

COLLECT THIS...

- four different projectiles such as marshmallows, marbles, balls of aluminium foil, cubes of metal
- popsicle sticks, large
- elastic bands
- bottle container lid or equivalent
- glue
- wedge
- ruler

DO THIS...

1 Using the elastic bands, popsicle sticks, lids and glue to create a catapult similar to that shown below.



2 Record the properties of the different projectiles you have chosen: mass (quantitative), estimated air resistance based on shape and surface (qualitative).

3 Place projectiles in the lids on the catapult. Pull the central stick down to a known position and let go. Record the distance the different projectiles travel. If you have chosen objects which can bounce or roll, how can you increase the precision of this measurement?

4 Repeat, changing one variable at a time:
  - launch height—place the catapult on a stack of books to change the height
  - launch angle—place a wedge under the catapult to change the angle but keep the launch height the same
  - launch energy—pull the central stick down to a different position to change the potential energy stored in the catapult.

RECORD THIS...

Describe how the different variables affected the motion of the different projectiles.

Present a table of your results.

REFLECT ON THIS...

How can models that are used to explain projectile motion be used to analyse and make predictions?

Where have you seen projectile motion before?

The previous section looked at projectiles that were launched horizontally. A more common situation is projectiles that are launched obliquely (at an angle), by being thrown forwards and upwards at the same time. An example of an oblique launch is shooting for a goal in basketball, such as in Figure 2.2.1. Once the ball is released, the only forces acting are gravity pulling it down to Earth and air resistance, a drag force which slows the ball's motion slightly.

FIGURE 2.2.1 The basketball was thrown upwards towards the ring, travelling in a smooth parabolic path.

PROJECTILES LAUNCHED AT AN ANGLE

As with projectiles launched horizontally, if air resistance is ignored, the only force that is acting on a projectile is gravity. Gravity acts vertically downwards, so it has no effect on the horizontal motion of a projectile. The vertical and horizontal components of the motion are independent of each other and once again must be treated separately.

In the vertical direction, a projectile accelerates due to the force of gravity. The effect of the force due to gravity is that the vertical component of the projectile's velocity decreases as the projectile rises. It is momentarily zero at the very top of the flight, and then it increases again as the projectile descends.

In the horizontal direction, a projectile has a uniform velocity since there are no forces acting in this direction (if air resistance is ignored).

Tips for solving problems involving projectiles launched at an angle

GO TO ➤ Section 2.1 page 52

GO TO ➤ SkillBuilder page 52

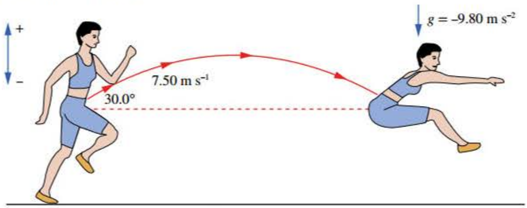
General rules for solving problems involving projectile motion were given in the previous section—see page 52 for a reminder.

For a projectile launched at an angle to the horizontal, trigonometry can be used to find the initial horizontal and vertical components of its velocity. Pythagoras' theorem can be used to determine the actual velocity of the projectile at any point, as well as its direction with respect to the horizontal plane.

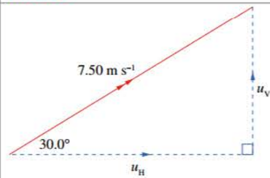
Worked example 2.2.1

LAUNCH AT AN ANGLE

An athlete in a long-jump event leaps with a velocity of $7.50\,\text{m s}^{-1}$ at an angle of 30.0° to the horizontal.



For the following questions, treat the athlete as a point mass, ignore air resistance and use $g = 9.80\,\text{m s}^{-2}$.

a What is the athlete's velocity at the highest point?	
Thinking	**Working**
First find the horizontal and vertical components of the initial speed. Remember, speed is a scalar quantity equal to the magnitude of the velocity.	 Using trigonometry: $u_H = 7.50 \cos 30$ $= 6.50\,\text{m s}^{-1}$ $u_V = 7.50 \sin 30$ $= 3.75\,\text{m s}^{-1}$

Projectiles that are launched at an angle move only horizontally at the highest point of their motion. The vertical component of the velocity at this point is therefore zero. The actual velocity is given by the horizontal component of the velocity at the highest point of their motion.	At maximum height: $v = 6.50\,\text{m s}^{-1}$ horizontally to the right.

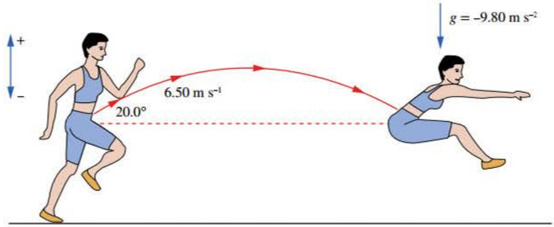
b What is the maximum height gained by the athlete during the jump?	
Thinking	**Working**
To find the maximum height gained by the athlete, you must use the vertical component. Recall that the vertical component of velocity at the highest point is zero.	Vertically, taking up as positive: $u = +3.75\,\text{m s}^{-1}$ $a = -9.80\,\text{m s}^{-2}$ $v = 0$ $s = ?$
Substitute these values into an appropriate equation for uniform acceleration.	$v^2 = u^2 + 2as$ $0 = (3.75)^2 + 2 \times (-9.80) \times s$
Rearrange and solve for s. Distance travelled is a scalar quantity.	$s = \frac{3.75^2}{2 \times 9.80}$ $= 71.7\,\text{cm}$

c Assuming the athlete returns to their original height, what is the total time they are in the air?	
Thinking	**Working**
As the motion is symmetrical, the time required to complete the motion will be double the time taken to reach the maximum height. First, the time it takes to reach the highest point must be found.	Vertically, taking up as positive: $u = +3.75\,\text{m s}^{-1}$ $a = -9.80\,\text{m s}^{-2}$ $v = 0$ $t = ?$
Insert these values into an appropriate equation for uniform acceleration.	$v = u + at$ $0 = 3.75 + (-9.80)t$
Rearrange the formula and solve for t.	$t = \frac{3.75}{9.80}$ $= 0.383\,\text{s}$
The time to complete the motion is double the time it takes to reach the maximum height.	Total time $= 2 \times 0.383$ $= 0.765\,\text{s}$

Worked example: Try yourself 2.2.1

LAUNCH AT AN ANGLE

An athlete in a long-jump event leaps with a velocity of $6.50\,m\,s^{-1}$ at 20.0° to the horizontal.



For the following questions, treat the athlete as a point mass, ignore air resistance and use $g = 9.80\,m\,s^{-2}$.

a What is the athlete's velocity at the highest point?

b What is the maximum height gained by the athlete during the jump?

c Assuming a return to the original height, what is the total time the athlete is in the air?

INITIAL CONDITIONS

When creating a model, the initial conditions are important. The initial speed and the launch angle both play a part in how high and how far a projectile will travel. In Figure 2.2.2a, the speed of a projectile is kept the same, but the launch angle is changed, while in Figure 2.2.2b, the launch angle is kept constant but the initial velocity is changed.

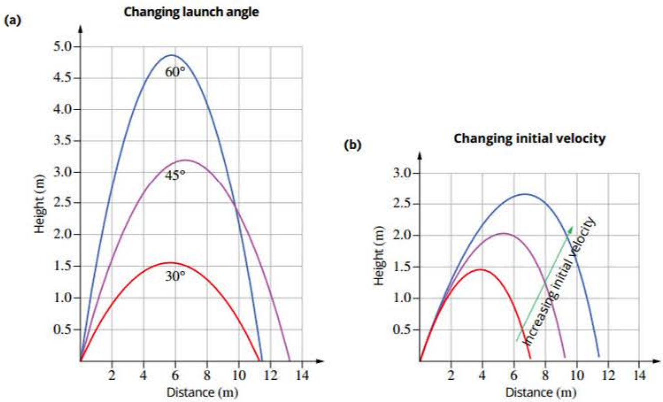


FIGURE 2.2.2 The initial conditions dictate the motion of the projectile. (a) The launch angle is increased while keeping the initial velocity constant. (b) The initial velocity is increased while keeping the launch angle constant.

+ ADDITIONAL

Analysing the effects of drag

Ship, aircraft and car designers look to minimise the effects of drag to allow their vehicles to travel as quickly and as economically as possible. These experts use computational fluid dynamics to make their calculations, but it is possible to make quite reasonable drag calculations for projectile motion with a simple spreadsheet calculator.

When ignoring drag, the only force acting on a projectile with mass m is that of gravity. The horizontal (H) and vertical (V) components of the projectile's acceleration are $a_H = 0$ and $a_V = -g = -9.8\,\text{m}\,\text{s}^{-2}$ (where upwards is positive).

The drag force, F_D, is approximately proportional to the square of the velocity, v, and D is a constant of proportionality related to the projectile's shape and the medium through which it travels, so $F_D = Dv^2$. The direction of the force opposes the direction of velocity, so $F_H = -Dv_H^2$ and $F_V = -Dv_V^2$.

Using Newton's second law to transform a force into an acceleration, the components of the acceleration (including both gravity and drag) are:

$$a_H = -\frac{D}{m}v_H^2$$

$$a_V = -g - \frac{D}{m}v_V^2$$

Clearly, the acceleration for both a_H and a_V change as the velocity changes, which means the simpler equations of motion can't be used. However, a numerical method can be used.

A numerical method repeats the velocity calculations after a short period—say every 0.1 s of the flight. At 0.1 s after the launch, the horizontal and vertical velocities can be recalculated. This gives a new position, a new acceleration and new velocity. Another 0.1 s later the calculation is repeated, and so the entire flight can be simulated.

For the example in Figure 2.2.3, a 900 g projectile was given an initial velocity of $10\,\text{m}\,\text{s}^{-1}$ and launched at 35° above the x-axis, with each point representing 0.016 s of time. The figure shows that the maximum height and total distance travelled were substantially reduced with drag included. In real-world examples, for example in sports or vehicle design, a calculation without drag will give unrealistic results.

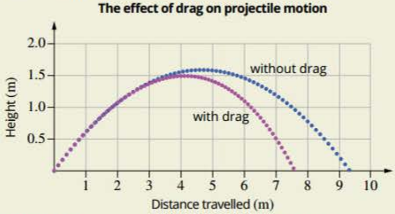


FIGURE 2.2.3 The calculated drag effect for a flight. Each dot represents a new numerical calculation. From here, it is simple to adjust the launch angle and launch velocity to explore the effects of drag, or to change the drag parameters to represent the shapes of different projectiles.

PHYSICSFILE ICT

Cannons

Cannons are useful, but difficult to use. Measuring angles and ranges are not precise, and the calculations can be slow. The force imparted on the projectile from the packed powder is imprecise and complicates making accurate repeated shots.

During World War I, the Western (French) and Italian fronts were the domain of artillery. In mid-1915 the Australian-born scientist William Bragg (Figure 2.2.4a) moved from his crystallography work to develop an accurate measurement of projectile ranging using acoustic locators. Acoustic locators (Figure 2.2.4b) work by using moveable microphones to find the angle that maximises the amplitude of sound received, which is also the bearing angle to the target. Two acoustic locators at different positions will generate two different bearings, which allows the use of triangulation to determine accurately where a projectile landed—or where the enemy launched one.

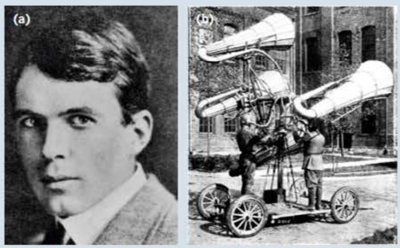


FIGURE 2.2.4 (a) William Lawrence Bragg, winner of the Nobel Prize for Physics 1915 and developer of sound-ranging methods. (b) An acoustic locator. The large horns amplified distant sounds, monitored through headphones worn by a crew member.

2.2 Review

SUMMARY

- For objects launched at an angle to the horizontal, the initial horizontal and vertical velocities can be calculated using trigonometry.
- At its highest point, the projectile is moving horizontally. Its velocity at this point is given by the horizontal component of its launch velocity. The vertical component of the velocity is zero at this point.

KEY QUESTIONS

For the following questions use an acceleration due to gravity of $g = 9.8\,m\,s^{-2}$ and ignore air resistance.

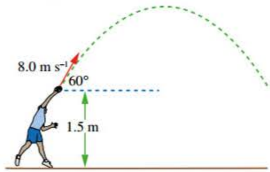
1 A javelin thrower launches her javelin at 40° above the horizontal. Select the correct statement about the javelin at the highest point of its path.

A It has zero acceleration.
B It is at its slowest speed.
C There are forward and downward forces acting on it.
D There are no forces acting on it since it is in free fall.

2 A basketballer shoots for a goal by launching the ball at $15\,m\,s^{-1}$ at 25° to the horizontal.

a Calculate the initial horizontal speed of the ball.
b What is the initial vertical speed of the ball?
c What are the magnitude and direction of the acceleration of the ball when it is at its maximum height?
d What is the velocity of the ball when it is at its maximum height?

3 In a shot-put event a 2.0 kg shot is launched from a height of 1.5 m, with an initial velocity of $8.0\,m\,s^{-1}$. The launch angle is 60° to the horizontal.



a What is the initial horizontal speed of the shot?
b What is the initial vertical speed of the shot?
c How long does it take the shot to reach its maximum height?
d What is the maximum height from the ground that is reached by the shot?
e What is the speed of the shot when it reaches its maximum height?

4 A projectile is launched at 60° from the horizontal and lands 50 m from the launch point. What was the projectile's speed at launch? (Hint: At the peak of the projectile's flight, the projectile is halfway to landing.)

Chapter review

02

KEY TERMS

air resistance
ballistic
free fall
magnitude
projectile
terminal velocity
vector notation
weight

REVIEW QUESTIONS

For the following questions, assume that the acceleration due to gravity is $9.8\,m\,s^{-2}$ and ignore the effects of air resistance unless otherwise stated.

1 A projectile is launched straight up. What is the initial velocity needed so that the projectile will reach 15 m above the ground?

2 A projectile is thrown upwards on Mars where $g = 3.80\,m\,s^{-2}$. What is the initial velocity needed to reach a point 20 m above the ground?

3 A squirrel is 9.0 m up in a tree and tosses a nut straight up with an initial velocity of $1.5\,m\,s^{-1}$. The squirrel climbs down the tree to the ground in 2.0 s. Does the squirrel reach the ground before the nut?

4 A stone is thrown horizontally at $5\,m\,s^{-1}$. Ignoring air resistance, which statement best describes the motion of the stone as it falls? More than one option can be correct.

A The stone travels in a circular path.
B The only force acting on the stone is gravity.
C There is a driving force acting on the stone.
D The stone's speed increases.

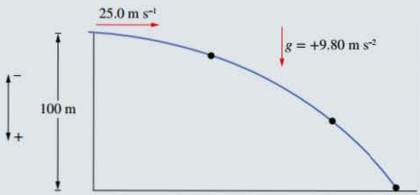
5 A skateboard travelling at $4.0\,m\,s^{-1}$ rolls off a horizontal bench that is 1.2 m high.

a How long does the board take to hit the ground?
b How far does the board land from the base of the bench?
c What is the magnitude and direction of the acceleration of the board just before it lands?

6 A marble travelling at $2.0\,m\,s^{-1}$ rolls off a horizontal bench and takes 0.75 s to reach the floor.

a How far does the marble travel horizontally before landing?
b What is the vertical component of the marble's speed as it lands?
c What is the speed of the marble as it lands?

7 A tourist stands on top of a sea cliff that is 100 m high. The tourist throws a rock horizontally at $25.0\,m\,s^{-1}$ into the sea.



a What is the vertical component of the final speed?
b At what angle is the rock travelling relative to the horizontal as it reaches the water?
c What is the velocity of the rock as it reaches the water?
d A buoy is located 120 m from shore. How much faster or slower would the tourist need to throw the rock for it to land near the buoy?

8 Two identical tennis balls A and B are hit horizontally from a point 2.0 m above the ground with different initial speeds: ball A has an initial speed of $5.0\,m\,s^{-1}$, while ball B has an initial speed of $10\,m\,s^{-1}$.

a Calculate the time it takes for ball A to strike the ground.
b Calculate the time it takes for ball B to strike the ground.
c How much further than ball A does ball B travel in the horizontal direction before bouncing?

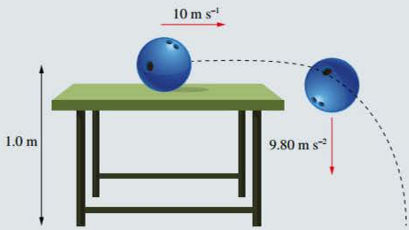
9 Joe throws a hockey ball horizontally at $5\,m\,s^{-1}$. He then throws a polystyrene ball of identical dimensions at the same speed horizontally. If air resistance is taken into account, which of the balls will travel further? Why?

10 An archer stands on top of a platform that is 20 m high and fires an arrow horizontally at 50 m s^{-1}.

a What is the speed of the arrow as it reaches the ground?

b At what angle relative to the horizontal is the arrow travelling as it reaches the ground?

11 A bowling ball of mass 7.5 kg travelling at 10 m s^{-1} rolls off a horizontal table 1.0 m high.



a Calculate the ball's horizontal and vertical velocity just as it strikes the floor.

b Calculate the velocity of the ball as it reaches the floor.

c What time interval has elapsed between the ball leaving the table and striking the floor?

d Calculate the horizontal distance travelled by the ball as it falls.

12 A toy car is moving at 2.5 m s^{-1} as it rolls off a horizontal platform. The car takes 1.0 s to reach the floor.

a How far does the car land from the foot of the table?

b What are the magnitude and direction of acceleration when the car is halfway to the floor?

c How long does the car take to reach a point halfway to the floor?

d What is the car's speed when it is halfway to the floor?

13 A rocket made from a plastic bottle is designed so that it is launched with a velocity of 18 m s^{-1} at an angle of 30° to the horizontal. It is fired towards a target 25 m away. If the target has a 1.0 m radius, does the rocket hit the target?

14 In a tennis match, a tennis ball is hit from a height of 1.2 m with an initial velocity of 16 m s^{-1} at an angle of 50° to the horizontal.

a What is the initial horizontal speed of the ball?

b What is the initial vertical speed of the ball?

c What is the maximum height that the ball reaches above the court surface?

15 A rugby player kicks for a goal by taking a place kick with the ball at rest on the ground. The ball is kicked at an angle of 30° to the horizontal at 20 m s^{-1}. At its highest point, what is the speed of the ball?

16 A ball is launched at an angle of 20° to the horizontal with an initial speed of 5.0 m s^{-1}. If the effects of air resistance are taken into account, which one of the following statements would be correct?

A The ball would have travelled a greater horizontal distance before striking the ground.

B The ball would have reached a greater maximum height.

C The ball's horizontal velocity would have been continually decreasing.

D The ball's vertical acceleration would have increased.

17 A child is holding a garden hose at ground level and the water stream from the hose is travelling at 15 m s^{-1}.

a Calculate how far away the water hits the ground when the angle the child holds the hose is equal to:

i 45°

ii 55°

iii 35°.

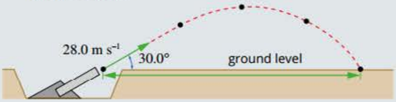
b A student is conducting an investigation into what effect the launch angle has on the displacement of a projectile and constructs the following hypothesis: If the launch angle of a projectile is 45° and the initial speed is kept constant, then the final displacement will be at a maximum.

i Do the results from part **a** seem to agree with this hypothesis?

ii Write a short description of an investigation which could test this hypothesis.

18 A computer game is designed in which the player has to launch different projectiles in order to knock down a castle. The game designer needs to calculate the distances travelled by changing the initial launch angle. A heavy red cannon ball is launched at 30° to the horizontal, and a lighter blue one is launched at 60°. Both cannon balls' initial speeds are the same. Which cannon travels further?

19 A senior physics class conducting a research project on projectile motion constructs a device that can launch a cricket ball. The launching device is designed so that the ball can be launched at ground level with an initial velocity of $28.0\,m\,s^{-1}$ at an angle of 30.0° to the horizontal.



a Calculate the horizontal component of the velocity of the ball:

i initially

ii after 1.0 s

iii after 2.0 s.

b Calculate the vertical component of the velocity of the ball:

i initially

i after 1.0 s

iii after 2.0 s.

c What is the speed of the cricket ball after 2.00 s?

d What is the speed of the ball as it lands?

e What horizontal distance does the ball travel before landing; that is, what is its range?

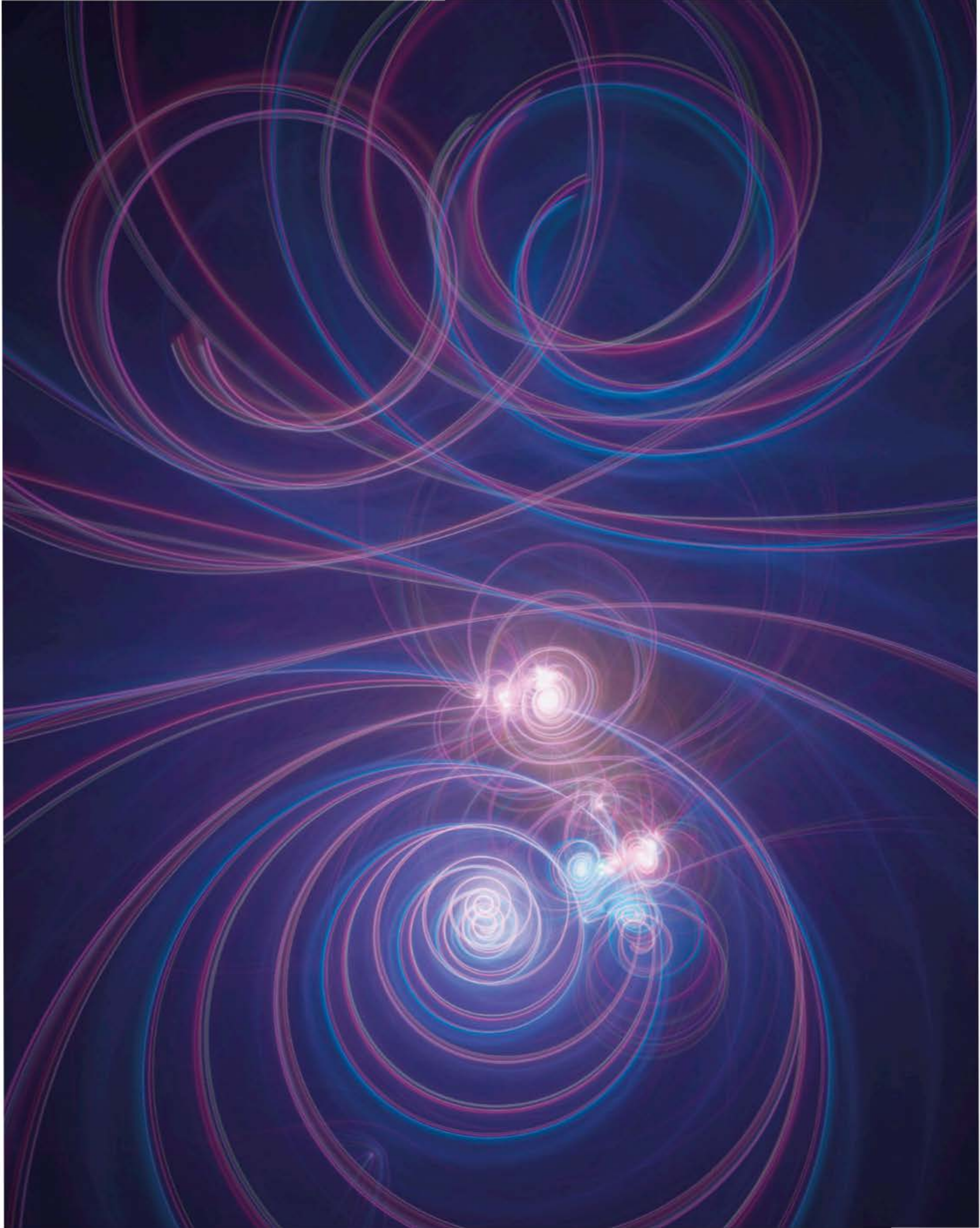
20 A ball is launched at $18\,m\,s^{-1}$ from the ground. The ball lands 20 m away.

a Derive a relationship between the time t and the angle θ by first calculating the initial horizontal velocity.

b Derive a relationship between the time t and the angle θ by first calculating the initial vertical velocity.

c At what angle was the ball launched? You will need to use the formula $2\sin\theta\cos\theta = \sin 2\theta$ (known as the double-angle formula).

21 After completing the activity on page 57, reflect on the inquiry question: How can models that are used to explain projectile motion be used to analyse and make predictions?



CHAPTER 03

Circular motion

An understanding of forces and fields has allowed humans to land on the Moon and to explore the outer reaches of the solar system. Satellites in orbit around the Earth have changed the way people live.

These advances have been achieved using Newton's laws of motion, which were published in the 17th century. Newton suggested that it should be possible to put satellites in orbit around the Earth almost 300 years before it was technically possible to do so. In this chapter Newton's laws will be used to analyse circular motion.

Content

INQUIRY QUESTION

Why do objects move in circles?

By the end of the chapter you will be able to:

- conduct investigations to explain and evaluate, for objects executing uniform circular motion, the relationships that exist between:
  - centripetal force
  - mass
  - speed
  - radius
- analyse the forces acting on an object executing uniform circular motion in a variety of situations; for example:
  - cars moving around horizontal circular bends
  - a mass on a string
  - objects on banked tracks (ACSPH100) CCT ICT
- solve problems, model and make quantitative predictions about objects executing uniform circular motion in a variety of situations, using the following relationships:
  - $a_c = \frac{v^2}{r}$
  - $v = \frac{2\pi r}{T}$
  - $F_c = \frac{mv^2}{r}$ ICT N
  - $\omega = \frac{\Delta\theta}{t}$
- investigate the relationship between the total energy and work done on an object executing uniform circular motion
- investigate the relationship between the rotation of mechanical systems and the applied torque $(\tau = r_{\perp}F = rF\sin\theta)$. ICT N

3.1 Circular motion

REVISION

The force of Newton

Newton's laws describe how forces can be used to explain the motion of bodies.

Newton's first law states that every object continues to be at rest or continues with constant velocity unless it experiences an unbalanced force. This is also called the law of inertia.

Newton's second law states that the acceleration of a body experiencing an unbalanced force is directly proportional to the net force acting on it and inversely proportional to the mass of the body: $\vec{F}_{net} = m\vec{a}$.

Newton's third law states that when one body exerts a force on another body (an action force), the second body exerts an equal force in the opposite direction on the first (a reaction force). Action–reaction pairs must act on different bodies. This can be written as $\vec{F}_{AB} = -\vec{F}_{BA}$.

PHYSICS INQUIRY CCT N

The force behind circular motion

Why do objects move in circles?

COLLECT THIS...

- tennis ball
- metre ruler
- hula hoop or other circular edge
- soft toy with string attached
- straw
- water bottle
- stopwatch

DO THIS...

Safety: Make sure the soft toy doesn't have any small objects attached, e.g. plastic or metal beads. Wear safety glasses.

Part 1

1. Place the hula hoop on a flat surface. Move the tennis ball around the inside edges of the hula hoop. Once the ball is moving in a circle, lift the hula hoop to remove the force that is keeping the ball moving in a circle. Record the direction the ball moves in.
2. Repeat three times to confirm the direction.

Part 2

3. Start the tennis ball rolling. Your task is to use the metre ruler to hit the ball to get it to travel in a circle. Record the direction of the force—towards the centre, away from the centre, tangent to the circle, or in another direction.

Part 3

4. Tie the string around the soft toy. Thread the string through the straw and tie it to a water bottle. Fill the bottle half full with water.
5. Record the mass of the toy and the mass of the water bottle.
6. Holding the straw vertically, spin the toy in a circle. When the water bottle stops moving, record the radius of the circle and the time to travel 10 revolutions.
7. Calculate the velocity of the stuffed toy using the time for one revolution, and the circumference of the circle.
8. Using the centripetal force equation, and the force of gravity on the water bottle, calculate the velocity of the stuffed toy.

RECORD THIS...

Describe the direction of the net force in each part.

Present a free-body diagram of the forces.

REFLECT ON THIS...

What forces create circular motion?

Where have you seen circular motion before, and what forces were in action that added to create a net centripetal force?

Circular motion is common throughout the universe. On a small scale, this could involve children moving in a circular path on a fair ride (Figure 3.1.1) or passengers in a car as it travels around a roundabout. In athletics, hammer throwers swing the hammer in a circular path before releasing it at high speed. On a much larger scale, the planets orbit the Sun in roughly circular paths; and on an even grander scale, stars can travel in circular paths around the centres of their galaxies. This section explains the nature of circular motion in a horizontal plane, and applies Newton's first and second laws to different circular-motion problems.

FIGURE 3.1.1 The people on this ride are travelling in a circular path at high speed.

UNIFORM CIRCULAR MOTION

Figure 3.1.2 shows an athlete in a hammer-throw event, swinging a steel ball in a horizontal circle with a constant speed of $25\,m\,s^{-1}$. As the hammer travels in its circular path, its speed is constant, but its velocity is continually changing.

Remember that velocity is a vector. Since the direction of the hammer is changing, so too is its velocity, even though its speed is not changing.

The velocity of the hammer at any instant is **tangential** (at a tangent) to its path. At one instant, the hammer is travelling at $25\,m\,s^{-1}$ north, then an instant later at $25\,m\,s^{-1}$ west, then $25\,m\,s^{-1}$ south, and so on.

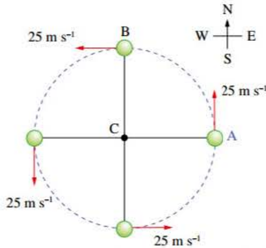


FIGURE 3.1.2 The velocity of the hammer (steel ball) at any instant is tangential to its path and is continually changing even though it has constant speed. This changing velocity means that the hammer is accelerating.

PERIOD AND FREQUENCY

Imagine that an object is moving in a circular path with a constant speed, v, and a radius of r metres, and it takes T seconds to complete one revolution. The time required to travel once around the circle is called the **period**, T, of the motion. The number of rotations each second is the **frequency**, f.

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

where

f is the frequency (Hz)

T is the period (s)

SPEED

An object that travels in a circle will travel a distance equal to the circumference of the circle, $C = 2\pi r$, with each revolution (Figure 3.1.3). Given that the time for each revolution is the period, T, the average speed of the object is:

$$\text{speed} = \frac{\text{distance}}{\text{time}} = \frac{\text{circumference}}{\text{period}}$$

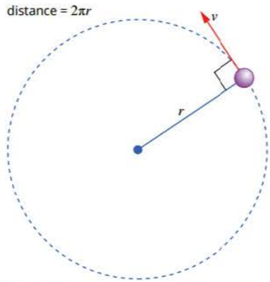


FIGURE 3.1.3 The average speed of an object moving in a circular path is given by the distance travelled in one revolution (the circumference) divided by the time taken (the period, T).

SKILLBUILDER N

Converting units

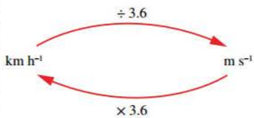
The usual unit in physics for velocity is $m\,s^{-1}$, but $km\,h^{-1}$ is often used in everyday life. So it is important to understand how to convert between them.

You should be familiar with $100\,km\,h^{-1}$ because it is the speed limit for most freeways and country roads in Australia. Since there are 1000 metres in 1 km and 3600 s in 1 hour (60 × 60 min), this is the same as travelling 100 000 m in 3600 s.

$$
\begin{aligned}
100\,km\,h^{-1} &= 100 \times 1000\,m\,h^{-1}\\
&= 100\,000\,m\,h^{-1}\\
&= \tfrac{100\,000}{3600}\,m\,s^{-1}\\
&= 27.8\,m\,s^{-1}
\end{aligned}
$$

Therefore $km\,h^{-1}$ can be converted to $m\,s^{-1}$ by multiplying by $\frac{1000}{3600}$ (or dividing by 3.6).

The diagram below summarises the conversion between $km\,h^{-1}$ and $m\,s^{-1}$.



In circular motion, this equation is represented as follows.

> The average speed of an object moving in a circular path is:
>
> $v = \frac{2\pi r}{T}$
>
> where
>
> v is the speed ($m\,s^{-1}$)
>
> r is the radius of the circle (m)
>
> T is the period of motion (s)

Worked example 3.1.1

CALCULATING SPEED

A wind turbine has blades 55.0 m in length that rotate at a frequency of 20 revolutions per minute. At what speed do the tips of the blades travel? Express your answer in $km\,h^{-1}$.

Thinking	Working
Calculate the period, T. Remember to express frequency in the correct units. Alternatively, recognise that 20 revolutions in 60 s means that each revolution takes 3 s.	20 revolutions per minute = $\frac{20}{60}$ = 0.333 Hz $T = \frac{1}{f}$ $= \frac{1}{0.333}$ $= 3.0$ s
Substitute r and T into the formula for speed and solve for v.	$v = \frac{2\pi r}{T}$ $= \frac{2 \times \pi \times 55.0}{3.0}$ $= 115.2\,m\,s^{-1}$
Convert $m\,s^{-1}$ into $km\,h^{-1}$ by multiplying by 3.6.	$115.2 \times 3.6 = 415\,km\,h^{-1}$

Worked example: Try yourself 3.1.1

CALCULATING SPEED

A water wheel has blades 2.0 m in length that rotate at a frequency of 10 revolutions per minute. At what speed do the tips of the blades travel? Express your answer in $km\,h^{-1}$.

ANGULAR VELOCITY

When objects travel in circular paths it can be convenient to measure the angle of rotation in a given time. The **angular velocity**, ω (Greek symbol omega), of an object travelling through an angle θ in a period of time, t, can be calculated using the following equation.

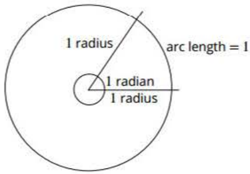
> $\omega = \frac{\Delta\theta}{t}$
>
> where
>
> ω is the angular velocity ($°\,s^{-1}$ or $rad\,s^{-1}$)
>
> $\Delta\theta$ is the angle travelled (° or rad)
>
> t is the time (s)

SKILLBUILDER N

Radians and degrees

Radians are another unit of measurement for angles. The unit symbol for radians is ^{c}, although often it is written as rad, as has been done in this chapter, to avoid confusion with the degrees symbol.

Radians are proportional to the radius of a unit circle. An angle of 1 radian is defined as the angle at the centre of a unit circle that marks out an arc length of 1 unit circle radius length, as shown in the diagram.



Using the circumference formula:

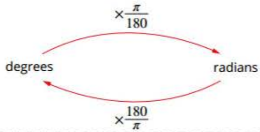
$C = 2\pi r$

$= 2\pi \times 1$ (when $r = 1$)

$= 2\pi$

In a unit circle, the circumference is 2π units. Because of this, angles (in degrees) can be related to the arc length (in radians) on a unit circle. In a complete circle, the angle 360° is equal to 2π rad.

Therefore degrees can be converted to radians by multiplying by $\frac{\pi}{180}$. Radians can be converted to degrees by multiplying by $\frac{180}{\pi}$.



SKILLBUILDER N

Angles in circles

In circular motion, in order to calculate the angular velocity you may need to first find the angle $\Delta\theta$ (also known as the angular displacement) the object travels through. Say you are given the radius of motion r and the object travels through a length l (the length around AB in the diagram below). The angle can be found using the following formula:

$$\Delta\theta = \frac{l}{r}$$

where

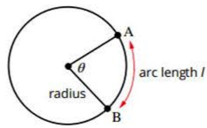
$\Delta\theta$ is the angular displacement (rad)

l is the arc length (m)

r is the radius (m)

For example, say you want to find the angle of a tennis ball travelling around a Totem Tennis pole. The radius of the string connecting the ball to the pole is 1.5 m, and the ball travels around a 40 cm arc (between points A and B in the diagram below). The angle the ball travels in this time is:

$$\Delta\theta = \frac{l}{r} = \frac{0.4}{1.5} = 0.27 \text{ rad}$$



Worked example 3.1.2

CALCULATING ANGULAR VELOCITY

A wind turbine has blades 55.0 m in length that rotate so the tips of the blade travel a distance of 150 m in 2 s. At what angular velocity does the turbine rotate? Express your answer in $°s^{-1}$.

Thinking	Working
Calculate the angle $\Delta\theta$ in radians.	$\Delta\theta = \frac{l}{r}$ $= \frac{150}{55}$ $= 2.7$ rad
Convert the angle to degrees.	$\Delta\theta = 2.7 \times \frac{180}{\pi}$ $= 156.3°$
Substitute $\Delta\theta$ and t into the formula for angular velocity and solve for ω.	$\omega = \frac{\Delta\theta}{t}$ $= \frac{156.3}{2}$ $= 78.2° s^{-1}$

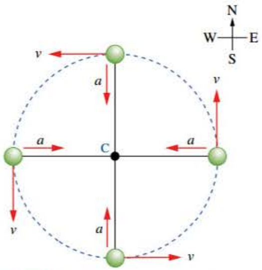


FIGURE 3.1.4 A body moving in a circular path has an acceleration towards the centre of the circle. This is known as a centripetal acceleration.

GO TO ➤ Section 4.2 page 115

Worked example: Try yourself 3.1.2

CALCULATING ANGULAR VELOCITY

A truck wheel of diameter 1 m travels over 8 m of ground in 3 s. What is the angular velocity of the wheels? Express your answer in $°s^{-1}$.

CENTRIPETAL ACCELERATION

When objects travel in circular paths, they can have a constant speed, yet at the same time have a velocity that is changing. This seeming contradiction arises because speed is a scalar quantity, whereas velocity is a vector.

Since the velocity of the object is changing, it is accelerating even though its speed is not changing. The object is continually deviating inwards from its straight-line direction and so has an acceleration towards the centre of its motion. This acceleration is known as **centripetal acceleration**, a_c. In Figure 3.1.4, the velocity vector of an object travelling in a circular path is shown with an arrow labelled v. Notice how it is at a tangent to the circular path. The acceleration is always towards the centre of the circular path.

However, as Figure 3.1.4 shows, even though the object is accelerating towards the centre of the circle, it never gets any closer to the centre. This is the same principle that applies to satellites in orbit, which will be studied in Chapter 4.

The centripetal acceleration a_c of an object moving in a circular path of radius r with a velocity v can be found from the relationship:

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

A substitution can be made for the speed of the object in this equation:

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

Then the centripetal acceleration is given by:

$$\begin{aligned} a_c &= \frac{v^2}{r} \\ &= \left(\frac{2\pi r}{T}\right)^2 \times \frac{1}{r} \\ &= \frac{4\pi^2 r}{T^2} \end{aligned}$$

Centripetal acceleration is always directed towards the centre of the circular path and is given by

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

where

a_c is the centripetal acceleration ($m\,s^{-2}$)
v is the magnitude of the velocity ($m\,s^{-1}$)
r is the radius of the circle (m)

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

where

a_c is the centripetal acceleration ($m\,s^{-2}$)
v is the speed ($m\,s^{-1}$)
r is the radius of the circle (m)
T is the period of motion (s)

FORCES THAT CAUSE CIRCULAR MOTION

As with all forms of motion, an analysis of the forces that are acting is needed to understand why circular motion occurs. In the hammer-throw event described earlier in this section, the hammer ball is continually accelerating. It follows from Newton's second law that there must be a net unbalanced force continuously acting on it. The net unbalanced force that gives the hammer ball its acceleration towards the centre of the circle is known as a **centripetal force**.

In every case of circular motion, a real force is necessary to provide the centripetal force. This force acts in the same direction as the acceleration, that is, towards the centre of the circle. This centripetal force can be provided in a number of ways. For the hammer in Figure 3.1.5a, the centripetal force is the tension force in the cable. Other examples of centripetal force are also shown in Figure 3.1.5.

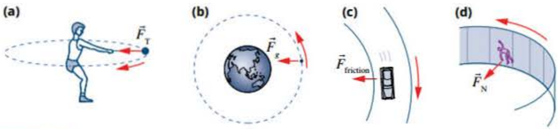


FIGURE 3.1.5 (a) In a hammer throw, tension in the cable provides the centripetal force. (b) For planets and satellites, the gravitational attraction to the central body provides the centripetal force. (c) For a car on a curved road, the friction between the tyres and the road provides the centripetal force. (d) For a person in the Gravitron ride, the normal force from the wall provides the centripetal force.

Now, consider the consequences if the unbalanced force ceases to act. In the example of the hammer thrower, if the tension in the wire becomes zero because the thrower releases the ball, there is no longer a force causing the ball to change direction. The result is that the ball then moves in a straight line tangential to its circular path, as would be expected from Newton's first law.

The centripetal force can then be found using Newton's second law, $F_{net} = ma$, and substituting in the centripetal acceleration.

Centripetal force is given by:

$$F_c = ma = \frac{mv^2}{r}$$

$$F_c = \frac{4\pi^2 rm}{T^2}$$

where

F_c is the net or centripetal force on the object (N)

m is the mass (kg)

a is the acceleration ($m\,s^{-2}$)

v is the velocity ($m\,s^{-1}$)

r is the radius of the circle (m)

T is the period of motion (s)

PHYSICSFILE L

Centripetal force or centrifugal force?

When going around a corner in a car you feel as if you are being pushed outwards. This force is called a centrifugal force, and comes from the Latin words *centrum* (meaning centre) and *fugere* (meaning to flee). If you were able to observe this from above, you would see something different. The passenger's body has inertia travelling forwards. As the car turns, the body initially continues in a straight line as shown in Figure 3.1.6. From the passenger's point of view, it would feel like they were being pushed outwards as they slid along the seat. When they reach the car door, the **normal reaction force** pushes the person towards the centre of the car. This is the net force that creates circular motion, and is called the centripetal force. The word centripetal comes from the Latin *centrum* (meaning centre) and *petere* (meaning to seek).

The centrifugal force is an apparent force, it does not really exist, and should not be used in calculations or explanations. Instead, the centripetal force is the force causing the circular motion.

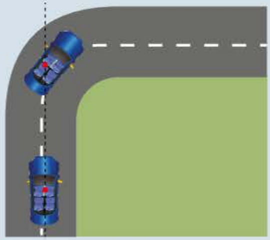


FIGURE 3.1.6 From an accelerating reference frame inside the car, it feels as if there is a force pushing outwards. When the same situation is viewed from above, the force that produces circular motion is seen to be inward, and the inertia of the body creates the illusion of a centrifugal force.

Worked example 3.1.3

CENTRIPETAL FORCES

An athlete in a hammer-throw event is swinging the ball of mass 7.0 kg in a horizontal circular path. The ball is moving at 20 m s^{-1} in a circle of radius 1.6 m.

a Calculate the magnitude of the acceleration of the ball.	
Thinking	**Working**
As the object is moving in a circular path, the centripetal acceleration is towards the centre of the circle. Write down the other variables that are given.	$v = 20\,\text{m s}^{-1}$ $r = 1.6\,\text{m}$ $a_c = ?$
Find the equation for centripetal acceleration that fits the information you have, and substitute the values.	$a_c = \frac{v^2}{r}$ $= \frac{20^2}{1.6}$ $= 250\,\text{m s}^{-2}$
The magnitude is required, so no direction is needed in the answer.	The acceleration of the ball is 250 m s^{-2}.

b Calculate the magnitude of the tensile (tension) force acting in the wire.	
Thinking	**Working**
Identify the unbalanced force that is causing the object to move in a circular path. Write down the information that you are given.	$m = 7.0\,\text{kg}$ $a = 250\,\text{m s}^{-2}$ $F_{net} = ?$
Select the equation for centripetal force, and substitute the variables you have.	Equation for centripetal force: $F_c = F_{net} = ma$ $= 7.0 \times 250$ $= 1.8 \times 10^3\,\text{N}$
The magnitude is required, so no direction is needed in the answer.	The force of tension in the wire is the unbalanced force that is causing the ball to accelerate. Tensile force $F_T = 1.8 \times 10^3\,\text{N}$

Worked example: Try yourself 3.1.3

CENTRIPETAL FORCES

PA 5.3

PA 5.3

An athlete in a hammer-throw event is swinging the ball of mass 7.0 kg in a horizontal circular path. The ball is moving at $25\,m\,s^{-1}$ in a circle of radius 1.2 m.

a Calculate the magnitude of the acceleration of the ball.

b Calculate the magnitude of the tensile force acting in the wire.

PHYSICS IN ACTION N

The Gravitron

The Gravitron (also known as the Vortex or Rotor) can rotate at 24 rpm and has a radius of 7 m. The centripetal acceleration can be over $40\,m\,s^{-2}$. This is caused by a very large centripetal force from the wall—i.e. the normal force, $\vec{F}_N$, which is greater than the weight force, $\vec{F}_g$. Since the wall exerts such a large force, the patrons are pinned firmly to the wall as an upward frictional force, $\vec{F}_{friction}$, acts to hold them up. The floor then drops away. It is important to note that there is no outward force acting. In fact, as you can see in Figure 3.1.7, these forces are unbalanced and the net force is equal in size and direction to the normal force towards the centre of the circle.

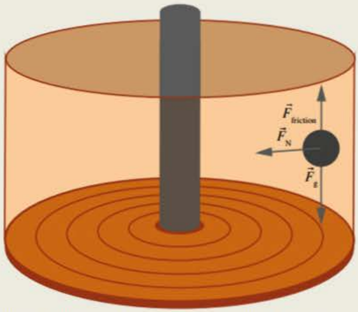


FIGURE 3.1.7 The forces acting on the person are unbalanced. There is an unbalanced force from the wall giving the person a centripetal acceleration.

3.1 Review

SUMMARY

- Frequency, f, is the number of revolutions each second and is measured in hertz (Hz).
- Period, T, is the time for one revolution and is measured in seconds.
- The relationship between T and f is:

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

- An object moving with a uniform speed in a circular path of radius r and with a period T has an average speed, v, that is given by:

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

- The velocity of an object moving (with a constant speed) in a circular path is continually changing. The velocity vector is always directed at a tangent to the circular path.
- The angular velocity of an object moving in a circular path can be calculated by:

$$\omega = \frac{\Delta\theta}{t}$$

- An object moving in a circular path (with a constant speed) has an acceleration due to its circular motion. This acceleration is directed towards the centre of the circular path and is called centripetal acceleration $\boldsymbol{a}_c$:

$$a_c = \frac{v^2}{r} \text{ and } a_c = \frac{4\pi^2 r}{T^2}$$

- Centripetal acceleration is a consequence of a centripetal force acting to make an object move in a circular path.
- Centripetal forces are directed towards the centre of the circle and their magnitude can be calculated by using Newton's second law:

$$F_c = \frac{mv^2}{r} \text{ and } F_c = \frac{4\pi^2 rm}{T^2}$$

- Centripetal force is always supplied by a real force, the nature of which depends on the situation. The real force is commonly friction, gravitation or the tension in a string or cable.

KEY QUESTIONS

1 Phil is standing inside a tram when it starts off suddenly. Len, who was sitting down, commented that Phil was 'thrown backwards' as the tram started moving. Is this a correct statement? Explain in terms of Newton's laws.

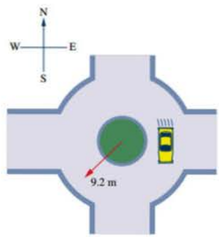
2 Explain why the normal force F_N and the weight force F_g are *not* an action–reaction pair for Newton's third law. State the third law pairs for a mug sitting on a table for both F_N and F_g.

3 A car is travelling with a constant speed around a roundabout. What is the centripetal force that is causing this circular motion?

A gravity
B friction
C drag
D tension

4 A boy is swinging a yo-yo in a horizontal circle five times each second. What is the period of the yo-yo?

5 A car of mass 1200 kg is travelling on a roundabout in a circular path of radius 9.2 m. The car moves with a constant speed of $8.0\,\text{m s}^{-1}$. The direction of the car is anticlockwise around the roundabout when viewed from above as shown.



a Which two of the following statements correctly describe the motion of the car as it travels around the roundabout?

A It has a constant speed.
B It has a constant velocity.
C It has zero acceleration.
D It has an acceleration that is directed towards the centre of the roundabout.

b When the car is in the position shown in the diagram, what is the:

i speed of the car
ii velocity of the car
iii magnitude and direction of the acceleration of the car?

c Calculate the magnitude and direction of the net force acting on the car at the position shown.

d Some time later, the car has travelled halfway around the roundabout. What is the:

i velocity of the car at this point

ii magnitude and direction of its acceleration at this point?

e If the driver of the car kept speeding up, what would eventually happen to the car as it travelled around the roundabout? Explain your answer.

6 An ice skater of mass 50 kg is skating in a horizontal circle of radius 1.5 m at a constant speed of $2.0\,m\,s^{-1}$.

a Determine the magnitude of the skater's acceleration.

b Are the forces acting on the skater balanced or unbalanced? Explain your answer.

c Calculate the magnitude of the centripetal force acting on the skater.

7 Jack and Jill are playing on a merry-go-round at the playground. The merry-go-round has a radius of 1.2 m. Jill pushes the outside of the merry-go-round a distance of 5.0 m.

a What angle of rotation does the merry-go-round travel through? Give your answer in degrees.

b What is the angular velocity if it takes Jill 0.50 s to travel the 5.0 m?

c Jack helps Jill push and they travel at an angular velocity of $50°\,s^{-1}$. How long does it take to do five revolutions?

3.2 Circular motion on banked tracks

The previous section focused on relatively simple situations involving uniform circular motion in a horizontal plane or vertical plane. However, there are more complex situations involving this type of motion. On many road bends, the road is not horizontal, but is at a small angle to the horizontal. This enables vehicles to travel at higher speeds without skidding. A more dramatic example of this effect is at a cycling velodrome like that shown in Figure 3.2.1. The Dunc Gray Velodrome used for the Sydney Olympics has banked or inclined corners that peak at 42°. This enables the cyclists to travel at much higher speeds than if the track were flat. This section examines the physics of conical pendulums and banked cornering, and applies Newton's laws to solving problems involving circular motion on banked tracks.

FIGURE 3.2.1 The Australian women's pursuit track cycling team in action on a banked velodrome track during the London Olympics in 2012.

BALL ON A STRING

You might have played Totem Tennis at one time. This is a game where a ball is attached to a pole by a string and can travel in a horizontal circle, although the string itself is not horizontal. This kind of motion is shown in Figure 3.2.2.

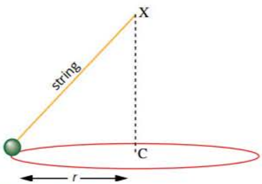


FIGURE 3.2.2 This ball is travelling in a horizontal circular path of radius r. The centre of its circular motion is at C.

If the ball at the end of the string was swinging slowly, the string would swing down at an angle closer to the pole. If the ball was swung faster, the string would become closer to being horizontal. In fact, it is not possible for the string to be absolutely horizontal, although as the speed increases, the closer to horizontal it becomes. This system is known as a conical pendulum.

If the angle of the conical pendulum is known, trigonometry can be used to find the radius of the circle and the forces involved.

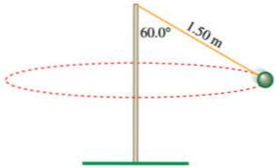
See the SkillBuilder on vector components for some useful trigonometric functions, and remember $\tan\theta = \frac{\text{opposite}}{\text{adjacent}}$.

GO TO ➤ SkillBuilder page 52

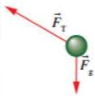
Worked example 3.2.1

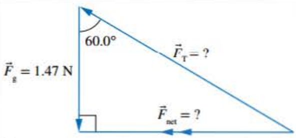
OBJECT ON THE END OF A STRING

During a game of Totem Tennis, the ball of mass 150 g is swinging freely in a horizontal circular path. The cord is 1.50 m long and is at an angle of 60.0° to the vertical, as shown in the diagram.



a Calculate the radius of the ball's circular path.	
Thinking	**Working**
The centre of the circular path is not the top end of the cord, but is where the pole is level with the ball. Use trigonometry to find the radius.	$r = 1.50 \sin 60.0$ $= 1.30\ \text{m}$

b Draw and identify the forces that are acting on the ball at the instant shown in the diagram.	
Thinking	**Working**
There are two forces acting—the tension in the cord, $\vec{F}_T$, and gravity, $\vec{F}_g$. These forces are unbalanced.	

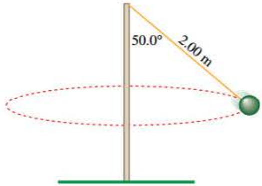
c Determine the net force that is acting on the ball at this time.	
Thinking	**Working**
First calculate the weight force $\vec{F}_g$.	$\vec{F}_g = m\vec{g}$ $= 0.150 \times 9.80$ $= 1.47\ \text{N}$ downwards
The ball has an acceleration that is towards the centre of its circular path. This is horizontal and towards the left at this instant. The net force will also lie in this direction at this instant. A force triangle and trigonometry can be used here.	 $\vec{F}_{net} = 1.47 \tan 60.0$ $= 2.55\ \text{N}$ towards the left

d Calculate the size of the tensile force in the cord.	
Thinking	**Working**
Use trigonometry to find F_T. The size of the force is a scalar and doesn't require a direction.	$F_T = \frac{1.47}{\cos 60.0}$ $= 2.94\ \text{N}$

Worked example: Try yourself 3.2.1

OBJECT ON THE END OF A STRING

During a game of Totem Tennis, the ball of mass 200 g is swinging freely in a horizontal circular path. The cord is 2.00 m long and is at an angle of 50.0° to the vertical, as shown in the diagram.



a Calculate the radius of the ball's circular path.

b Draw and identify the forces that are acting on the ball at the instant shown in the diagram.

c Determine the net force that is acting on the ball at this time.

d Calculate the size of the tensile force in the cord.

BANKED CORNERS

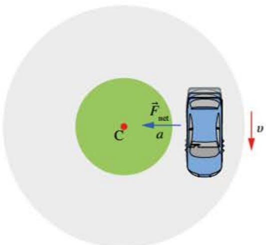


FIGURE 3.2.3 The car is travelling in a circular path on a horizontal track.

Cars and bikes can travel much faster around corners when the road or track surface is inclined or banked at some angle to the horizontal. **Banked tracks** are most obviously used at cycling velodromes or motor sport events such as NASCAR races. Road engineers also design roads to be banked in places where there are sharp corners such as exit ramps on freeways.

When cars travel in circular paths on horizontal roads, they are relying on the force of friction between the tyres and the road to provide the sideways force that keeps the car turning in the circular path.

Consider a car travelling clockwise around a horizontal roundabout with a constant speed, v. As can be seen in Figure 3.2.3, the car has an acceleration towards the centre of the circle, C, and so the net force is also sideways on the car towards C.

The forces acting on the car are shown in Figure 3.2.4. As you can see, the vertical forces (gravity and the normal reaction force) are balanced. The only horizontal force is the sideways force that the road exerts on the car tyres. This is a force of friction, $\vec{F}_{friction}$, and is unbalanced, so this is equal to the net force, $\vec{F}_{net}$.

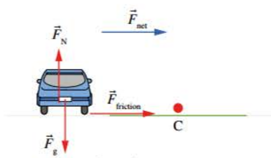


FIGURE 3.2.4 The vertical forces balance, and it is friction between the tyres and the road that enables the car to turn the corner.

If the car drove over an icy patch, there would be no friction and the car would not be able to turn. It would skid in a straight line at a tangent to the circular path.

Banking the road reduces the need for a sideways frictional force and allows cars to travel faster without skidding off the road and away from the circular path. Consider the same car travelling around a circular, banked road with constant speed, v, as shown in Figure 3.2.5. It is possible for the car to travel at a speed so that there is no sideways frictional force. This is called the **design speed** and it is dependent on the angle, θ, at which the track is banked. At this speed, the car exhibits no tendency to drift higher or lower on the track.

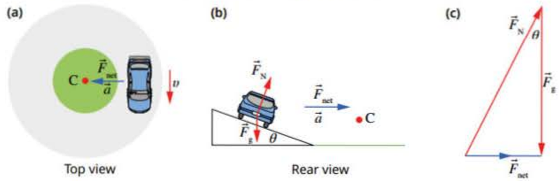


FIGURE 3.2.5 (a) The car is travelling in a circular path on a banked track. (b) The acceleration and net force are towards C. The banked track means that the normal force ($\vec{F}_N$) has an inward component. This is what enables the car to turn the corner. (c) Vector addition gives the net force ($\vec{F}_{net}$) as acting horizontally towards the centre of the circle of motion.

The car still has an acceleration towards the centre of the circle, C, and so there must be an unbalanced force in this direction. Due to the banking, there are now only two forces acting on the car: its weight, $\vec{F}_g$, and the normal force, $\vec{F}_N$, from the track.

As can be seen in Figure 3.2.5c, these forces are unbalanced. They add together to give a net force that is horizontal and directed towards C.

At the design speed, the angle of bank, θ, of the road or track can be found by using:

$$\tan\theta = \frac{F_{net}}{F_g}$$

where

F_{net} is the force acting towards the centre of the circle (N)

F_g is the force due to gravity on the object (N)

Extending this equation by substituting $F_{net} = F_c = \frac{mv^2}{r}$ and $F_g = mg$ gives:

$$\tan\theta = \frac{v^2}{rg} \text{ and hence } \theta = \tan^{-1}\left(\frac{v^2}{rg}\right)$$

where

v is the speed of the vehicle ($m\,s^{-1}$)

r is the radius of the track (m)

θ is the angle of bank (°)

g is the magnitude of the acceleration due to gravity ($9.80\,m\,s^{-2}$ near the surface of the Earth)

FIGURE 3.2.6 Australian cyclist Anna Meares on this banked velodrome track is cornering at speeds far higher than she could use on a flat track. Cyclists on a velodrome do not need to rely on friction to turn because they experience a larger normal force than usual.

If the angle and weight are known, trigonometry can be used to calculate the net force (Figure 3.2.5c) and therefore the design speed.

Rearranging $\tan\theta = \frac{v^2}{rg}$ to make the design speed, v, the subject gives:

$$v^2 = rg\tan\theta$$

$$v = \sqrt{rg\tan\theta}$$

It is worth noting that the normal force will be larger here than on a flat track. In the case of a cyclist, the rider and bike would feel a larger force acting from the road when they are on a banked track compared to when they are cycling on a flat track (Figure 3.2.6).

PHYSICSFILE N

Forces on a banking aeroplane

The main forces acting on an aeroplane as it is flying are weight, lift, drag and thrust. In Figure 3.2.7 the aeroplane is seen front-on—the thrust force would be pushing it out of the page and the drag force would be pulling it into the page. When the aeroplane is travelling in a straight level line (Figure 3.2.7a) the lift force is equal in magnitude but opposite in direction to the weight force. The net force acting on the plane is 0 N. When an aeroplane turns a corner, which is called banking, the aeroplane tilts one wing down and one wing up. This changes the orientation of the lift force as shown in Figure 3.2.7b, providing an unbalanced force radially inward. The angle the aeroplane tilts at is called the bank angle. At the same speed, different bank angles will result in different turning radii.

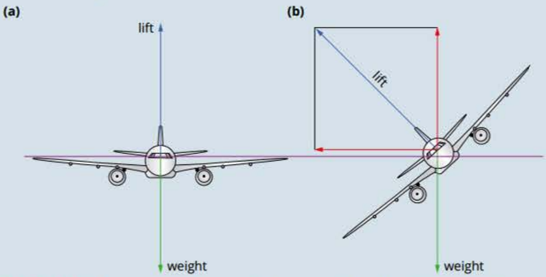


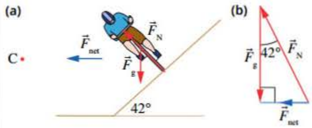
FIGURE 3.2.7 (a) Forces on an aeroplane travelling in a straight level line. The net force is 0 N. (b) Forces on an aeroplane banking, or turning a corner. The net force is radially inwards.

Worked example 3.2.2

BANKED CORNERS

A curved section of track on an Olympic velodrome has a radius of 50 m and is banked at an angle of 42° to the horizontal. A cyclist of mass 75 kg is riding on this section of track at the design speed.

a Calculate the net force acting on a cyclist at this instant if they are riding at the design speed.

Thinking	Working
Draw a force diagram and include all forces acting on the cyclist.	The forces acting on the cyclist are gravity and the normal force from the track, and these are unbalanced. The net force is horizontal and towards the centre of the circular track as shown in diagram a and the force triangle of diagram b. 
Calculate the weight force, $\vec{F}_g$.	$\vec{F}_g = m\vec{g}$ $= 75 \times 9.80$ $= 735$ N downwards
Use the force triangle and trigonometry to work out the net force, $\vec{F}_{net}$.	$\tan\theta = \frac{\vec{F}_{net}}{\vec{F}_g}$ $\tan 42 = \frac{\vec{F}_{net}}{735}$ $\vec{F}_{net} = 0.90 \times 735$ $= 661.8$ $= 660$ N horizontally towards the centre of the circle

b Calculate the design speed for this section of the track.

Thinking	Working
List the known values.	$m = 75$ kg $r = 50$ m $\theta = 42°$ $\vec{F}_g = 735$ N down $\vec{F}_{net} = 660$ N towards the centre of the circle $v = ?$
Use the design speed formula.	$v = \sqrt{rg\tan\theta}$ $= \sqrt{50 \times 9.80 \times \tan 42}$ $= 21\ \text{m s}^{-1}$

Worked example: Try yourself 3.2.2

BANKED CORNERS

A curved section of track on an Olympic velodrome has a radius of 40 m and is banked at an angle of 37° to the horizontal. A cyclist of mass 80 kg is riding on this section of track at the design speed.

a Calculate the net force acting on a cyclist at this instant as they are riding at the design speed.

b Calculate the design speed for this section of the track.

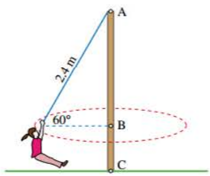
3.2 Review

SUMMARY

- A conical pendulum is a mass that travels in a horizontal circle on a string. The tension in the string supplies the force to counter the force of gravity and the centripetal force.
- A banked track is one where the track is inclined at some angle to the horizontal. This enables vehicles to travel at higher speeds when cornering, compared with around a horizontal curved path.
- Banking a track eliminates the need for a sideways frictional force to turn. When the speed and angle are such that there is no sideways frictional force, the speed is known as the design speed.
- The forces acting on a vehicle travelling at the design speed on a banked track are gravity and the normal force from the track. These forces are unbalanced and add to give a net force directed towards the centre of the circular motion.
- At the design speed, the bank angle of the track, $\theta = \tan^{-1}\left(\frac{v^2}{rg}\right)$.
- For a given bank angle and curve radius, the design speed is given by $v = \sqrt{rg\tan\theta}$.

KEY QUESTIONS

1 A child of mass 30 kg is playing on a maypole swing in a playground. The rope is 2.4 m long and at an angle of 60° to the horizontal as she swings freely in a circular path. Ignore the mass of the rope in your calculations.



a Calculate the radius of her circular path.

b Identify the forces that are acting on her as she swings freely.

c What is the direction of her acceleration when she is at the position shown in the diagram?

d Calculate the net force acting on the girl.

e What is her speed as she swings?

2 A cyclist is riding along a circular section of a velodrome where the radius is 30 m and the track is inclined at 30° to the horizontal. The cyclist is riding at the design speed and maintains a constant speed. Describe the direction of the acceleration on the cyclist.

3 An architect is designing a velodrome and the original plans have semicircular sections of radius 15 m and a bank angle of 30°. The architect is asked to make changes to the plans that will increase the design speed for the velodrome. What two design elements could the architect change in order to meet this requirement?

4 A racing car is travelling around a circular banked track which has a design speed of 100 km h^{-1}. On one lap, the car travels at 150 km h^{-1}. At this higher speed, the car would tend to travel in a different position along the banked surface. Would the car travel higher or lower up the banked track? Explain your answer.

5 A racing car travels at high speed along a horizontal track and tries to turn a corner. The car skids and loses control. The racing car then travels along a banked track and is able to travel much faster around the corners without skidding at all. Complete the sentences below by choosing the correct term from those given in bold.

On the horizontal track, the car is depending on the force of **friction/weight** to turn the corner. The size of the **friction/normal** force is equal to the **weight/friction** of the car, so these vertical forces are **balanced/unbalanced**. When driving on the banked track, the **normal/weight** force is not vertical and so is not balanced by the **weight/normal** force. In both cases, the forces acting on the car are unbalanced.

6 Copy and complete the following diagram by drawing and labelling the normal force, weight force and net force acting on the bicycle.

7 A car racing track is banked so that when the cars corner at $40\,m\,s^{-1}$, they experience no sideways frictional forces. The track is circular with a radius of 150 m. Calculate the angle to the horizontal at which the track is banked.

8 A section of track at a NASCAR raceway is banked to the horizontal. The track section is circular with a radius of 80 m and it has a design speed of $18\,m\,s^{-1}$. A car of mass 1200 kg is being driven around the track at $18\,m\,s^{-1}$.

a i Calculate the magnitude of the net force acting on the car (in kN).

ii Calculate the angle to the horizontal at which the track is banked.

b The driver now drives around the track at $30\,m\,s^{-1}$. What would the driver have to do to maintain their circular path around the track?

3.3 Work and energy

In previous sections, the motion of objects travelling in circular paths was discussed. It was explained that a body moving with constant speed in a horizontal circular path has an acceleration that is directed towards the centre of the circle. The same applies for vertical circular paths.

When you travel on a roller-coaster, you can experience quite strong forces pushing you down into the seat as you fly through the dips. On the other hand, as you travel over the humps, you tend to lift off the seat. These forces relating to circular motion in a vertical plane will be discussed in this section. As in the previous sections, Newton's laws are used to solve problems involving this type of circular motion.

UNIFORM HORIZONTAL MOTION

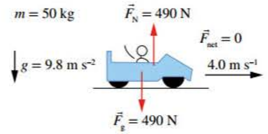


FIGURE 3.3.1 The vertical forces are in balance in this situation, i.e. $\vec{F}_N = \vec{F}_g$.

Theme park rides make you appreciate that the forces you experience throughout a ride can vary greatly. First, consider the case of a person in a roller-coaster cart, like that shown in Figure 3.3.1, travelling horizontally at $4.0\,\mathrm{m\,s^{-1}}$. If the person's mass is 50 kg and the gravitational field strength is $9.8\,\mathrm{m\,s^{-2}}$, the forces acting on the person can be calculated. These forces are the weight, $\vec{F}_g$, and the normal reaction force, $\vec{F}_N$, from the seat.

The person is moving in a straight line with a constant speed, so there is no unbalanced force acting. The weight force balances the normal reaction force from the seat. The normal force is therefore 490 N up, which is what usually acts upwards on this person when moving horizontally and they would feel the same as their usual weight.

Circular motion: travelling through dips

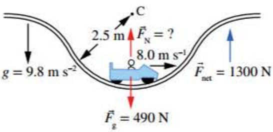


FIGURE 3.3.2 The person has a centripetal acceleration that is directed upwards towards the centre of the circle, and so the net force is also upwards. In this case, the magnitude of the normal force, $\vec{F}_N$, is greater than the weight, $\vec{F}_g$, and produces a situation where the rider feels heavier than usual.

Now consider the forces that act on the person after the cart has reached the bottom of a circular dip of radius 2.5 m and is moving at $8.0\,\mathrm{m\,s^{-1}}$. Figure 3.3.2 illustrates these forces.

The person will have a centripetal acceleration due to the circular path. This centripetal acceleration is directed towards the centre, C, of the circular path—in this case, vertically upwards. The person's centripetal acceleration, a_c, is:

$$\begin{aligned} a_c &= \frac{v^2}{r} \\ &= \frac{8.0^2}{2.5} \\ &= 26\,\mathrm{m\,s^{-2}} \text{ towards C, or upwards} \end{aligned}$$

The net (centripetal) force acting on the person is given by:

$$\begin{aligned} \vec{F}_{net} &= m\vec{a} \\ &= 50 \times 26 \\ &= 1300\,\mathrm{N} \text{ upwards} \end{aligned}$$

The normal force and the weight force are not in balance anymore. They add together to give an upward force of 1300 N. This indicates that the normal force must be greater than the weight force by 1300 N. In other words, the normal force is $\vec{F}_N = \vec{F}_{net} - \vec{F}_g = 1300 - (-490) = 1790\,\mathrm{N}$ up. This is over three times larger than the normal force of 490 N that usually acts. That is the reason why, when in a ride, you feel the seat pushing up against you much more strongly at this point. The normal force of 1790 N in this instance is equal to the **apparent weight** of the person and indicates they would feel much heavier than usual.

CIRCULAR MOTION: TRAVELLING OVER HUMPS

Now consider the situation as the cart moves over the top of a hump of radius 2.5 m with a lower speed of $2.0\,\mathrm{m\,s^{-1}}$, as illustrated in Figure 3.3.3.

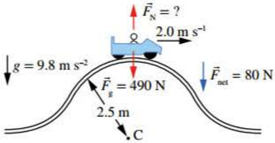


FIGURE 3.3.3 The centripetal acceleration is downwards towards the centre of the circle, and so the net force is also in that direction. At this point, the magnitude of the normal force, $\vec{F}_N$, is less than the weight, $\vec{F}_g$, of the person.

The person now has a centripetal acceleration that is directed vertically downwards towards the centre, C, of the circle. Therefore, the net force acting at this point is directed vertically downwards. The centripetal acceleration is:

$$\begin{aligned} a_c &= \frac{v^2}{r} \\ &= \frac{2.0^2}{2.5} \\ &= 1.6\,\mathrm{m\,s^{-2}} \text{ towards C, or downwards} \end{aligned}$$

The net (centripetal) force is:

$$\begin{aligned} \vec{F}_{net} &= m\vec{a} \\ &= 50 \times 1.6 \\ &= 80\,\mathrm{N} \text{ downwards} \end{aligned}$$

As in the dip, the weight force and the normal force are not in balance. They add to give a net force of 80 N down. The weight force must therefore be 80 N greater than the normal force. This tells us that the normal force is:

$$\vec{F}_N = \vec{F}_{net} - \vec{F}_g = -80 - (-490) = 410\,\mathrm{N} \text{ up.}$$

How the normal force varies during the ride

It is interesting to compare the normal forces that act on the person in these three situations.

- The normal force when travelling horizontally is 490 N upwards.
- At the bottom of the dip, the normal force is 1790 N upwards. In other words, in the dip, the seat pushes into the person with a greater force than usual. This gives the person an apparent weight of 1790 N and makes the person feel much heavier than normal. If the person had been sitting on weighing scales at this time, it would have shown a higher than usual reading.
- At the top of the hump, the normal force is 410 N upwards. In other words, over the hump, the seat pushes into the person with a smaller force than usual. This gives the person an apparent weight of 410 N and gives them the sensation of feeling lighter.

The weight of the person has not changed. $\vec{F}_g$ is 490 N throughout the duration of the ride; it is the normal force acting on them that varies. The normal force is equal to the person's apparent weight, and this makes the person 'feel' heavier and lighter as they travel through the dips and humps respectively.

REVISION

Energy

Kinetic energy is energy associated with motion. Any moving object will have kinetic energy. It is a scalar variable and can be calculated with the formula:

$K = \frac{1}{2}mv^2$

where

K is the kinetic energy (J)

m is the mass (kg)

v is the speed ($\mathrm{m\,s^{-1}}$)

The **potential energy** of an object is associated with its position relative to another object or within a field. For example, an object suspended by a crane has gravitational potential energy because of its position in the Earth's gravitational field. The gravitational potential energy can be calculated with the formula:

$\Delta U = mg\Delta h$

where

ΔU is the change in gravitational potential energy (J)

m is the mass (kg)

g is the acceleration due to gravity ($\mathrm{m\,s^{-2}}$)

Δh is the change in height of the object (m)

CONSERVATION OF MECHANICAL ENERGY

The sum of the potential and kinetic energy of an object is its **mechanical energy**, and this is constant unless work is done by an external force. The total mechanical energy is said to be **conserved** if the final energy is equal to the initial energy. There is frequently a transformation of energy between potential and kinetic energy. A child dropping from the branch of a tree onto a trampoline loses gravitational potential energy, but gains kinetic energy. On striking the trampoline, kinetic energy is transformed to strain potential energy in the springs, and in an ideal case would be returned as kinetic energy on the rebound.

Sometimes energy is dissipated, or transformed into heat, light and/or sound, and thus the energy remaining in the system is reduced, although the conservation of energy still applies in these situations. Spacecraft have to dissipate huge amounts of kinetic and gravitational potential energy as they re-enter the Earth's atmosphere and slow down to make a landing. Meteors, or so-called shooting stars, burn up in the upper atmosphere because of the heat generated by friction.

Principle of conservation of mechanical energy

Given that, in a system of bodies, there are no other forms of energy except kinetic energy and potential energy, then the total mechanical energy of the system is constant.

$$(E_m)_{initial} = (E_m)_{final}$$

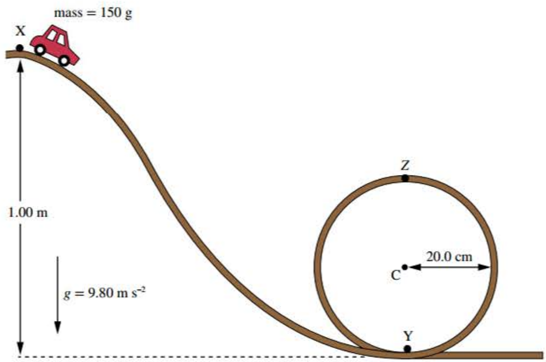
$$K_{initial} + U_{initial} = K_{final} + U_{final}$$

During vertical circular motion, energy is transformed from kinetic energy to gravitational potential as the object goes up, and from gravitational potential to kinetic energy as it goes down. The principle of conservation of energy must always hold.

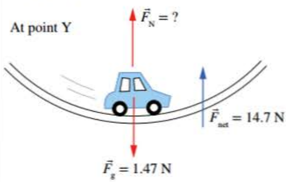
Worked example 3.3.1

VERTICAL CIRCULAR MOTION

A student arranges a toy car track with a vertical loop of radius 20.0 cm, as shown. A toy car of mass 150 g is released from rest at a height of 1.00 m at point X. The car rolls down the track and travels inside the loop. Assume g is $9.80\ m\ s^{-2}$, and ignore friction.



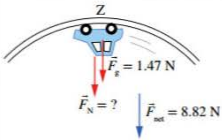
a Calculate the speed of the car as it reaches the bottom of the loop, point Y.	
Thinking	**Working**
Note all the variables given to you in the question.	At X: $m = 150\,g = 0.150\,kg$ $h = 1.00\,m$ $v = 0$ $g = 9.80\,m\,s^{-2}$
Use an energy approach to calculate the speed. Calculate the total mechanical energy first.	The initial speed is zero, so K at X is zero. The total mechanical energy, E_m, at X is: $E_m = K + U$ $= \frac{1}{2}mv^2 + mg\Delta h$ $= 0 + (0.150 \times 9.80 \times 1.00)$ $= 1.47\,J$
Use conservation of energy ($E_m = K + U$) to determine the velocity at point Y. As the car rolls down the track, it loses its gravitational potential energy and gains kinetic energy. At the bottom of the loop (Y), the car has zero potential energy.	At Y: $E_m = 1.47\,J$ $h = 0$ $U_g = 0$ $E_m = K + U$ $E_m = \frac{1}{2}mv^2 + mg\Delta h$ $1.47 = 0.5 \times 0.150\,v^2 + 0$ $v^2 = \frac{1.47}{0.0750}$ $v = \sqrt{19.6}$ $= 4.43\,m\,s^{-1}$

b Calculate the normal reaction force from the track at point Y.	
Thinking	**Working**
To solve for $\vec{F}_N$, start by working out the net, or centripetal, force. At Y, the car has a centripetal acceleration towards C (i.e. upwards), so the net (centripetal) force must also be vertically upwards at this point.	$F_{net} = F_c = \frac{mv^2}{r}$ $= \frac{0.150 \times 4.43^2}{0.200}$ $= 14.7\,N$ up
Calculate the weight force, $\vec{F}_g$, and add it to a force diagram.	$\vec{F}_g = m\vec{g}$ $= 0.150 \times 9.80$ $= 1.47\,N$ down 

Work out the normal force using vectors. Note up as positive and down as negative for your calculations. The forces acting are unbalanced, as the car has a centripetal acceleration upwards (towards C). The upward (normal) force must be larger than the downward force.	$\vec{F}_{net} = \vec{F}_g + \vec{F}_N$ $+14.7 = -1.47 + \vec{F}_N$ $\vec{F}_N = +14.7 + 1.47$ $= 16.2\,\text{N up}$ Note that the force the track exerts on the car is much greater (by about ten times) than the weight force. If the car were travelling horizontally, the normal force would be just 1.47 N up.

c What is the speed of the car as it reaches point Z?	
Thinking	**Working**
Calculate the speed from the values you have, using $E_m = K + U$.	At Z: $m = 0.150\,\text{kg}$ $\Delta h = 2 \times 0.200 = 0.400\,\text{m}$ Mechanical energy is conserved, so $E_m = 1.47\,\text{J}$ $E_m = K + U$ $= \frac{1}{2}mv^2 + mg\Delta h$ $1.47 = \frac{1}{2} \times 0.150 \times v^2 + 0.150 \times 9.80 \times 0.400$ $1.47 = 0.075v^2 + 0.588$ $0.075v^2 = 1.47 - 0.588$ $v^2 = 11.76$ $v = \sqrt{11.76}$ $= 3.43\,\text{m}\,\text{s}^{-1}$

d What is the normal force acting on the car at point Z?	
Thinking	**Working**
To find $\vec{F}_N$, start by working out the net, or centripetal, force. At Z, the car has a centripetal acceleration towards C (i.e. downwards), so the net (centripetal) force must also be vertically downwards at this point.	$F_{net} = F_c = \frac{mv^2}{r}$ $= \frac{0.150 \times 3.43^2}{0.200}$ $= 8.82\,\text{N down}$

<table>
<tr>
<td>Work out the normal force using vectors. Note up as positive and down as negative for your calculations.</td>
<td>

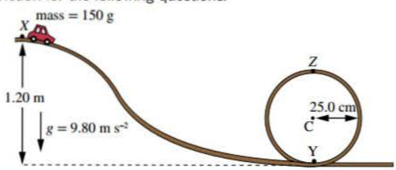
$\vec{F}_{net} = \vec{F}_g + \vec{F}_N$
$-8.82 = -1.47 + \vec{F}_N$
$\vec{F}_N = -8.82 + 1.47$
$= -7.35$
$= 7.35$ N down
Note that there is still strong contact between the car and the track from the normal force, but that is only around half the size compared to at the bottom of the track.
If the car had progressively lower speeds, the normal force at Z would decrease and would eventually drop to zero. At this point, the car would lose contact with the track and would fall with its acceleration equal to g.</td>
</tr>
</table>

Worked example: Try yourself 3.3.1

VERTICAL CIRCULAR MOTION

A student arranges a toy car track with a vertical loop of radius 25.0 cm, as shown.

A toy car of mass 150 g is released from rest at a height of 1.20 m at point X. The car rolls down the track and travels around the loop. Assume g is $9.80\,m\,s^{-2}$, and ignore friction for the following questions.



a Calculate the speed of the car as it reaches the bottom of the loop, point Y.
b Calculate the normal reaction force from the track at point Y.
c What is the speed of the car as it reaches point Z?
d What is the normal force acting on the car at point Z?

The symbol || signifies that the direction of the force is parallel to the displacement.

WORK

Work is the transfer of energy from one object to another and/or the transformation of energy from one form to another. A force does work on an object when it acts on a body causing a displacement in the direction of the force. Where the force is constant, the work done by the force is $W = F_{\parallel}s$.

If the force is applied at an angle to the displacement, only the component of the force in the direction of the displacement contributes to the work done. That is, if the force and displacement vectors are at an angle θ with respect to each other, then $F\cos\theta$ is the component of force that does work.

$W = F_{\parallel}s = Fs\cos\theta$

where

W is the work done by the force (J)

F is the magnitude of the constant force (N)

s is the displacement (m)

θ is the angle between the force vector and the displacement vector

While both force and displacement are vectors, work and energy are scalar quantities that are measured in joules (J).

To find the work done on an object, use the net force. For example, if a man pushes a heavy couch across a carpeted floor, the work done on the couch depends on the force applied by the man, less the frictional force which opposes the motion:

$$W = \Delta E = Fs\cos\theta$$

The energy, ΔE, gained by the couch depends on the net force acting on it.

When a force performs no work

It is important to remember that work is only done when a force, or a component of force, is applied in the direction of displacement. Hence it is possible to exert a force and feel very tired without doing work. This would mean no energy has been transferred. For example, if you hold a heavy object with your arms out in front of you, you will get tired very quickly but you are not doing any work on the object.

Similarly, an object moving in a circular path in a horizontal plane is constantly accelerated by the centripetal force. Because this force is perpendicular to the displacement at each instant, the force does no work, and no energy is transferred to the object. It does not get faster or slower, it only changes direction as shown in Figure 3.3.4.

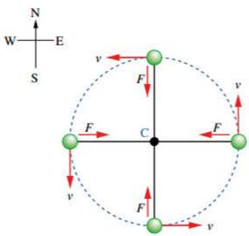


FIGURE 3.3.4 A body moving in a circular path has a force directed towards the centre. The displacement is in the direction of the velocity. There is therefore no force in the direction of the displacement.

Kinetic energy, K, is dependent on the magnitude of the velocity (or the speed) of an object. During uniform circular motion, the velocity is constantly changing but the magnitude of the velocity is constant. Therefore, as an object undergoes uniform circular motion, its kinetic energy is constant. Mechanical work is defined as the change in kinetic energy of an object. Objects undergoing uniform circular motion, with a constant kinetic energy, have no work being done to them despite the application of a force that produces movement.

3.3 Review

SUMMARY

- Kinetic energy is the energy of motion of a body: $K = \frac{1}{2}mv^2$.
- The sum of the kinetic and potential energy (total mechanical energy) of an isolated system is always conserved.
- Close to the surface of the Earth, where the force of gravity can be assumed to be constant, the change in gravitational potential energy of an object of mass m is $\Delta U = mg\Delta h$.
- Work, W, is a scalar and is measured in joules (J).
- A centripetal force does no work on an orbiting object, as the force and displacement are perpendicular.

KEY QUESTIONS

1 A horse pulls a carriage of mass 100 kg around a horizontal circular track at a constant speed of 4 m s^{-1}. The circular track has a radius of 15 m. What is the work done on the carriage as it moves from one side of the circular track to the other?

2 A ball of mass 0.20 kg rolls over the top of a crest at a velocity of 1.5 m s^{-1}. The crest has a radius of 2.0 m. What is the normal reaction force on the ball?

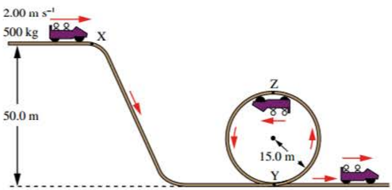
3 A yo-yo has a mass of 80 g and travels in a vertical circle with radius of 0.50 m. At the top of the circle the speed is 4.0 m s^{-1}.

a Calculate the tension in the string at the top of the circle.

b Calculate the speed of the yo-yo at the bottom of the circle.

c Calculate the tension in the string at the bottom of the circle.

4 A popular amusement park ride is the 'loop-the-loop', in which a cart descends a steep incline at point X, enters a circular loop at point Y, and makes one complete revolution of the circular loop. The car, whose total mass is 500 kg, carries the passengers with a speed of 2.00 m s^{-1} when it begins its descent at point X from a vertical height of 50.0 m.



a Calculate the speed of the car at point Y.

b What is the speed of the car at point Z?

c Calculate the normal force acting on the car at Z.

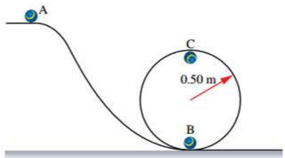
d What is the minimum speed that the car can have at point Z and still stay in contact with the track?

5 A skateboarder of mass 55 kg is practising on a half-pipe of radius 2.0 m. At the lowest point of the half-pipe, the speed of the skater is 6.0 m s^{-1}. Ignore air resistance and friction.

a What is the acceleration of the skater at this point? Indicate both magnitude and direction.

b Calculate the normal force acting on the skater at this point.

6 A ball bearing of mass 25 g is rolled along a smooth track in the shape of a loop-the-loop. The ball bearing is given a launch speed at A so that it just maintains contact with the track as it passes through point C. Ignore air resistance and friction.



a Determine the magnitude of the acceleration of the ball bearing as it passes point C.

b How fast is the ball bearing travelling at point C?

7 A bucket of mass 1 kg is on a 0.8 m long rope being swung in a vertical circle. What is the minimum speed to keep the bucket travelling on the circular path?

8 A pendulum with a mass of 0.15 kg on a string of length 0.4 m is held so the string is parallel to the ground. The pendulum is let go.

a What is the speed of the mass at the bottom of the swing?

b What is the maximum tension in the string?

3.4 Torque

Many situations involve objects that rotate about a **pivot point**, such as closing a door, using a spanner or turning a steering wheel. In these situations, a force acts to provide a turning effect or a **torque** (τ). Torque is a vector so it has a magnitude and a direction. Newton's laws use the concept of a force to help understand changes in the linear (straight line) motion of an object. The concept of torque is used in exactly the same way to explain a change in the rotational (turning) motion of an object.

TORQUE

Consider the steering wheel in Figure 3.4.1. When a turning effect is applied to the steering wheel there are a number of factors that must work together, causing it to turn. For all turning objects, there must be a pivot point around which the object will rotate. There must be a force applied to the object in such a way as to cause the object to rotate. This means that the force applied must not be aligned with the pivot point. There must be some distance between the **line of action of the force** (an imaginary line through the force vector) and the pivot point.

FIGURE 3.4.1 Applying a torque to a steering wheel will cause it to turn.

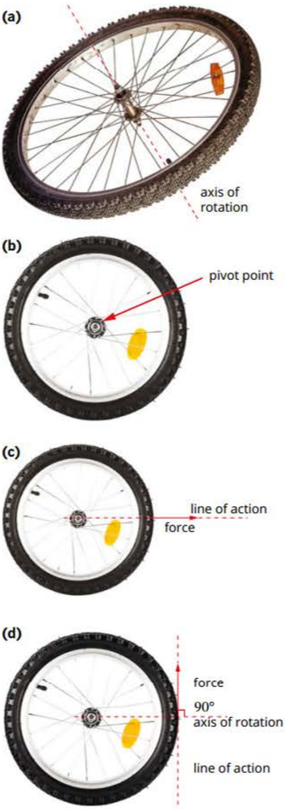


FIGURE 3.4.2 (a) The axis of rotation. (b) The pivot point. (c) When the line of action of the force passes through the pivot point, the wheel will not turn. (d) When the line of action of the force is at 90° to a line drawn from the pivot point to the point of application, the torque is at a maximum.

Force and the pivot point

When analysing a rotating system, the position of the pivot point or **axis of rotation** is an important consideration. A wheel, for example, moves in a circular path around its axle. An imaginary line along the length of the axle is called the axis of rotation and is shown in Figure 3.4.2a.

The pivot point is the point on a two-dimensional representation of the object through which the axis of rotation passes. As an example, the pivot point of a wheel is shown in Figure 3.4.2b.

A force applied directly towards or directly away from the pivot point of the wheel will not create a turning effect on the wheel. So, for the example in Figure 3.4.2c, if the force acted along the line labelled 'line of action', the wheel would not turn.

Torque can be achieved by applying a force on the wheel where the line of action of the force does not pass through the axis of rotation or the pivot point. The maximum effect is achieved when the force applied is at 90° to a line drawn from the pivot point to the point of application (the point at which the force is applied). This is shown in Figure 3.4.2d.

Magnitude of the force and torque

The torque (represented by τ, the Greek symbol tau) on an object is directly proportional to the magnitude of the force. If all other things are equal, a larger force will result in a larger torque. This is illustrated in Figure 3.4.3.

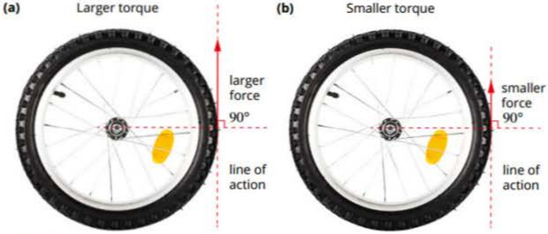


FIGURE 3.4.3 The magnitude of the force affects the torque on an object. The wheel in (a) will experience a larger torque than the wheel in (b).

Distance from the pivot point and torque

The amount of torque on an object is directly proportional to the perpendicular distance between the pivot point and the line of action of the force. This perpendicular distance is called the **force arm**. The force arm is given the symbol r and is shown in Figure 3.4.4. Assuming that everything else is constant, then the larger the force arm (r), the larger the torque (τ).

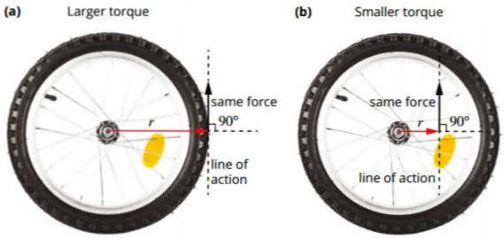


FIGURE 3.4.4 The perpendicular distance from the pivot point to the line of action of the force affects the torque on an object. The wheel in (a) will experience a larger torque than the wheel in (b).

It was stated previously that maximum torque occurs when the line of action of the force is perpendicular to a line drawn from the pivot point to the point of application. You can now use the concept of a force arm to understand this. The force arm is maximised when the line of action is perpendicular to the line between the pivot point and the point of application, and therefore the torque is maximised (Figure 3.4.5).

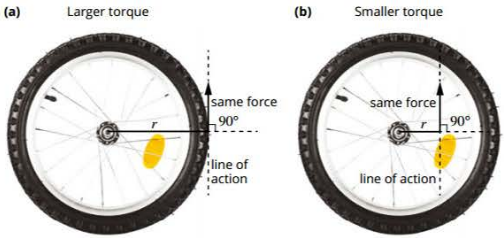


FIGURE 3.4.5 The direction of the applied force affects the size of the force arm, and therefore the torque on the object. The wheel in (a) will experience a larger torque than the wheel in (b).

The torque equation

The magnitude of the torque increases or decreases as the force increases or decreases. The magnitude of the torque also increases or decreases as the force arm or the perpendicular distance from the pivot to the line of action of the force ($r_\perp$) increases or decreases.

The formula for calculating torque is:

> $\tau = r_\perp F$
>
> where
>
> τ is the torque (N m)
>
> $r_\perp$ is the force arm perpendicular to the force (m)
>
> F is the force (N)

Torque is a vector quantity. This enables us to distinguish its unit from the joule, the scalar unit for work and energy, which can also be written with the unit N m. A rotating body rotates either clockwise or anticlockwise. When more than one torque is acting on a body and the net torque needs to be found, you will need to specify a direction convection. While this is often chosen to describe a clockwise rotation as negative and an anticlockwise rotation as positive, this is an arbitrary choice. As with any direction convention it is only important to be consistent with your approach.

Worked example 3.4.1

CALCULATING TORQUE

A bus driver applies a force of 45.0 N on the steering wheel of a bus as it turns a right-hand corner. The radius of the steering wheel is 30.0 cm. If the force is applied at 90° to the radius, calculate the torque on the steering wheel.

Thinking	Working
Identify the variables involved and state them in their standard form.	$\tau = ?$ $r_\perp = 0.300$ m $F = 45.0$ N
Apply the equation for torque. State the answer with the appropriate direction.	$\tau = r_\perp F$ $= 0.300 \times 45.0$ $= 13.5$ N m clockwise

Worked example: Try yourself 3.4.1

CALCULATING TORQUE

A force of 255 N is required to apply a torque on a sports car steering wheel as it turns left. The force is applied at 90° to the 15.5 cm radius of the steering wheel. Calculate the torque on the steering wheel.

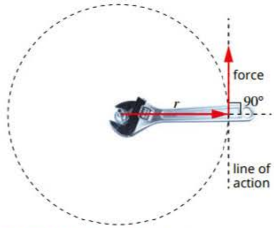


FIGURE 3.4.6 Although the adjustable spanner is not a wheel or circle, torque can still be applied to the nut.

Torque on different objects

Torque doesn't only act on circular objects. Any object can rotate about a point if a force is applied where the line of action of the force is not acting through the pivot point.

Spanners, like the one in Figure 3.4.6, apply a torque to a nut or bolt: the pivot point is the bolt and a force is applied at right angles to the spanner.

The reason a spanner is an effective hand tool is because it increases the force arm when turning a nut. If you try unscrewing a nut with your hands you will probably find that you are unable to provide enough force to create the torque required to turn the bolt. Longer spanners can apply a greater torque on a nut than shorter spanners. Some wheel-nut spanners, like the one in Figure 3.4.7, have handles which can extend so the force arm can be increased. This provides extra torque for loosening very tight nuts or for tightening the nuts with the correct torque.

FIGURE 3.4.7 Using an extended-handle spanner will increase the torque on the nut of this wheel.

Doors are also good examples of torque in action, with the hinges forming the axis of rotation. If force is applied to the handle and the line of action of the force is perpendicular to the door, then the distance between the hinge and the handle represents the force arm. This is shown in Figure 3.4.8.

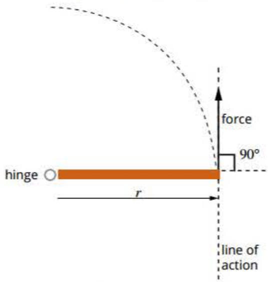


FIGURE 3.4.8 A door can have a torque applied to it, as long as the line of action of the force is not through the axis of rotation.

NON-PERPENDICULAR CALCULATIONS OF TORQUE

When the force causing a torque acts along a line that is at an angle other than 90° to an object (such as the door in Figure 3.4.9), then the torque is reduced. In these circumstances, we can calculate the torque by two approaches: either finding the component of the force acting perpendicular to the door, or by finding the perpendicular distance from the pivot point to the line of action of the force.

Recall that the formula for torque (τ) on an object is:

$$\tau = r_{\perp}F$$

This equation calculates the torque (τ) when the force (F) and the distance from the pivot to the line of action of the force (r) are perpendicular to each other. It really doesn't matter whether the radius is perpendicular to the line of action of the force, or if the force is perpendicular to the radius. The result is the same either way. That is, $\tau = r_{\perp}F$ and also $\tau = rF_{\perp}$.

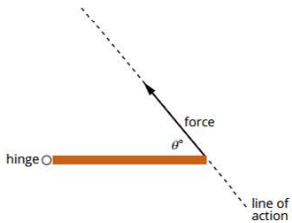


FIGURE 3.4.9 When the force causing a torque is not perpendicular to a door, the torque is reduced.

Calculating torque using perpendicular force

The components of any force can be calculated using trigonometry. To find the component of the force that is perpendicular to a door, for example, use the magnitude of the force and the angle between the door and the line of action of the force. This is shown in Figure 3.4.10.

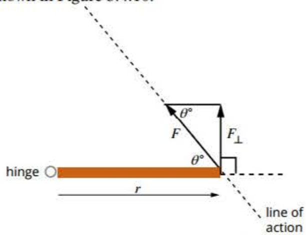


FIGURE 3.4.10 Finding the component of the force that is perpendicular to a door using the magnitude of the force and the angle between the door and the line of action of the force.

In this case, $\tau = rF_{\perp}$ and $F_{\perp} = F\sin\theta$.

This combines to give:

$\tau = r\,F\sin\theta$

where

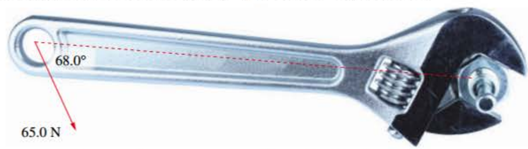
- τ is the torque (N m)
- r is the force arm (m)
- F is the applied force (N)
- θ is the angle between the applied force and the force arm

Your strategy for solving questions of this type may be to calculate the perpendicular component of the force and then apply the torque equation, or to use the combined equation. To begin with, it is recommended that you calculate the perpendicular component and then use the torque equation. When you have gained confidence with that strategy, try using the combined equation.

Worked example 3.4.2

CALCULATING TORQUE FROM THE PERPENDICULAR COMPONENT OF FORCE

A student uses a 42.0 cm long adjustable spanner to loosen a nut on her bike. She applies a force of 65.0 N at an angle of 68.0° to the spanner.



Calculate the anticlockwise torque that the student applies to the nut.

Thinking	Working
Use the trigonometric relationship $F_{\perp} = F \sin\theta$ to determine the force perpendicular to the spanner.	$F_{\perp} = F \sin\theta$ $= 65.0 \sin 68.0$ $= 60.3$ N perpendicular to the force arm
Convert the variables to their standard units.	$r = 42.0$ cm $= 0.420$ m
Apply the equation for torque: $\tau = r_{\perp}F = rF_{\perp}$ State the answer with the appropriate units and direction.	$\tau = rF_{\perp}$ $= 0.420 \times 60.3$ $= 25.3$ N m anticlockwise

Worked example: Try yourself 3.4.2

WS 5.7

CALCULATING TORQUE FROM THE PERPENDICULAR COMPONENT OF FORCE

A mechanic uses a 17.0 cm long spanner to tighten a nut on a winch. He applies a force of 104 N at an angle of 75.0° to the spanner.

Calculate the magnitude of the torque that the mechanic applies to the nut.

Calculating torque using perpendicular radius

The components of any distance can be calculated using either Pythagoras' theorem or trigonometry. To find the component of a length that is perpendicular to the line of action of the force acting on a door, construct a line from the pivot point to the line of action of the force so that it intersects the line of action at right angles. An example is shown in Figure 3.4.11.

In this case, $\tau = r_{\perp}F$ and $r_{\perp} = r\sin\theta$.

This combines to give:

$$\tau = r\sin\theta F$$

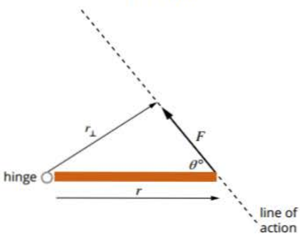


FIGURE 3.4.11 Determining the components of a distance.

The equation $\tau = rF\sin\theta$ is identical to $\tau = r\sin\theta F$, so either method would be appropriate for calculating the torque on an object when the force is not at right angles to the object. In either method, the component of the distance or the component of the force is always going to be less than the total distance or the total force itself. This will result in a smaller torque being applied to the object. The maximum torque will always be when the line of action of the force is perpendicular to the distance from the pivot point to the point of application.

PHYSICS IN ACTION ICT

The torque wrench

The extent to which a nut or bolt is tightened can be critical to the safe operation of machinery or motors. If a nut or bolt is too loose then it could fall out. If it is too tight then it could either distort the part or the bolt could break off. Both of these situations could require expensive repairs. To avoid nuts and bolts being too loose or too tight, manufacturers use different tools and methods to estimate the amount of torque required to tighten a nut or bolt to the correct tightness. Some examples of these tools are shown in Figure 3.4.12.

The beam wrench is the simplest type of torque wrench. It has a flexible lever arm with a bar and scale, separating the wrench head and handle. When torque is applied, a pointer on the scale moves to indicate the amount of torque being applied in newton metres (N m).

The click-type torque wrench can be set to apply a fixed amount of torque. When the required amount of torque has been achieved, the wrench 'clicks' and releases itself, preventing any further tightening from being applied.

More recently, electronic torque wrenches have been developed. The signal generated is converted to a torque reading (in N m) and is shown on the digital readout screen. Measurements can also be stored within the instrument's memory and transferred to a computer.

FIGURE 3.4.12 Three types of wrenches commonly used to measure the torque applied to a nut or bolt: (a) beam torque wrench, (b) click-type torque wrench, and (c) digital torque wrench.

3.4 Review

SUMMARY

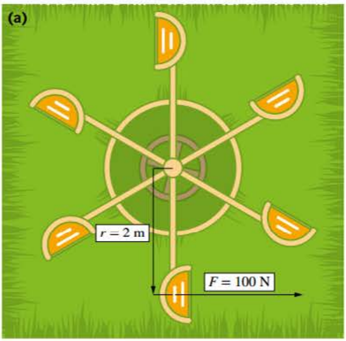
- Torque is a measurement of the tendency of a force to cause an object to rotate around an axis.
- The formula for calculating torque is $\tau = r_{\perp}F$.
- Torque occurs when the acting force is not applied directly through the pivot point of the object.
- Maximum torque occurs when the acting force applied is perpendicular to the force arm (r).
- The larger the force acting on the object, the larger the torque will be.
- The longer the force arm, the greater the torque will be.
- If torque is generated by an acting force that is not perpendicular to the force arm of the object, then either:
  - the component force perpendicular to the length of the object is used to calculate torque:
    $$\tau = rF_{\perp}$$
    or
  - the distance from the pivot point perpendicular to the line of action of the force is used to calculate torque: $\tau = r_{\perp}F$.
- Both strategies for determining torque from non-perpendicular situations equate to $\tau = rF\sin\theta$.

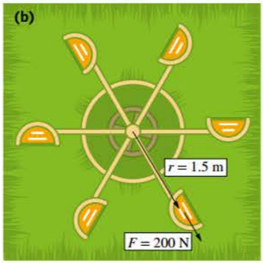
KEY QUESTIONS

Use $g = 9.80\,\text{m}\,\text{s}^{-2}$ to answer these questions.

1 Use the concept of torque to explain the following.
  a It is easier to open a heavy door by pushing it at the handle rather than in the middle of the door.
  b It is possible to move very heavy rocks in the garden by using a long crowbar.

2 Calculate the torque exerted on the roundabouts shown. Include the direction where appropriate.





3 The magnitude of the torque required to tighten a bicycle wheel is 15 N m. Calculate the force arm required if a 30 N force is applied perpendicularly.

4 A student pushes a heavy door at a point that is 50 cm from the hinges such that it creates a torque of 9 N m. With what magnitude of force does the student push? Assume the student pushes perpendicular to the surface of the door.

5 A spanner with a length of 40 cm is used to tighten a nut on a car wheel. If the magnitude of the perpendicular force applied is 225 N, calculate the maximum torque that can be created on this wheel nut.

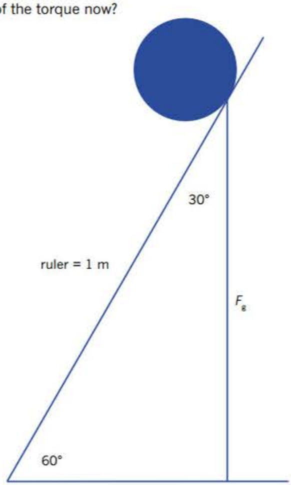
3.4 Review continued

6 Nikki is investigating torque using a metre ruler and a 1.0 kg mass. She uses a rubber band to attach the mass to the ruler at the 50 cm mark. Nikki first holds the ruler at one end so that it is horizontal.

a What is the size of the torque that is acting?

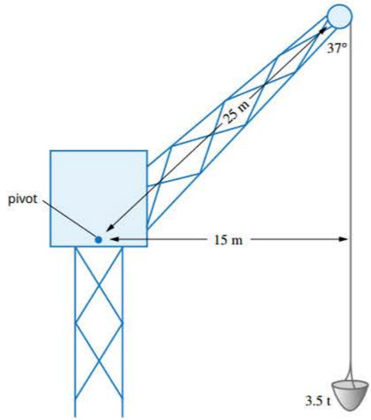
b She now moves the mass so that it is right at the far end of the ruler. How much torque is acting now?

c Finally, she lifts the ruler so that it makes an angle of 60° to the horizontal. What is the size of the torque now?



7 A mechanic uses a spanner of length 30 cm to tighten a bolt head. The mechanic applies a force of 300 N at an angle of 30° to the length of the spanner. Calculate the magnitude of the torque created.

8 A crane is being used to lift a skip of concrete with a total mass of 3.5 tonnes (3500 kg). The lever arm of the crane is 25 m long and makes an angle of 37° with the vertical, as shown in the diagram. Ignore the mass of the cable when answering these questions.



a What is the total weight of the skip?

b The skip is winched up so that it is near the top of the crane. Does the torque around the pivot created by this load increase, decrease or remain the same as the load is lifted?

c Calculate the magnitude and direction of the torque about the pivot that the skip exerts on the crane when the skip is at the highest point.

Chapter review

03

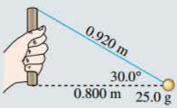
KEY TERMS

angular velocity
apparent weight
axis of rotation
banked track
centripetal acceleration
centripetal force
conserved
design speed
force arm
frequency
kinetic energy
line of action of the force
mechanical energy
Newton's first law
Newton's second law
Newton's third law
normal reaction force
period
pivot point
potential energy
radian
tangential
torque
work

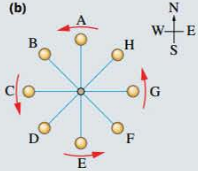
REVIEW QUESTIONS

The following information applies to questions 1 and 2.
During a high-school physics experiment, a copper ball of mass 25.0 g was attached to a very light piece of steel wire 0.920 m long and was whirled in a circle at 30.0° to the horizontal, as shown in diagram (a). The ball moves in a circular path of radius 0.800 m with a period of 1.36 s. The top view of the resulting motion of the ball is shown in diagram (b).

(a)



(b)



1 **a** Calculate the orbital speed of the ball.
b What is the centripetal acceleration of the ball?
c What is the magnitude of the centripetal force acting on the ball?

2 **a** Draw a diagram similar to diagram (a) that shows all the forces acting on the ball at this time.
b What is the magnitude of the tension in the wire?

3 A radio-controlled car is travelling in a circular path of radius 10 m at a constant speed of $5.0\,m\,s^{-1}$.
a What is the acceleration of the toy car?
b What force is creating the circular motion of the car?

4 The tip of a propeller on an airplane has a radius of 0.90 m and experiences a centripetal acceleration of $8.88 \times 10^4\,m\,s^{-2}$. What is the frequency of rotation of the propeller?

5 A cycling track has a turn that is banked at 40° to the horizontal. The radius of the track at this point is 30 m. Determine the speed at which a cyclist of mass 60 kg would experience no sideways force on their bike as they rode this section of track.

6 A cycling velodrome has a turn that is banked at 33° to the horizontal. The radius of the track at this point is 28 m.
a Determine the speed (in $km\,h^{-1}$) at which a cyclist of mass 55 kg would experience no sideways force on their bike as they rode this section of track.
b Calculate the size of the normal force that is acting on the cyclist.
c How does this compare with the normal force that would act on the cyclist if they were riding on a flat track?

7 A Totem Tennis ball has been hit and is travelling in a horizontal circle of radius 0.80 m. The ball has a speed of $5.0\,m\,s^{-1}$. What is the angular velocity of the ball?

8 An object that moves in uniform circular motion on a horizontal plane has a centripetal acceleration of $13\,m\,s^{-2}$. If the radius of motion is 0.020 m, what is the frequency of motion?

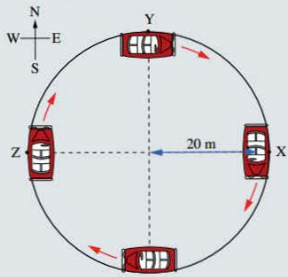
9 Fiona and Mark are flying their remote-controlled model plane. It has a mass of 1.6 kg and travels in a horizontal circular path of radius 62 m with a speed of $50\,km\,h^{-1}$. A radio transmitter controls the plane so there are no strings attached. Answer the following questions about the plane's motion.
a Calculate the period of the model plane.
b Determine the magnitude of the centripetal force that is acting on the plane.

10 An athlete competing at a junior sports meet swings a 2.5 kg hammer in a horizontal circle of radius 0.80 m at 2.0 revolutions per second. Assume that the wire is horizontal at all times.
a What is the period of rotation of the ball?
b What is the orbital speed of the ball?
c What is the magnitude of the acceleration of the ball?
d What is the magnitude of the net force acting on the ball?

11 For an object in uniform circular motion, what is the effect on the magnitude of the centripetal acceleration if:
- **a** the speed doubles
- **b** the radius triples
- **c** the mass halves?

12 A geostationary communications satellite is at an altitude of 3.60×10^4 m. The Earth has an average radius 6.37×10^6 m and a period of rotation of 23 hours, 56 minutes and 5 seconds. Calculate the centripetal acceleration of the satellite.

13 A car of mass 1500 kg is driven at a constant speed of $10\,m\,s^{-1}$ around a level, circular roundabout. The centre of mass of the car is always 20 m from the centre of the road.



- **a** What is the velocity of the car at point X?
- **b** What is the speed of the car at point Y?
- **c** What is the period of revolution for this car?
- **d** What is the acceleration of the car at point X?
- **e** Determine the size and direction of the unbalanced frictional force acting on the tyres at point X.

14 A proton moves into a region of uniform magnetic field 0.250 T directed perpendicular to the velocity vector. If it travels into the field at $3.50 \times 10^6\,m\,s^{-1}$, calculate the radius of curvature of its path. Note that $m_p = 1.67 \times 10^{-27}$ kg, $q = 1.60 \times 10^{-19}$ C and $F = qvB$.

15 A track cyclist is riding at high speed on the steeply banked section of a velodrome ($\theta = 37°$). Which statement describes the size of the normal force acting on the cyclist at this point?
- **A** greater than the weight of the cyclist
- **B** zero
- **C** less than the weight of the cyclist
- **D** equal to the weight of the cyclist

16 The Moon orbits the Earth once in 27.3 days in a circular orbit of radius 3.84×10^8 m.
- **a** Calculate the orbital speed of the Moon.
- **b** Calculate the angular velocity of the Moon.
- **c** Calculate the net force keeping the Moon in orbit if the mass of the Moon is 7.36×10^{22} kg.

17 A bicep muscle is able to produce 500 N of force. Given that the muscle attaches 4.0 cm from the elbow and the hand is 35 cm from the elbow, how heavy an object can be held?

18 A 50 kg student sits 1.25 m from the centre of a seesaw and a box is placed at the opposite end 1.5 m from the centre. What is the mass and weight of the box in order for this system to be in equilibrium?

19 A bolt requires 15 N m of torque to tighten it. If you are capable of producing 40 N of force, how long does the torque wrench need to be to provide the necessary torque to the bolt to tighten it?

20 After completing the activity on page 68, reflect on the inquiry question: Why do objects move in circles?

CHAPTER

04 Motion in gravitational fields

Gravity is, quite literally, the force that drives the universe. It was gravity that first caused particles to coalesce into atoms, and atoms to congregate into nebulas, planets and stars. An understanding of gravity is fundamental to understanding the universe.

This chapter centres on Newton's law of universal gravitation. This will be used to predict the size of the force experienced by an object at various locations on the Earth and other planets. It will also be used to develop the idea of a gravitational field. Since the field concept is also used to describe other basic forces such as electromagnetism and the strong and weak nuclear forces, this will provide an important foundation for further study in physics.

Content

INQUIRY QUESTION

How does the force of gravity determine the motion of planets and satellites?

By the end of this chapter you will be able to:

- apply qualitatively and quantitatively Newton's law of universal gravitation to:
  - determine the force of gravity between two objects $F = \frac{GMm}{r^2}$
  - investigate the factors that affect the gravitational field strength $g = \frac{GM}{r^2}$
  - predict the gravitational field strength at any point in a gravitational field, including at the surface of a planet (ACSPH094, ACSPH095, ACSPH097)
- investigate the orbital motion of planets and artificial satellites when applying the relationships between the following quantities: CCT ICT N
  - gravitational force
  - centripetal force
  - centripetal acceleration
  - mass
  - orbital radius
  - orbital velocity
  - orbital period
- predict quantitatively the orbital properties of planets and satellites in a variety of situations, including near the Earth and geostationary orbits, and relate these to their uses (ACSPH101) ICT N
- investigate the relationship of Kepler's laws of planetary motion to the forces acting on, and the total energy of, planets in circular and non-circular orbits using: (ACSPH101)
  - $v = \frac{2\pi r}{T}$
  - $\frac{r^3}{T^2} = \frac{GM}{4\pi^2}$ ICT N
- derive quantitatively and apply the concepts of gravitational force and gravitational potential energy in radial gravitational fields to a variety of situations, including but not limited to: ICT N
  - the concept of escape velocity $v_{esc} = \sqrt{\frac{2GM}{r}}$
  - total potential energy of a planet or satellite in its orbit $U = -\frac{GMm}{r}$
  - total energy of a planet or satellite in its orbit $U + K = -\frac{GMm}{2r}$
  - energy changes that occur when satellites move between orbits (ACSPH096)
  - Kepler's laws of planetary motion (ACSPH101).

4.1 Gravity

FIGURE 4.1.1 Sir Isaac Newton was one of the most influential physicists who ever lived.

In 1687, Sir Isaac Newton (Figure 4.1.1) published a book that changed the world. Entitled *Philosophiæ Naturalis Principia Mathematica* (Mathematical Principles of Natural Philosophy), Newton's book (shown in Figure 4.1.2) used a new form of mathematics now known as calculus and outlined his famous laws of motion.

The *Principia* also introduced Newton's law of universal gravitation. This was particularly significant because, for the first time in history, it scientifically explained the motion of the planets. This led to a change in humanity's understanding of its place in the universe.

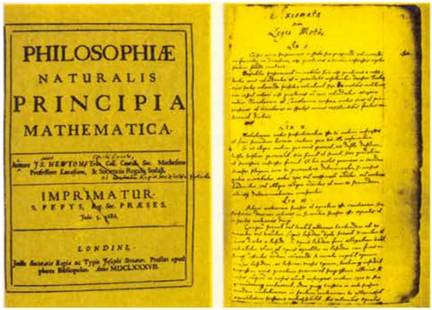


FIGURE 4.1.2 The *Principia* is one of the most influential books in the history of science.

NEWTON'S LAW OF UNIVERSAL GRAVITATION

Newton's law of universal gravitation states that any two bodies in the universe attract each other with a force that is directly proportional to the product of their masses and inversely proportional to the square of the distance between them.

Mathematically, Newton's law of universal gravitation can be expressed as:

$$F = \frac{GMm}{r^2}$$

where

F is the gravitational force (N)

M is the mass of object 1 (kg)

m is the mass of object 2 (kg)

r is the distance between the centres of objects 1 and 2 (m)

G is the gravitational constant, $6.67 \times 10^{-11}\ \text{N m}^2\ \text{kg}^{-2}$

The gravitational force is always an attractive force. The **gravitational constant**, G, was first accurately measured by the British scientist Henry Cavendish in 1798, over a century after Newton's death. As its name suggests, the law of universal gravitation predicts that any two objects that have mass will attract each other. However, because the value of G is so small, the **gravitational force** between two everyday objects is too small to be noticed.

The fact that r appears in the denominator of Newton's law of universal gravitation indicates an inverse relationship. Since r is also squared, this relationship is known as an **inverse square law**. The implication is that as r increases, F will decrease dramatically. This law will reappear again later in the chapter when gravitational fields are examined in detail.

Worked example 4.1.1

GRAVITATIONAL ATTRACTION BETWEEN SMALL OBJECTS

A man with a mass of 90 kg and a woman with a mass of 75 kg have a distance of 80 cm between their centres. Calculate the force of gravitational attraction between them.

Thinking	Working
Recall the formula for Newton's law of universal gravitation.	$F = \frac{GMm}{r^2}$
Identify the information required, and convert values into appropriate units when necessary.	$G = 6.67 \times 10^{-11}\,N\,m^2\,kg^{-2}$ $M = 90\,kg$ $m = 75\,kg$ $r = 80\,cm = 0.80\,m$ between the man and the woman
Substitute the values into the equation.	$F = 6.67 \times 10^{-11} \times \frac{90 \times 75}{0.80^2}$
Solve the equation.	$F = 7.0 \times 10^{-7}\,N$ towards one another.

Worked example: Try yourself 4.1.1

GRAVITATIONAL ATTRACTION BETWEEN SMALL OBJECTS

Two bowling balls are sitting next to each other on a shelf so that the centres of the balls are 60 cm apart. Ball 1 has a mass of 7.0 kg and ball 2 has a mass of 5.5 kg. Calculate the force of gravitational attraction between them.

For the gravitational force to become significant, at least one of the objects must have a very large mass—for example, a planet (Figure 4.1.3).

FIGURE 4.1.3 Gravitational forces become significant when at least one of the objects has a large mass; for example, the Earth and the Moon.

Worked example 4.1.2

GRAVITATIONAL ATTRACTION BETWEEN MASSIVE OBJECTS

Calculate the force of gravitational attraction between the Sun and the Earth given the following data:

$m_{Sun} = 2.0 \times 10^{30}$ kg

$m_{Earth} = 6.0 \times 10^{24}$ kg

$r_{Sun-Earth} = 1.5 \times 10^{11}$ m

Thinking	Working
Recall the formula for Newton's law of universal gravitation.	$F = \frac{GMm}{r^2}$
Identify the information required.	$G = 6.67 \times 10^{-11}\,N\,m^2\,kg^{-2}$ $M = 2.0 \times 10^{30}$ kg $m = 6.0 \times 10^{24}$ kg $r = 1.5 \times 10^{11}$ m between the Sun and the Earth
Substitute the values into the equation.	$F = 6.67 \times 10^{-11} \times \frac{2.0 \times 10^{30} \times 6.0 \times 10^{24}}{(1.5 \times 10^{11})^2}$
Solve the equation.	$F = 3.6 \times 10^{22}$ N between the Sun and the Earth.

Worked example: Try yourself 4.1.2

GRAVITATIONAL ATTRACTION BETWEEN MASSIVE OBJECTS

Calculate the force of gravitational attraction between the Earth and the Moon, given the following data:

$m_{Earth} = 6.0 \times 10^{24}$ kg

$m_{Moon} = 7.3 \times 10^{22}$ kg

$r_{Moon-Earth} = 3.8 \times 10^{8}$ m

The forces in Worked example 4.1.2 are much greater than those in Worked example 4.1.1, illustrating the difference in the gravitational force when at least one of the objects has a very large mass.

+ ADDITIONAL N

Multi-body systems

So far, only gravitational systems involving two objects have been considered, such as the Moon and the Earth. In reality, objects experience gravitational force from every other object around them. Usually, most of these forces are negligible and only the gravitational effect of the largest object nearby (i.e. the Earth) needs to be considered.

When there is more than one significant gravitational force acting on a body, the gravitational forces must be added together as vectors to determine the net gravitational force (Figure 4.1.4). The direction and relative magnitude of the net gravitational force in a multi-body system depends entirely on the masses and positions of the attracting objects (i.e. m_1, m_2 and m_3 in Figure 4.1.4).

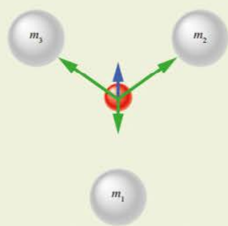


FIGURE 4.1.4 For the three masses $m_1 = m_2 = m_3$, the gravitational forces acting on the central red ball are shown by the green arrows. The vector sum of the green arrows is shown by the blue arrow. This will be the direction of the net (or resultant) gravitational force on the red ball due to the other three masses.

EFFECT OF GRAVITY

According to Newton's third law of motion, forces occur in action–reaction pairs. An example of such a pair is shown in Figure 4.1.5. The Earth exerts a gravitational force on the Moon and, conversely, the Moon exerts an equal and opposite force on the Earth. Using Newton's second law of motion, you can see that the effect of the gravitational force of the Moon on the Earth will be much smaller than the corresponding effect of the Earth on the Moon. This is because of the Earth's larger mass.

FIGURE 4.1.5 The Earth and Moon exert gravitational forces on each other.

> Remember that the gravitational field strength (units $N\,kg^{-1}$) is equivalent to the acceleration due to gravity, which has units of $m\,s^{-2}$. This is explained in further detail on page 110.

Worked example 4.1.3

ACCELERATION CAUSED BY A GRAVITATIONAL FORCE

The force of gravitational attraction between the Moon and the Earth is approximately 2.0×10^{20} N. Calculate the accelerations of the Earth and the Moon caused by this attraction. Compare these accelerations by calculating the ratio $\frac{a_{Moon}}{a_{Earth}}$.

Use the following data:

$m_{Earth} = 6.0 \times 10^{24}$ kg

$m_{Moon} = 7.3 \times 10^{22}$ kg

Thinking	Working
Recall the formula for Newton's second law of motion.	$\vec{F}_{net} = m\vec{a}$
Transpose the equation to make $\vec{a}$ the subject.	$\vec{a} = \frac{\vec{F}_{net}}{m}$
Substitute values into this equation to find the accelerations of the Moon and the Earth. The ratio uses the scalar acceleration so no direction is needed.	$a_{Earth} = \frac{2.0 \times 10^{20}}{6.0 \times 10^{24}} = 3.3 \times 10^{-5}\,N\,kg^{-1}$ $a_{Moon} = \frac{2.0 \times 10^{20}}{7.3 \times 10^{22}} = 2.7 \times 10^{-3}\,N\,kg^{-1}$
Compare the two accelerations.	$\frac{a_{Moon}}{a_{Earth}} = \frac{2.7 \times 10^{-3}}{3.3 \times 10^{-5}} = 82$ The acceleration of the Moon is 82 times greater than the acceleration of the Earth.

Worked example: Try yourself 4.1.3

ACCELERATION CAUSED BY A GRAVITATIONAL FORCE

The force of gravitational attraction between the Sun and the Earth is approximately 3.6×10^{22} N. Calculate the accelerations of the Earth and the Sun caused by this attraction. Compare these accelerations by calculating the ratio $\frac{a_{Earth}}{a_{Sun}}$.

Use the following data:

$m_{Earth} = 6.0 \times 10^{24}$ kg

$m_{Sun} = 2.0 \times 10^{30}$ kg

Gravity in the solar system

Although the accelerations caused by gravitational forces in Worked example 4.1.3 are small, over billions of years they created the motion of the solar system.

In the Earth–Moon system, the acceleration of the Moon is many times greater than that of the Earth, which is why the Moon orbits the Earth. Although the Moon's gravitational force causes a much smaller acceleration of the Earth, it does have other significant effects, such as the tides.

Similarly, the Earth and other planets orbit the Sun because their masses are much smaller than the Sun's mass. However, the combined gravitational effect of the planets of the solar system (and Jupiter in particular) causes the Sun to wobble slightly as the planets orbit it.

PHYSICSFILE N

Extrasolar planets

In recent years, scientists have been interested in discovering whether other stars have planets like those in our own solar system. One of the ways in which these 'extrasolar planets' (or 'exoplanets') can be detected is from their gravitational effect.

When a large planet (i.e. Jupiter-sized or larger) orbits a star, it causes the star to wobble. This causes variations in the star's appearance, which can be detected on Earth. Hundreds of exoplanets have been discovered using this technique.


FIGURE 4.1.6 The solar system is a complex gravitational system.

GRAVITATIONAL FIELDS

In the 18th century, to simplify the process of calculating the effect of simultaneous gravitational forces, scientists developed a mental construct known as the gravitational field. In the following centuries, the idea of a **field** was also applied to other forces and has become a very important concept in physics.

A **gravitational field** is a region in which a gravitational force is exerted on all matter within that region. Complex systems like the solar system involve a number of objects (i.e. the Sun and planets shown in Figure 4.1.6) that are all exerting attractive forces on each other at the same time. Every physical object has an accompanying gravitational field. For example, the space around your body contains a gravitational field because any other object that comes into this region will experience a (small) force of gravitational attraction towards your body.

The gravitational field around a large object like a planet is much more significant than that around a small object. The Earth's gravitational field exerts a significant influence on objects on its surface and even up to thousands of kilometres into space.

Representing gravitational fields

Over time, scientists have developed a commonly understood method of representing fields using a series of arrows known as field lines (Figure 4.1.7). For gravitational fields, these are constructed as follows:

- the direction of the arrowhead indicates the direction of the gravitational force
- the space between the arrows indicates the relative magnitude of the field:
  - closely spaced arrows indicate a strong field
  - widely spaced arrows indicate a weaker field
  - parallel field lines indicate constant or **uniform** field strength.

An infinite number of field lines could be drawn, so only a few are chosen to represent the rest. The size of the gravitational force acting on a mass in the region of a gravitational field is determined by the strength of the field, and the force acts in the direction of the field.

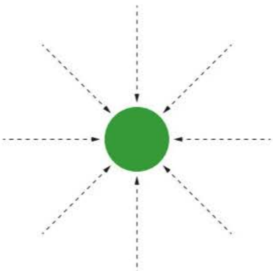


FIGURE 4.1.7 The arrows in this gravitational field diagram indicate that objects will be attracted towards the mass in the centre; the spacing of the lines shows that force will be strongest at the surface of the central mass and weaker further away from it.

GRAVITATIONAL FIELD STRENGTH

In theory, gravitational fields extend infinitely out into space. However, since the magnitude of the gravitational force decreases with the square of distance, eventually these fields become so weak as to become negligible.

The constant for the **acceleration due to gravity** g can be derived directly from the dimensions of the Earth. An object with mass m sitting on the surface of the Earth is a distance of 6.4×10^6 m from the centre of the Earth.

Given that the Earth has a mass of 6.0×10^{24} kg, then:

$$\text{weight} = \text{gravitational force}$$

$$\therefore \quad mg = G\frac{M_{\text{Earth}}m}{(r_{\text{Earth}})^2}$$

$$= mG\frac{M_{\text{Earth}}}{(r_{\text{Earth}})^2}$$

$$\therefore \quad g = G\frac{M_{\text{Earth}}}{(r_{\text{Earth}})^2}$$

$$= 6.67 \times 10^{-11} \times \frac{6.0 \times 10^{24}}{(6.4 \times 10^6)^2}$$

$$= 9.8\,\text{m s}^{-2} \text{ towards the centre of the Earth}$$

So, the rate of acceleration of objects near the surface of the Earth is a result of the Earth's mass and radius. A planet with a different mass and/or different radius will therefore have a different value for g.

The variable g can also be used as a measure of the strength of the gravitational field. When understood in this way, it is written with the equivalent units of N kg^{-1} rather than m s^{-2}. This means $g_{\text{Earth}} = 9.8\,\text{N kg}^{-1}$.

PHYSICSFILE N

$\text{N kg}^{-1} = \text{m s}^{-2}$

It is a simple matter to show that N kg^{-1} and m s^{-2} are equivalent units.

From Newton's second law, $\vec{F}_{\text{net}} = m\vec{a}$, you will remember that:

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

$$\therefore \quad 1\,\text{N kg}^{-1} = 1\,\text{kg m s}^{-2} \times \text{kg}^{-1} = 1\,\text{m s}^{-2}$$

These units indicate that an object on the surface of the Earth experiences 9.8 N of gravitational force for every kilogram of its mass.

Accordingly, the familiar equation $F_g = mg$ can be transposed so that the **gravitational field strength** g can be calculated: $g = \frac{F_g}{m} = \frac{GMm}{r^2} \times \frac{1}{m}$

$g = \frac{GM}{r^2}$

where

g is gravitational field strength ($N\,kg^{-1}$)

M is the mass of an object in the field (kg)

G is the gravitational constant, $6.67 \times 10^{-11}\,N\,m^2\,kg^{-2}$

r is the distance from the centre of M (m)

The gravitational field strength represents a vector quantity where its direction is always pointing towards the centre of the mass, i.e. downwards on Earth. Once you have chosen a specific direction convention, g may be represented as either negative or positive.

If the surface of the Earth is considered a flat surface as it appears in everyday life, then the gravitational field lines are approximately parallel, indicating a uniform field (Figure 4.1.8).

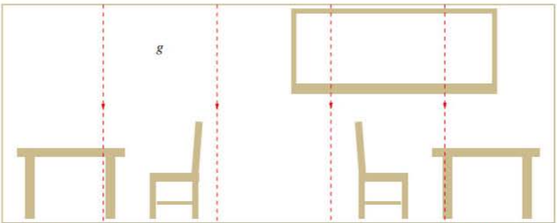


FIGURE 4.1.8 The uniform gravitational field, g, is represented by evenly spaced parallel lines in the direction of the force.

However, when the Earth is viewed from a distance as a sphere, it becomes clear that the Earth's gravitational field is not uniform at all (Figure 4.1.9). The increasing distance between the field lines indicates that the field becomes progressively weaker out into space.

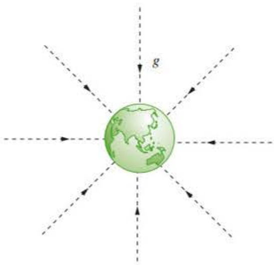


FIGURE 4.1.9 The Earth's gravitational field becomes progressively weaker out into space.

This is because gravitational field strength, like gravitational force, is governed by the inverse square law given above.

The gravitational field strength at different **altitudes** can be calculated by adding the altitude to the radius of the Earth to calculate the distance of the object from the Earth's centre (Figure 4.1.10).

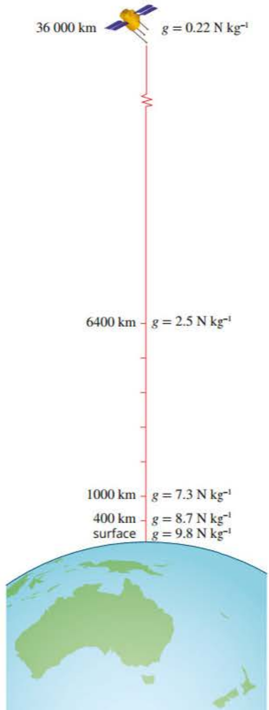


FIGURE 4.1.10 The Earth's gravitational field strength is weaker at higher altitudes.

The strength of the Earth's gravitational field at a particular altitude is given by:

$g = \frac{GM_{Earth}}{(r_{Earth} + altitude)^2}$

PHYSICS IN ACTION N

Variations in gravitational field strength of the Earth

The gravitational field strength of the Earth g is usually assigned a value of $9.81\,N\,kg^{-1}$. However, the field strength experienced by objects on the surface of the Earth can actually vary between $9.76\,N\,kg^{-1}$ and $9.83\,N\,kg^{-1}$, depending on the location.

The Earth's gravitational field strength is not the same at every point on the Earth's surface. As the Earth is not a perfect sphere (Figure 4.1.11), objects near the equator are slightly further from the centre of the Earth than objects at the poles. This means that the Earth's gravitational field is slightly stronger at the poles than at the equator.

Geological formations can also create differences in gravitational field strength, depending on their composition. Geologists use a sensitive instrument known as a **gravimeter** that detects small local variations in gravitational field strength to indicate underground features. Rocks with above-average density, such as those containing mineral ores, create slightly stronger gravitational fields, whereas less-dense sedimentary rocks produce weaker fields.

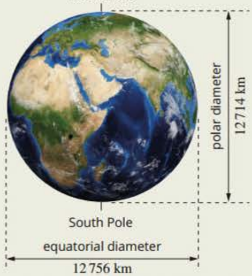


FIGURE 4.1.11 The Earth is a flattened sphere, which means its gravitational field is slightly stronger at the poles.

Worked example 4.1.4

CALCULATING GRAVITATIONAL FIELD STRENGTH AT DIFFERENT ALTITUDES

Calculate the strength of the Earth's gravitational field at the top of Mt Everest using the following data:

$r_{Earth} = 6.38 \times 10^6\,m$

$m_{Earth} = 5.97 \times 10^{24}\,kg$

height of Mt Everest = 8850 m

Thinking	Working
Recall the formula for gravitational field strength.	$g = \frac{GM}{r^2}$
Add the height of Mt Everest to the radius of the Earth.	$r = 6.38 \times 10^6 + 8850\,m$ $= 6.389 \times 10^6\,m$
Substitute the values into the formula.	$g = \frac{GM}{r^2}$ $= 6.67 \times 10^{-11} \times \frac{5.97 \times 10^{24}}{(6.389 \times 10^6)^2}$ $= 9.76\,N\,kg^{-1}$ towards the centre of the Earth

Worked example: Try yourself 4.1.4

CALCULATING GRAVITATIONAL FIELD STRENGTH AT DIFFERENT ALTITUDES

Commercial airlines typically fly at an altitude of 11 000 m. Calculate the gravitational field strength of the Earth at this height using the following data:

$r_{Earth} = 6.38 \times 10^6\,m$

$m_{Earth} = 5.97 \times 10^{24}\,kg$

GRAVITATIONAL FIELD STRENGTHS OF OTHER PLANETS

The gravitational field strength on the surface of the Moon is much less than on Earth, at approximately 1.6 $N\,kg^{-1}$. This is because the Moon's mass is smaller than the Earth's (Figure 4.1.12).

The formula $g = \frac{GM}{r^2}$ can be used to calculate the gravitational field strength on the surface of any astronomical object.

FIGURE 4.1.12 The gravitational field strength on the surface of the Moon is different to the gravitational field strength on the surface of the Earth.

Worked example 4.1.5

GRAVITATIONAL FIELD STRENGTH ON ANOTHER PLANET OR MOON

Calculate the strength of the gravitational field on the surface of the Moon given that the Moon's mass is 7.35×10^{22} kg and its radius is 1740 km.
Give your answer correct to three significant figures.

Thinking	Working
Recall the formula for gravitational field strength.	$g = \frac{GM}{r^2}$
Convert the Moon's radius to m.	$r = 1740$ km $= 1740 \times 1000$ m $= 1.74 \times 10^6$ m
Substitute values into the formula.	$g = \frac{GM}{r^2}$ $= 6.67 \times 10^{-11} \times \frac{7.35 \times 10^{22}}{(1.74 \times 10^6)^2}$ $= 1.62\ N\,kg^{-1}$ towards the centre of the Moon

Worked example: Try yourself 4.1.5

GRAVITATIONAL FIELD STRENGTH ON ANOTHER PLANET OR MOON

Calculate the strength of the gravitational field on the surface of Mars.
$m_{Mars} = 6.42 \times 10^{23}$ kg
$r_{Mars} = 3390$ km
Give your answer correct to three significant figures.

4.1 Review

SUMMARY

- All objects with mass attract one another with a gravitational force.
- The gravitational force acts equally on each of the masses.
- The magnitude of the gravitational force is given by Newton's law of universal gravitation:

$$F = \frac{GMm}{r^2}$$

- Gravitational forces are usually negligible unless one of the objects is massive, e.g. a planet.
- The weight of an object on the Earth's surface is due to the gravitational attraction of the Earth.
- A gravitational field is a region in which a gravitational force is exerted on all matter within that region.
- A gravitational field can be represented by a gravitational field diagram:
  - The arrowheads indicate the direction of the gravitational force.
  - The spacing of the lines indicates the relative strength of the field. The closer the line spacing, the stronger the field.
- The strength of a gravitational field can be calculated using the following formulae:

$$g = \frac{F_g}{m} \text{ or } g = \frac{GM}{r^2}$$

- The acceleration due to gravity of an object near the Earth's surface can be calculated using the dimensions of the Earth:

$g = \frac{Gm_{Earth}}{(r_{Earth})^2} = 9.8\,\text{m}\,\text{s}^{-2}$ towards the centre of the Earth

- The gravitational field strength on the Earth's surface is approximately $9.8\,\text{N}\,\text{kg}^{-1}$. This varies from location to location and with altitude.
- The gravitational field strength on the surface of any other planet depends on the mass and radius of the planet.

KEY QUESTIONS

1 What are the proportionalities in Newton's law of universal gravitation?

2 Calculate the force of gravitational attraction between the Sun and Mars given the following data:
$m_{Sun} = 2.0 \times 10^{30}\,\text{kg}$
$m_{Mars} = 6.4 \times 10^{23}\,\text{kg}$
$r_{Sun–Mars} = 2.2 \times 10^{11}\,\text{m}$

3 The force of gravitational attraction between the Sun and Mars is $1.8 \times 10^{21}\,\text{N}$. Calculate the acceleration of Mars given that $m_{Mars} = 6.4 \times 10^{23}\,\text{kg}$.

4 On 14 April 2014, Mars came within 93 million km of Earth. Its gravitational effect on the Earth was the strongest it had been for over 6 years. Use the following data to answer the questions below.
$m_{Sun} = 2.0 \times 10^{30}\,\text{kg}$
$m_{Earth} = 6.0 \times 10^{24}\,\text{kg}$
$m_{Mars} = 6.4 \times 10^{23}\,\text{kg}$

a Calculate the gravitational force between the Earth and Mars on 14 April 2014.

b Calculate the force of the Sun on the Earth if the distance between them was 153 million km.

c Compare your answers to parts (a) and (b) above by expressing the Mars–Earth force as a percentage of the Sun–Earth force.

5 The gravitational field strength, g, is measured as $5.5\,\text{N}\,\text{kg}^{-1}$ at a distance of 400 km from the centre of a planet. The distance from the centre of the planet is then increased to 1200 km. What would the ratio of the magnitude of the gravitational field strength be at this new distance compared to the original measurement?

6 On 12 November 2014, the *Rosetta* spacecraft landed a probe on the comet 67P/Churyumov–Gerasimenko. Assuming this comet is a roughly spherical object with a mass of $1 \times 10^{13}\,\text{kg}$ and a diameter of 1.8 km, calculate the gravitational field strength on its surface.

7 The masses and radii of three planets are given in the following table.

Planet	Mass (kg)	Radius (m)
Mercury	3.30×10^{23}	2.44×10^{6}
Saturn	5.69×10^{26}	6.03×10^{7}
Jupiter	1.90×10^{27}	7.15×10^{7}

Calculate the magnitude of the gravitational field strength, g, at the surface of each planet.

4.2 Satellite motion

PHYSICS INQUIRY CCT N

Satellite motion

How does the force of gravity determine the motion of planets and satellites?

COLLECT THIS...

- string
- washer
- retort stand or tape to attach string to ceiling
- ball
- support for ball (beaker, roll of tape or rolled paper)

DO THIS...

1 Place the ball on the support. Tie the string to the washer, and fix the string so that it is directly above the centre of the ball as shown below, and falls on the centre line or equator of the ball.
2 Starting with the washer against the ball, tap the washer. Try different directions and strengths of the tap to try to get the washer to orbit around the ball for the longest time.

RECORD THIS...

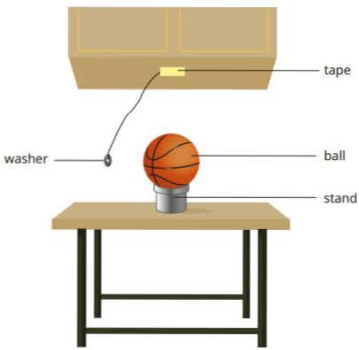
Describe your observations of the orbit for different initial forces.

Present a free-body diagram of the washer orbiting the ball. Refer back to the coverage of conical pendulums in Chapter 3 on page 78 for additional assistance if required.

REFLECT ON THIS...

How does the force of gravity determine the motion of planets and satellites?

What do your observations tell you about how satellites are launched into orbit?



Isaac Newton's development of the law of universal gravitation built on work previously done by Nicolaus Copernicus, Johannes Kepler and Galileo Galilei. Copernicus had proposed a Sun-centred (heliocentric) solar system. Galileo had developed laws relating to motion near the Earth's surface and Kepler had devised rules concerned with the motion of the planets. Kepler's laws on the motion of planets were published 80 years before Newton published his law of universal gravitation.

In this section, you will look at how Newton synthesised the work of Galileo and Kepler and proposed that the force that was causing an apple to fall to the Earth was the same force that was keeping the Moon in its orbit. Newton was the first to propose that satellites could be placed in orbit around Earth, almost 300 years before it was technically possible to do this. Now, thousands of artificial satellites are in orbit around Earth and are an essential part of modern life (Figure 4.2.1).

FIGURE 4.2.1 Astronauts on a repair mission to the Hubble Space Telescope in 1994. The satellite initially malfunctioned, but the repair was successful and the telescope is still going strongly.

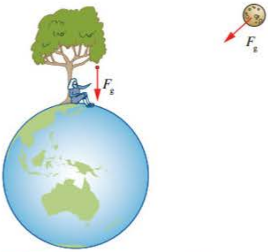


FIGURE 4.2.2 Newton realised that the gravitational attraction of the Earth was determining the motions of both the Moon and the apple.

NEWTON'S THOUGHT EXPERIMENT

A **satellite** is an object in a stable orbit around another object. Isaac Newton developed the notion of satellite motion while working on his theory of gravitation. He was comparing the motion of the Moon with the motion of a falling apple and realised that it was the gravitational force of attraction towards the Earth that determined the motion of both objects (Figure 4.2.2). He reasoned that if this force of gravity was not acting on the Moon, the Moon would move at constant speed in a straight line at a tangent to its orbit.

Newton proposed that the Moon, like the apple, was also falling. It was continuously falling to the Earth without actually getting any closer to the Earth. He devised a thought experiment in which he compared the motion of the Moon with the motion of a cannonball fired horizontally from the top of a high mountain.

Newton's thought experiment is illustrated in Figure 4.2.3. In this thought experiment, if the cannonball was fired at a low speed, it would not travel a great distance before gravity pulled it to the ground (see the shortest dashed line in Figure 4.2.3b). If it was fired with a greater velocity, it would follow a less curved path and land a greater distance from the mountain (see the next two dashed lines in Figure 4.2.3b). Newton reasoned that, if air resistance was ignored and if the cannonball was fired fast enough, it could travel around the Earth and reach the place from where it had been launched (shown by the solid circular line in Figure 4.2.3b). At this speed, it would continue to circle the Earth indefinitely even though the cannonball has no propulsion system.

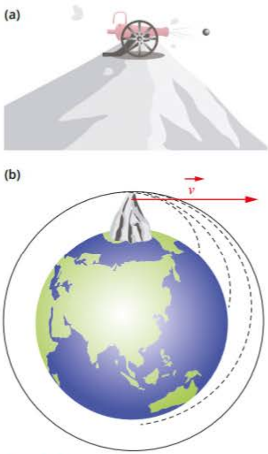


FIGURE 4.2.3 These diagrams show (a) how a projectile that was fired fast enough from a very high mountain (b) would fall all the way around the Earth and become an Earth satellite.

In reality, satellites could not orbit the Earth at low altitudes, because of air resistance. Nevertheless, Newton had proposed the notion of an artificial satellite hundreds of years before one was actually launched. Any object placed at the right altitude with enough speed would simply continue in its orbit.

PHYSICS IN ACTION CCT L

Apparent weightlessness

Your apparent weight is a contact reaction force that acts upwards on you from a surface because gravity is pulling you down on that surface. So if you are not standing on a surface, then you will experience zero apparent weight or **apparent weightlessness**. This means that you will experience apparent weightlessness the moment you step off the top platform of a diving pool or as you skydive from a plane.

Astronauts also experience apparent weightlessness in the International Space Station, which orbits about 370 km above the surface of the Earth (about the horizontal distance from Sydney to the town of Gundagai).

Whenever you are in free fall, you experience apparent weightlessness. It follows then that whenever you experience apparent weightlessness, you must be in free fall. When astronauts experience apparent weightlessness, they are not floating in space as they orbit the Earth. They are actually in free fall. Astronauts and their spacecraft are both falling, but not directly towards the Earth. The astronauts are actually moving horizontally, as shown in Figure 4.2.4. Astronauts are moving at a velocity relative to the Earth so they are moving across the sky at the same rate as they are falling. The combined effect is that they fall in a curved path that exactly mirrors the curve of the Earth. So they fall, but continually miss the Earth as the surface of the Earth curves away from their path.

Importantly there is a significant difference between apparent weightlessness and true weightlessness. True weightlessness only occurs when the gravitational field strength is zero. This only occurs in deep space, far enough away from any planets that their gravitational effect is zero. Apparent weightlessness, however, can occur when still under the influence of a gravitational field.

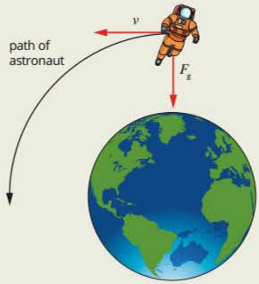


FIGURE 4.2.4 Astronauts are in free fall while orbiting the Earth.

NATURAL SATELLITES

WS 5.5

Natural satellites have existed throughout the universe for billions of years. The planets and asteroids of the solar system are natural satellites of the Sun (Figure 4.2.5).

The Earth has one natural satellite: the Moon. The largest planets—Jupiter and Saturn—have more than 60 natural satellites each in orbit around them. Most of the stars in the Milky Way galaxy have planets and more of these exoplanets are being discovered each year.

FIGURE 4.2.5 The planets are natural satellites of the Sun.

ARTIFICIAL SATELLITES

Since the Space Age began in 1957 with the launch of Sputnik, about 6000 **artificial satellites** have been launched into orbit around the Earth. Today there are around 4000 still in orbit, although only around 1200 of these are operational.

Satellites in orbit around the Earth are classified as low, medium or high orbit.

- Low orbit: 180 km to 2000 km altitude. Most satellites orbit in this range (an example is shown in Figure 4.2.6). These include the Hubble Space Telescope, which is used by astronomers to view objects right at the edge of the universe.
- Medium orbit: 2000 km to 36 000 km altitude. The most common satellites in this region are the Global Positioning System (GPS) satellites used to run navigation systems.
- High orbit: 36 000 km altitude or greater. Australia uses the Optus satellites for communications, and deep-space weather pictures come from the Japanese MTSAT-1R satellite. The satellites that sit at an altitude of about 36 000 km and orbit with a period of 24 hours are known as **geostationary satellites** (or geosynchronous satellites). Most communications satellites are geostationary.

FIGURE 4.2.6 A low-orbit satellite called the Soil Moisture and Ocean Salinity (SMOS) probe was launched in August 2014. Its role is to measure water movements and salinity levels on Earth as a way of monitoring climate change. It was launched from northern Russia by the European Space Agency (ESA).

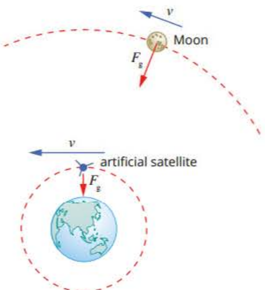


FIGURE 4.2.7 The only force acting on these artificial and natural satellites is the gravitational attraction of the Earth. Both orbit with a centripetal acceleration equal to the gravitational field strength at their locations.

GO TO ➤ Section 3.1 page 72

Earth satellites can have different orbital paths depending on their function:

- equatorial orbits, where the satellite always travels above the equator
- polar or near-polar orbits, where the satellite travels over or close to the North and South poles as it orbits
- inclined orbits, which lie between equatorial and polar orbits.

Satellites are used for a multitude of different purposes, with 60 per cent used for communications.

As discussed in Chapter 3, all objects travelling in circular motion require a centripetal force. Artificial and natural satellites are not propelled by rockets or engines. They orbit in free fall and the only force acting on them is the gravitational attraction between themselves and the body about which they orbit. This means the gravitational attraction is equivalent to the centripetal force of the satellite's motion. The satellites therefore have a centripetal acceleration that is equal to the gravitational field strength at their location (Figure 4.2.7). Centripetal acceleration is covered in more detail in Chapter 3.

Artificial satellites are often equipped with tanks of propellant that are squirted in the appropriate direction when the orbit of the satellite needs to be adjusted.

PHYSICSFILE ICT S

Space junk

Today there are around 1200 satellites that are still in operation. There are also around 2800 satellites that have reached the end of their operational life or have malfunctioned but are still in orbit.

In 2007, a Chinese satellite was deliberately destroyed by a missile, creating thousands of pieces of debris. In 2009, a collision between the defunct Russian Cosmos 2251 and operational US Iridium 33 created even more debris. This debris and the defunct satellites are classified as space junk (Figure 4.2.8).

The presence of this fast-moving space junk puts the other satellites and the International Space Station at risk from collision. Currently around 22 000 pieces of space junk are being tracked and monitored. There have been a number of occasions where satellites have been moved to avoid collisions with space junk.

The UN has passed a resolution to remove defunct satellites from low-Earth orbits by placing them in much higher orbits, or bringing them back to Earth and allowing them to burn up in the atmosphere.

FIGURE 4.2.8 An exaggerated map showing the location of space debris and abandoned satellites in near-Earth orbits.

PHYSICSFILE ICT

See the International Space Station (ISS) and other satellites

It is easy to see low-orbit satellites if you are away from city lights. The best time to look is just after sunset. If you can, go outside and look for any slow-moving objects passing across the star background.

There are also many websites that will allow you to track and predict the real-time paths of satellites. You can use the NASA 'Spot the Station' website to see when the ISS is passing over your part of the planet. The ISS is so bright that it is easy to see from most locations.

KEPLER'S LAWS

Johannes Kepler, a German astronomer (depicted in Figure 4.2.9), published his three laws regarding the motion of planets in 1609. This was about 80 years before Newton's law of universal gravitation was published. Kepler analysed the motion of the planets in orbit around the Sun, but Kepler's laws can be used for any satellite in orbit around any central mass.

Kepler's laws are as follows:

1 The planets move in elliptical orbits with the Sun at one focus.
2 The line connecting a planet to the Sun sweeps out equal areas in equal intervals of time (Figure 4.2.10).
3 For every planet, the ratio of the cube of the average orbital radius, r, to the square of the period, T, of revolution is the same, i.e. $\frac{r^3}{T^2}$ = a constant, k.

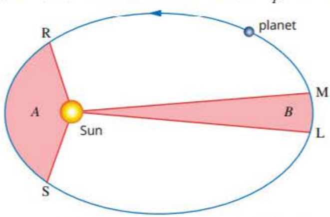


FIGURE 4.2.10 The planets, which are natural satellites of the Sun, orbit in elliptical paths with the Sun at one focus. Their speeds vary continually, and they are fastest when closest to the Sun. A line joining a planet to the Sun will sweep out equal areas in equal times. So, for example, the time it takes to move from R to S is equal to the time it takes to move from L to M, and so area *A* is the same as area *B*.

FIGURE 4.2.9 Johannes Kepler was the first to work out that the planets do not travel in circular paths, but rather in elliptical paths.

Kepler's first two laws proposed that planets move in elliptical paths and that they move faster when they are closer to the Sun. It took Kepler many months of laborious calculations to arrive at his third law. Newton used Kepler's laws to justify the inverse square relationship that he used in his law of universal gravitation.

CALCULATING THE ORBITAL PROPERTIES OF SATELLITES

The speed, v, of a satellite can be calculated from its motion for one revolution. For simplicity, we will assume that the orbital paths are approximately circular. This means that a satellite will travel a distance equal to the circumference of the orbit, $2\pi r$, in the time of one period, T.

> The speed, v, of a satellite in a circular orbit is given by:
>
> $$v = \frac{\text{distance}}{\text{time}} = \frac{2\pi r}{T}$$
>
> where
>
> r is the radius of the orbit (m)
>
> T is the time for one revolution, or the period (s)

The centripetal acceleration of a satellite can be determined from the gravitational field strength at its location. Satellites are in free fall; therefore, the only force acting is gravity, F_g. For example, the International Space Station (ISS) is in orbit at a distance from Earth where g is $8.8\,\text{N}\,\text{kg}^{-1}$, and so it orbits with a centripetal acceleration of $8.8\,\text{m}\,\text{s}^{-2}$.

The centripetal acceleration, a_c, of the satellite can also be calculated by considering its circular motion. The equation for speed given above can be substituted into the centripetal acceleration formula:

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

Since $v = \frac{2\pi r}{T}$ then $\frac{v^2}{r} = \frac{4\pi^2 r}{T^2}$.

Since the centripetal acceleration of the satellite is equal to the gravitational field strength at the location of its orbit, and using the gravitational field strength equation from earlier, we can give the following expression.

> The centripetal acceleration, a_c, of a satellite in circular orbit is given by:
>
> $$a_c = \frac{v^2}{r} = \frac{4\pi^2 r}{T^2} = \frac{GM}{r^2} = g$$
>
> where
>
> v is the speed of the satellite ($\text{m}\,\text{s}^{-1}$)
>
> r is the radius of the orbit (m)
>
> T is the period of orbit (s)
>
> M is the central mass (kg)
>
> g is the gravitational field strength at r ($\text{N}\,\text{kg}^{-1}$)
>
> G is the gravitational constant, $6.67 \times 10^{-11}\,\text{N}\,\text{m}^2\,\text{kg}^{-2}$

GO TO ➤ Section 3.1 page 72

Remember, acceleration can be a vector or a scalar quantity. From Chapter 3, if you are to write the centripetal acceleration as a vector, the direction will always be towards the centre of the circular motion—the same direction as the net force.

These relationships can be manipulated to determine any feature of a satellite's motion: its speed, radius of orbit or period of orbit. They can also be used to find the mass, M, of the central body around which the satellite orbits.

In the same way as with freely falling objects at the Earth's surface, the mass of the satellite itself has no effect on any of these orbital properties.

Worked example 4.2.1

WORKING WITH KEPLER'S LAWS

Determine the orbital speed of the Moon, assuming it is in a circular orbit of radius 384 000 km around the Earth. Take the mass of the Earth to be 5.97×10^{24} kg and use $G = 6.67 \times 10^{-11}$ N m^2 kg^{-2}.

Thinking	Working
Ensure that the variables are in their standard units.	$r = 384\,000 \text{ km} = 3.84 \times 10^8 \text{ m}$
Choose the appropriate relationship between the orbital speed, *v*, and the data that has been provided.	$a_c = g = \frac{GM}{r^2} = \frac{v^2}{r}$ $\therefore \frac{GM}{r^2} = \frac{v^2}{r}$ $\therefore \frac{GM}{r} = v^2$
Make *v*, the orbital speed, the subject of the equation.	$v = \sqrt{\frac{GM}{r}}$
Substitute in values and solve for the orbital speed, *v*.	$v = \sqrt{\frac{GM}{r}}$ $= \sqrt{\frac{(6.67 \times 10^{-11}) \times (5.97 \times 10^{24})}{3.84 \times 10^8}}$ $= 1.02 \times 10^3 \text{ m s}^{-1}$

Worked example: Try yourself 4.2.1

WORKING WITH KEPLER'S LAWS

Determine the orbital speed of a satellite, assuming it is in a circular orbit of radius of 42 100 km around the Earth. Take the mass of the Earth to be 5.97×10^{24} kg and use $G = 6.67 \times 10^{-11}$ N m^2 kg^{-2}.

HOW NEWTON DERIVED KEPLER'S THIRD LAW USING ALGEBRA

It took Kepler many months of trial-and-error calculations to arrive at his third law:

$$\frac{r^3}{T^2} = \text{constant}$$

Newton was able to use some clever algebra to derive this from his law of universal gravitation. Using Newton's law for the magnitude of the gravitational force:

$$F_g = mg = m\frac{4\pi^2 r}{T^2} = \frac{GMm}{r^2}$$

$$\therefore \frac{r^3}{T^2} = \frac{GM}{4\pi^2}$$

For any central mass, M, the term $\frac{GM}{4\pi^2}$ is constant and the ratio $\frac{r^3}{T^2}$ is equal to this constant value for all of its satellites (Figure 4.2.11).

So, for example, if you know the orbital radius, r, and period, T, of one of the moons of Saturn, you could calculate $\frac{r^3}{T^2}$ and use this as a constant value for all of Saturn's moons. If you knew the period, T, of a different satellite of Saturn, it would then be straightforward to calculate its orbital radius, r.

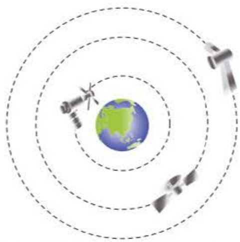


FIGURE 4.2.11 These three satellites are at different distances from Earth and hence according to Kepler's third law will have different orbital periods. For all three, the ratio of $\frac{r^3}{T^2}$ will equal the same constant value.

Worked example 4.2.2

SATELLITES IN ORBIT

Ganymede is the largest of Jupiter's moons. It has a mass of 1.66×10^{23} kg, an orbital radius of 1.07×10^6 km and an orbital period of 7.15 days (6.18×10^5 s).

a Use Kepler's third law to calculate the orbital radius (in km) of Europa, another moon of Jupiter, which has an orbital period of 3.55 days.

Thinking	Working
Note down the values for the known satellite. You can work in days and km as this question involves a ratio.	Ganymede: $r = 1.07 \times 10^6$ km $T = 7.15$ days
For all satellites of a central mass, $\frac{r^3}{T^2}$ = constant. Work out this ratio for the known satellite.	$\frac{r^3}{T^2}$ = constant $= \frac{(1.07 \times 10^6)^3}{7.15^2}$ $= 2.40 \times 10^{16}$
Use this constant value with the ratio for the satellite in question. Make sure T is in days to match the ratio calculated in the previous step.	Europa: $T = 3.55$ days, $r = ?$ $\frac{r^3}{T^2}$ = constant $\frac{r^3}{3.55^2} = 2.40 \times 10^{16}$
Make r^3 the subject of the equation.	$r^3 = 2.40 \times 10^{16} \times 3.55^2$ $= 3.02 \times 10^{17}$
Solve for r. The unit for r is km as the original ratio was calculated using km.	$r = \sqrt[3]{3.02 \times 10^{17}}$ $= 6.71 \times 10^5$ km Note: Europa has a shorter period than Ganymede so you should expect Europa to have a smaller orbit than Ganymede.

b Use the orbital data for Ganymede to calculate the mass of Jupiter.

Thinking	Working
Note down the values for the known satellite. You must work in SI units to find the mass value in kg.	Ganymede/Jupiter: $r = 1.07 \times 10^9$ m $T = 6.18 \times 10^5$ s $m = 1.66 \times 10^{23}$ kg $G = 6.67 \times 10^{-11}$ N m^2 kg^{-2} $M = ?$
Select the expressions from the equation for centripetal acceleration that best suit your data. $a_c = \frac{v^2}{r} = \frac{4\pi^2 r}{T^2} = \frac{GM}{r^2} = g$	Use the 3rd and 4th terms of the expression. $\frac{4\pi^2 r}{T^2} = \frac{GM}{r^2}$ These two expressions use the given variables r and T, and the constant G, so that a solution may be found for M.
Transpose the equation to make M the subject.	$M = \frac{4\pi^2 r^3}{GT^2}$
Substitute values and solve.	$M = \frac{4\pi^2 (1.07 \times 10^9)^3}{6.67 \times 10^{-11} \times (6.18 \times 10^5)^2}$ $= 1.90 \times 10^{27}$ kg

c Calculate the orbital speed of Ganymede in km s^{-1}.	
Thinking	**Working**
Note values you will need to use in the equation $v = \frac{2\pi r}{T}$.	Ganymede: $r = 1.07 \times 10^6$ km $T = 6.18 \times 10^5$ s $v = ?$
Substitute values and solve. The answer will be in $km\,s^{-1}$ if r is expressed in km.	$v = \frac{2\pi r}{T}$ $= \frac{2\pi \times 1.07 \times 10^6}{6.18 \times 10^5}$ $= 10.9\,km\,s^{-1}$

Worked example: Try yourself 4.2.2

SATELLITES IN ORBIT

Callisto is the second largest of Jupiter's moons. It is about the same size as the planet Mercury. Callisto has a mass of 1.08×10^{23} kg, an orbital radius of 1.88×10^6 km and an orbital period of 1.44×10^6 s (16.7 days).

a Use Kepler's third law to calculate the orbital radius (in km) of Europa, another moon of Jupiter, which has an orbital period of 3.55 days.

b Use the orbital data for Callisto to calculate the mass of Jupiter.

c Calculate the orbital speed of Callisto in $km\,s^{-1}$.

PHYSICSFILE ICT

SuitSat1

One of the more unusual satellites was launched from the International Space Station on 3 February 2006. It was an obsolete Russian spacesuit into which the astronauts had placed a radio transmitter, batteries and some sensors. Its launch involved simply being pushed off by one of the astronauts while on a spacewalk. SuitSat1 (Figure 4.2.12) was meant to transmit signals that would be picked up by ham radio operators on Earth for a few weeks, but transmissions ceased after just a few hours. The spacesuit burned up in the atmosphere over Western Australia in September 2006.

SuitSat2 was launched in August 2011 and contained experiments created by school students. It re-entered Earth's atmosphere in January 2012 after 5 months in orbit.

FIGURE 4.2.12 This photograph does not show an astronaut drifting off to certain death in space. This is SuitSat1, one of the strangest satellites ever launched, at the start of its mission.

PHYSICS IN ACTION A ICT

Three satellites

Geostationary Meteorological Satellite MTSAT-1R

The Japanese MTSAT-1R satellite was launched in February 2005, and orbits at approximately 35 800 km directly over the equator. At its closest point to the Earth, known as the **perigee**, its altitude is 35 776 km. At its furthest point from the Earth, known as the **apogee**, it is at 35 798 km. MTSAT-1R orbits at a longitude of 140° E, so it is just to the north of Cape York and ideally located for use by Australia's weather forecasters. It has a period of 24 hours, so is in a geostationary orbit.

Signals from MTSAT-1R are transmitted every 2 hours and are received by a satellite dish on the roof of the head office at the Bureau of Meteorology in Perth. Infrared images show the temperature variations in the atmosphere and are invaluable in weather forecasting. MTSAT-1R is box-like and measures about 2.6 m along each side. It has a mass of 1250 kg and is powered by solar panels that, when deployed, take its overall length to over 30 m.

Hubble Space Telescope (HST)

This cooperative venture between NASA and the European Space Agency (ESA) was launched by the crew of the space shuttle *Discovery* on 25 April 1990. Hubble is a permanent unoccupied space-based observatory with a 2.4 m diameter reflecting telescope, spectrographs and a faint-object camera. It orbits above the Earth's atmosphere, producing images of distant stars and galaxies far clearer than those from ground-based observatories (Figure 4.2.13). The HST is in a low-Earth orbit inclined at 28° to the equator. Its expected life span was originally around 15 years, but service and repair missions have extended its life and it is still in use today.

FIGURE 4.2.13 The Hubble Space Telescope has produced spectacular images of stars that are clearer than from any ground-based telescope. This is an image of the spiral galaxy known as NGC 3344.

National Oceanic and Atmospheric Administration Satellite (NOAA-19)

Many of the US-owned and operated NOAA satellites are located in low-altitude near-polar orbits. This means that they pass close to the poles of the Earth as they orbit. NOAA-19 was launched in February 2009 and orbits at an inclination of 99° to the equator. Its low altitude means that it captures high-resolution pictures of small bands of the Earth. The data is used in local weather forecasting as well as to provide enormous amounts of information for monitoring global warming and climate change.

The properties of these satellites are summarised in Table 4.2.1.

TABLE 4.2.1 A comparison of the three satellites discussed in this section

Satellite	Orbit	Inclination	Perigee (km)	Apogee (km)	Period
MTSAT-1R	equatorial	0°	35 776	35 798	1 day
Hubble	inclined	28°	591	599	96.6 min
NOAA-19	near polar	99°	846	866	102 min

4.2 Review

SUMMARY

- A satellite is an object that is in a stable orbit around a larger central mass.
- The only force acting on a satellite is the gravitational attraction between it and the central body.
- Satellites are in continual free fall. They move with a centripetal acceleration that is equal to the gravitational field strength at the location of their orbit.
- The speed of a satellite, v, is given by:

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

- For a satellite in a circular orbit:

$$a_c = \frac{v^2}{r} = \frac{4\pi^2 r}{T^2} = \frac{GM}{r^2} = g$$

- The magnitude of the gravitational force acting on a satellite in a circular orbit is given by:

$$F_g = \frac{mv^2}{r} = \frac{4\pi^2 rm}{T^2} = \frac{GMm}{r^2} = mg$$

- For any central body of mass, M: $\frac{r^3}{T^2} = \frac{GM}{4\pi^2}$ = constant, so knowing another satellite's orbital radius, r, enables its period, T, to be determined.

KEY QUESTIONS

1 Which of the following is correct?
  A Earth is a satellite of Mars.
  B The Moon is a satellite of the Sun.
  C Earth is a satellite of the Sun.
  D The Sun is a satellite of Earth.

2 A geostationary satellite orbits above Singapore, which is on the equator. Which of the following statements about the satellite is correct?
  A It is in a low orbit.
  B It is in a high orbit.
  C It passes over the north pole.
  D It is not moving.

3 The gravitational field strength at the location where the Optus D1 satellite is in stable orbit around the Earth is equal to $0.22\,\text{N kg}^{-1}$. The mass of this satellite is 2.3×10^3 kg.
  a Using only the information given, calculate the magnitude of the acceleration of this satellite as it orbits.
  b Calculate the net force acting on this satellite as it orbits.

4 One of Saturn's moons is Atlas, which has an orbital radius of 1.37×10^5 km and a period of 0.60 days. The largest of Saturn's moons is Titan. It has an orbital radius of 1.20×10^6 km. What is the orbital period of Titan in days?

5 Different types of satellite have different types of orbit, as shown in the table below. Calculate the strength of the Earth's gravitational field in each orbit. Give all of your answers to three significant figures.

$r_{\text{Earth}} = 6380$ km

$m_{\text{Earth}} = 5.97 \times 10^{24}$ kg

	Type of orbit	Altitude (km)
a	low Earth orbit	2000
b	medium Earth orbit	10 000
c	semi-synchronous orbit	20 200
d	geosynchronous orbit	35 786

4.3 Gravitational potential energy

FIGURE 4.3.1 Satellites in orbit have gravitational potential energy.

GO TO ➤ Revision page 87

The concept of gravitational potential energy should be familiar to you from Year 11 Physics. However, the nature of a gravitational field means that a more sophisticated understanding of gravitational potential energy is needed when considering the motion of objects like rockets or satellites (Figure 4.3.1).

ENERGY IN A CONSTANT GRAVITATIONAL FIELD

Up until now (see Chapter 3), our consideration of energy in gravitational fields has been simplified by the assumption that the Earth's gravitational field is constant. Under this assumption, the **gravitational potential energy** of an object, U, is directly proportional to the mass of the object, m, its height above the surface of the planet, h, and the strength of the gravitational field, g.

$U = mgh$

where

- U is the gravitational potential energy of an object (J)
- m is the mass of the object (kg)
- g is the magnitude of the gravitational field strength (N kg^{-1}; 9.80 N kg^{-1} near the surface of the Earth)
- h is the height of the object above a reference point (m)

$K = \frac{1}{2}mv^2$

where

- K is the kinetic energy of an object (J)
- m is the mass of the object (kg)
- v is the speed of the object (m s^{-1})

ENERGY IN A NON-CONSTANT GRAVITATIONAL FIELD

Newton's law of universal gravitation indicates that the strength of the Earth's gravitational field changes with altitude: the field is stronger close to the ground and weaker at high altitudes (Figure 4.3.2).

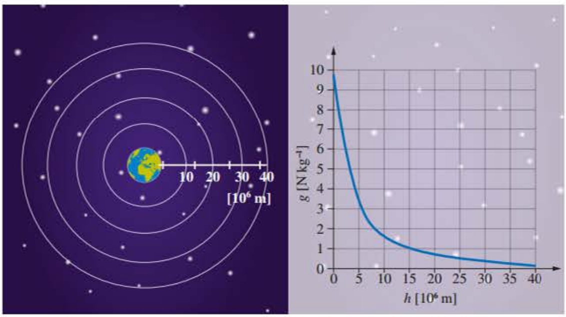


FIGURE 4.3.2 As the distance from the surface of the Earth is increased from 0 to 40 000 km, the value for g decreases rapidly from 9.8 N kg^{-1}, according to the inverse square law. The blue line on the graph gives the value of g at various altitudes (h).

Clearly it is not sufficient to assume a constant value for the Earth's gravitational field when considering objects like satellites or moons that orbit at high altitudes. In section 4.1, we saw that the magnitude of the gravitational field strength is given by the formula $g = \frac{GM}{r^2}$. Substituting this into the formula for gravitational potential energy:

$$U = mgh = m\frac{GM}{r^2}r$$

$U = -\frac{GMm}{r}$

where

U is the gravitational potential energy of an object (J)

m is the mass of the object (kg)

M is the mass of the Earth (or, more generally, the central body; kg)

G is the gravitational constant, $6.67 \times 10^{-11}\,\mathrm{N\,m^2\,kg^{-2}}$

r is the radius of the orbit of the object (the sum of the radius of the central body and the altitude of the orbit; in m)

Note that this formula contains a negative sign. Gravitational potential energy is measured against a reference level and a negative value means that it has moved below this level. By convention, the gravitational potential energy of a satellite is considered to be zero when it has escaped the gravitational field of the Earth. This only occurs when the satellite is a very large distance away (in theory, an infinite distance). As the satellite moves closer to the Earth, its distance from the Earth becomes less than the reference level and the potential energy of the satellite becomes a negative value.

Worked example 4.3.1

GRAVITATIONAL POTENTIAL ENERGY IN A NON-CONSTANT FIELD

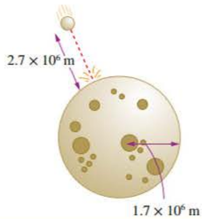
A communications satellite with a mass of 800 kg is orbiting the Earth at an altitude of 35 800 km. Calculate the gravitational potential energy of the satellite. (Use $r_{Earth} = 6.38 \times 10^6$ m and $m_{Earth} = 5.97 \times 10^{24}$ kg.)

Thinking	Working
Determine the radius of the satellite's orbit.	$r = 3.58 \times 10^7 + 6.38 \times 10^6$ $= 4.22 \times 10^7$ m
Recall the formula for the gravitational potential energy of a satellite.	$U = -\frac{GMm}{r}$
Substitute the values into the formula.	$U = -\frac{6.67 \times 10^{-11} \times 5.97 \times 10^{24} \times 800}{4.22 \times 10^7}$ $= -7.55 \times 10^9$ J

Worked example: Try yourself 4.3.1

GRAVITATIONAL POTENTIAL ENERGY IN A NON-CONSTANT FIELD

A 500 kg lump of space junk is plummeting towards the Moon (see the figure below). The Moon has a radius of 1.7×10^6 m and a mass of 7.3×10^{22} kg. Calculate the gravitational potential energy of the space junk when it is 2.7×10^6 m away from the surface of the Moon.



TOTAL ENERGY IN A NON-CONSTANT GRAVITATIONAL FIELD

A satellite has two important forms of energy: gravitational potential energy, U, and kinetic energy, K. As you saw in Chapter 3, it is often useful to consider the sum of these two energies (sometimes known as total mechanical energy), i.e. $E = U + K$.

The centripetal force of a satellite is equal to the force due to gravity. Combining this with the magnitude of the force from Newton's second law, and remembering the formula for the centripetal acceleration:

$$F = \frac{GMm}{r^2} = \frac{mv^2}{r}$$

$$\therefore\ mv^2 = \frac{GMm}{r}$$

Using the equation $K = \frac{1}{2}mv^2$, the kinetic energy is then equal to:

$$K = \frac{1}{2}mv^2 = \frac{GMm}{2r}$$

Combining this with the gravitational potential energy of a satellite, $U = -\frac{GMm}{r}$, the total energy is then

$$E = K + U = \frac{GMm}{2r} - \frac{GMm}{r} = -\frac{GMm}{2r}$$

$$E = K + U = -\frac{GMm}{2r}$$

where

- E is the total energy of an object (J)
- K is the kinetic energy of the object (J)
- U is the potential energy of the object (J)
- m is the mass of the object (kg)
- M is the mass of the Earth (or, more generally, the central body; kg)
- G is the gravitational constant, 6.67×10^{-11} N m^2 kg^{-2}
- r is the radius of the orbit of the object (the sum of the radius of the central body and the altitude of the orbit; in m)

Worked example 4.3.2

TOTAL ENERGY OF A SATELLITE

A communications satellite with a mass of 1390 kg is orbiting the Earth at an altitude of 357 km. (Use $r_{Earth} = 6.38 \times 10^6$ m and $m_{Earth} = 5.97 \times 10^{24}$ kg.)

a Calculate the total mechanical energy of this satellite.	
Thinking	**Working**
Determine the radius of the satellite's orbit.	$r = 357 \times 10^3 + 6.38 \times 10^6$ $= 6737000$ $= 6.74 \times 10^6$ m
Use the definition for total energy: $E = -\frac{GMm}{2r}$	$E = -\frac{GMm}{2r}$ $= -\frac{6.67 \times 10^{-11} \times 5.97 \times 10^{24} \times 1390}{2 \times 6.74 \times 10^6}$ $= -4.11 \times 10^{10}$ J

b Calculate the speed of the satellite.	
Thinking	**Working**
Recall the equation for the kinetic energy of a satellite.	$K = \frac{GMm}{2r}$
Substitute the known values and solve for K.	$K = \frac{GMm}{2r}$ $= \frac{6.67 \times 10^{-11} \times 5.97 \times 10^{24} \times 1390}{2 \times 6.74 \times 10^6}$ $= 4.11 \times 10^{10}$ J
Remember that the kinetic energy of an object can also be calculated with the equation: $K = \frac{1}{2}mv^2$ Use this to solve for the speed v.	$K = \frac{1}{2}mv^2$ $4.11 \times 10^{10} = \frac{1}{2} \times 1390 \times v^2$ $v^2 = 5.91 \times 10^7$ $v = 7690$ m s^{-1}

Worked example: Try yourself 4.3.2

TOTAL ENERGY OF A SATELLITE

Sputnik 1 was the first artificial satellite to be put into orbit. It had a mass of 84.0 kg and orbited the Earth at an altitude of 577 km. (Use $r_{Earth} = 6.38 \times 10^6$ m and $m_{Earth} = 5.97 \times 10^{24}$ kg.)

a Calculate the total mechanical energy of this satellite.

b Calculate the speed of the satellite.

It is important to note that the total energy of a satellite remains constant throughout its orbit, even though its distance from the attracting object may vary. For example, Kepler's first law states that planets travel in elliptical orbits around the Sun (Figure 4.3.3).

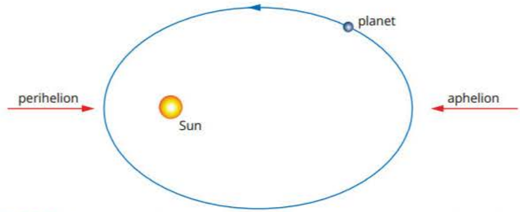


FIGURE 4.3.3 The total energy of a planet remains constant throughout its orbit, even though its distance from the Sun changes.

A planet's gravitational potential energy is greatest when it is at its **aphelion**, i.e. the point on its orbit when it is furthest from the Sun. As the planet moves towards its **perihelion** (i.e. point of closest approach to the Sun), its gravitational potential energy will decrease. However, according to Kepler's second law, the velocity of the planet will increase as it approaches the Sun. This means that the kinetic energy of the planet will increase in a way that exactly balances the decrease in gravitational potential energy, keeping the total energy of the planet constant. This pattern holds true for all satellites.

ENERGY CHANGES IN A NON-CONSTANT GRAVITATIONAL FIELD

The familiar formula for change in gravitational potential energy $\Delta U = mg\Delta h$ is developed assuming that work is done against a constant force of gravity: $\Delta U = W = F_{\parallel} s$. While this assumption holds true for objects close to the surface of a planet, it is not adequate for objects like satellites that move to altitudes at which the gravitational field of the planet becomes significantly diminished.

Consider the example of a 10 kg meteor falling towards the Earth from deep space as shown in Figure 4.3.4. As the meteor gets closer to the Earth, it moves through regions of increasing gravitational field strength. So the gravitational force, F_g, on the meteor increases as it approaches Earth. Since the force is not constant, this means that the work done on the meteor (which corresponds to its change in gravitational potential energy) cannot be found by simply multiplying the gravitational force by the distance travelled; it must be calculated directly from the formula $U = -\frac{GMm}{r}$.

FIGURE 4.3.4 As a meteor approaches Earth, it moves through an increasingly stronger gravitational field and so is acted upon by a greater gravitational force.

Note that the energy change of the meteor will be the same regardless of whether the meteor falls directly towards the planet (Figure 4.3.5a) or follows a more indirect path (Figure 4.3.5b).

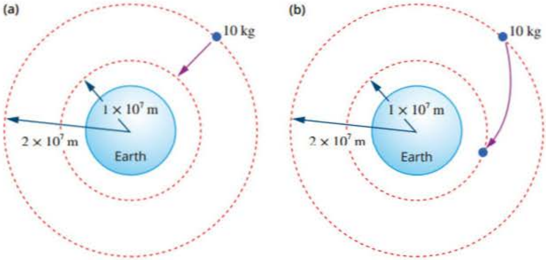


FIGURE 4.3.5 The change in gravitational potential energy will be the same whether the satellite takes (a) a direct path or (b) a curved path.

Worked example 4.3.3

CHANGES IN GRAVITATIONAL POTENTIAL ENERGY

A meteor with a mass of 10 kg is orbiting the Earth with an orbital radius of 2×10^7 m. It moves into a lower orbit with a radius of 1×10^7 m. Calculate the change in gravitational potential energy of the satellite.
(Use $m_{Earth} = 5.97 \times 10^{24}$ kg.)

Thinking	Working
Use the formula for gravitational potential energy to calculate the initial gravitational potential energy, U_i.	$U_i = -\frac{GMm}{r}$ $= -\frac{6.67 \times 10^{-11} \times 5.97 \times 10^{24} \times 10}{2 \times 10^7}$ $= -2.0 \times 10^8$ J
Use the formula for gravitational potential energy to calculate the final gravitational potential energy, U_f.	$U_f = -\frac{GMm}{r}$ $= -\frac{6.67 \times 10^{-11} \times 5.97 \times 10^{24} \times 10}{1 \times 10^7}$ $= -4.0 \times 10^8$ J
Calculate the change in gravitational potential energy.	$\Delta U = U_f - U_i$ $= -4.0 \times 10^8 - (-2.0 \times 10^8)$ $= -2 \times 10^8$ J

Worked example: Try yourself 4.3.3

CHANGES IN GRAVITATIONAL POTENTIAL ENERGY

A satellite with a mass of 500 kg is orbiting the Earth with an orbital radius of 7100 km. It moves into a lower orbit at an altitude of 6800 km. Calculate the change in gravitational potential energy of the satellite.
(Use $m_{Earth} = 5.97 \times 10^{24}$ kg.)

+ ADDITIONAL

Using the force–distance graph

When a free-falling body is acted upon by a varying gravitational force, the energy changes of the body can also be analysed by using a gravitational force–distance graph. As with other force–distance graphs, the area under the graph is equal to the work done, i.e. the energy change of the body. The area under the graph has units of newton metres (N m), which are equivalent to joules (J).

The shaded area in Figure 4.3.6 represents the decrease in gravitational potential energy of the 10 kg meteor as it falls from a distance of 2.0×10^7 m to 1.0×10^7 m from the centre of the Earth. This area also represents the amount of kinetic energy that the meteor gains as it approaches Earth.

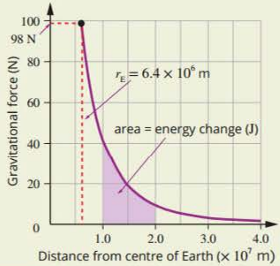


FIGURE 4.3.6 The gravitational force acting on a 10 kg meteor at different distances from the Earth. The shaded region represents the work done by the gravitational field as the body moves between 2.0×10^7 m and 1.0×10^7 m from the centre of the Earth.

According to the law of conservation of mechanical energy, when an object loses gravitational potential energy, its kinetic energy must increase. This is also true for satellites.

Consider the 10 kg meteor discussed above. When it moves from a higher orbit to a lower orbit, it loses 200 MJ of gravitational potential energy. It will correspondingly gain 200 MJ of kinetic energy and therefore orbit at a higher speed.

Escape velocity

When an object like a rocket has enough kinetic energy to escape the Earth's gravitational field, it is said to have reached **escape velocity**.

Since an object that has escaped a gravitational field has a potential energy of $U = 0$, due to the conservation of mechanical energy the escape velocity occurs when the kinetic energy of the object is equal in magnitude to its gravitational potential energy. In other words, when $K = U$.

$$\frac{1}{2}mv_{esc}^2 = \frac{GMm}{r}$$

$$v_{esc}^2 = \frac{2GM}{r}$$

$$v_{esc} = \sqrt{\frac{2GM}{r}}$$

The escape velocity is given by the equation:

$$v_{esc} = \sqrt{\frac{2GM}{r}}$$

where

- M is the mass of the Earth (or, more generally, the central body; in kg)
- G is the gravitational constant, 6.67×10^{-11} N m^2 kg^{-2}
- r is the radius of the orbit of the object (the sum of the radius of the central body and the altitude of the orbit; in m)

Note that escape velocity is independent of the mass of the object; i.e. regardless of how heavy the object is, its escape velocity will be the same.

Worked example 4.3.4

ESCAPE VELOCITY

Calculate the escape velocity for a 500 kg rocket being launched from the surface of the Earth. (Note: the Earth has a radius of 6.38×10^6 m and a mass of 5.97×10^{24} kg.)

Thinking	Working
Recall the definition of the escape velocity.	$v_{esc} = \sqrt{\frac{2GM}{r}}$
Identify the information required, and convert values into appropriate units where necessary.	$G = 6.67 \times 10^{-11}\,N\,m^2\,kg^{-2}$ $M = 5.97 \times 10^{24}\,kg$ $r = 6.38 \times 10^6\,m$
Substitute the values into the equation and solve for v_{esc}.	$v_{esc} = \sqrt{\frac{2 \times 6.67 \times 10^{-11} \times 5.97 \times 10^{24}}{6.38 \times 10^6}}$ $= 1.12 \times 10^4\,m\,s^{-1}$

Worked example: Try yourself 4.3.4

ESCAPE VELOCITY

Calculate the escape velocity for an 11 900 kg spacecraft being launched from the surface of the Moon. (Note: the Moon has a radius of 1.74×10^6 m and a mass of 7.32×10^{22} kg.)

WS 5.6

PHYSICS IN ACTION ICT

Voyager space probes

Two of the earliest human-made objects to achieve escape velocity are the Voyager space probes (Figure 4.3.7). These were launched in 1977 to explore the outer planets of the solar system. These probes have not only escaped the Earth's gravitational field, but are also now escaping the Sun's gravitational field as they travel out beyond the orbits of the most distant planets.

FIGURE 4.3.7 The Voyager probes have achieved escape velocity from the Sun's gravitational field.

4.3 Review

SUMMARY

- The gravitational potential energy formula $U = mgh$ assumes that the Earth's gravitational field is constant. This is approximately true for objects that are within a few kilometres of the Earth's surface.
- The strength of the Earth's gravitational field decreases as altitude increases.
- The gravitational potential energy of an object in a non-constant gravitational field is given by $U = -\frac{GMm}{r}$.
- The total energy of an object in a non-constant gravitational field is given by $E = K + U = -\frac{GMm}{2r}$.
- The total energy of a satellite remains constant throughout its orbit.
- Escape velocity is the velocity required for an object to escape a gravitational field: $v_{esc} = \sqrt{\frac{2GM}{r}}$.

KEY QUESTIONS

Where necessary, assume that the Earth has a radius of 6.38×10^6 m and a mass of 5.97×10^{24} kg.

1 Which one of the following statements is correct? A satellite in a circular orbit around the Earth will have:

A varying potential energy as it orbits

B varying kinetic energy as it orbits

C constant kinetic energy and constant potential energy

2 The path of a meteor plunging towards the Earth is as shown. Ignore air resistance when answering these questions.

a How does the gravitational field strength of the Earth change from point A to point D?

b How will the acceleration of the meteor change as it travels along the path shown?

c Which one or more of the following statements is correct?

A The kinetic energy of the meteor increases as it travels from A to D.

B The gravitational potential energy of the meteor decreases as it travels from A to D.

C The total energy of the meteor remains constant.

D The total energy of the meteor increases.

3 The *Saturn V* rocket that took the first astronauts to the Moon had a mass of 3000 tonnes. Its Stage I rockets fired for 6 minutes and took the rocket to an altitude of 67 km. Calculate the gravitational potential energy of the rocket at its final altitude.

4 A communications satellite of mass 240 kg is launched from a space shuttle that is in orbit 600 km above the Earth's surface. The satellite travels directly away from the Earth and reaches a maximum distance of 8000 km from the centre of the Earth before stopping due to the influence of the Earth's gravitational field. Calculate the change in gravitational potential energy of this satellite as it was launched.

5 Calculate the escape velocity (in $km\,s^{-1}$) for a spacecraft leaving Mars which has a radius of 3390 km and a mass of 6.42×10^{23} kg.

Chapter review

04

KEY TERMS

acceleration due to gravity
altitude
aphelion
apogee
apparent weightlessness
artificial satellite
escape velocity
field
geostationary satellite
gravimeter
gravitational constant
gravitational field
gravitational field strength
gravitational force
gravitational potential energy
inverse square law
natural satellite
Newton's law of universal gravitation
perigee
perihelion
satellite
uniform

REVIEW QUESTIONS

Where necessary, assume that the Earth has a radius of 6.4×10^6 m and a mass of 6.0×10^{24} kg.

1 Use Newton's law of universal gravitation to calculate the gravitational force acting on a person with a mass of 75 kg.

2 The gravitational force of attraction between Saturn and Dione, a moon of Saturn, is equal to 2.79×10^{20} N. Calculate the orbital radius of Dione. Use the following data:
mass of Dione = 1.05×10^{21} kg
mass of Saturn = 5.69×10^{26} kg

3 Of all the planets in the solar system, Jupiter exerts the largest force on the Sun: 4.2×10^{23} N. Calculate the scalar acceleration of the Sun due to this force, using the following data: $m_{Sun} = 2.0 \times 10^{30}$ kg.

4 The acceleration of the Moon caused by the gravitational force of the Earth is much larger than the acceleration of the Earth due to the gravitational force of the Moon. What is the reason for this?

5 Calculate the acceleration due to gravity on the surface of Mars if it has a mass of 6.4×10^{23} kg and a radius of 3400 km.

6 A comet of mass 1000 kg is plummeting towards Jupiter. Jupiter has a mass of 1.90×10^{27} kg and a planetary radius of 7.15×10^7 m. If the comet is about to crash into Jupiter, calculate the:

a magnitude of the gravitational force that Jupiter exerts on the comet
b magnitude of the gravitational force that the comet exerts on Jupiter
c acceleration of the comet towards Jupiter
d acceleration of Jupiter towards the comet.

7 A person standing on the surface of the Earth experiences a gravitational force of 900 N. What gravitational force will this person experience at a height of two Earth radii above the Earth's surface?

A 900 N
B 450 N
C 100 N
D zero

8 Calculate the weight of a 65 kg cosmonaut standing on the surface of Mars, given that the planet has a mass of 6.4×10^{23} kg and a radius of 3.4×10^6 m.

9 During a space mission, an astronaut of mass 80 kg initially accelerates at 30 m s^{-2} upwards, then travels in a stable circular orbit at an altitude where the gravitational field strength is 8.2 N kg^{-1}.

a What is the total force acting on the astronaut during lift-off?
A zero
B 660 N
C 780 N
D 3200 N

b During the lift-off phase, the astronaut will feel:
A lighter than usual
B heavier than usual
C the same as usual

c During the orbit phase, the gravitational force acting on the astronaut is:
A zero
B 660 N
C 780 N
D 3200 N

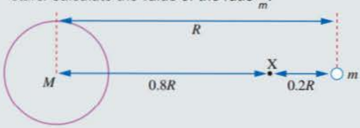
10 What are the main steps to follow when drawing gravitational field lines?

11 A group of students use a spring balance to measure the weight of a 150 g set of slotted masses to be 1.4 N. According to this measurement, what is the gravitational field strength in their classroom?

12 The Earth is a flattened sphere. Its radius at the poles is 6357 km compared to 6378 km at the equator. The Earth's mass is 5.97×10^{24} kg.

a Calculate the Earth's gravitational field strength at the equator.
b Using the information in part (a), calculate how much stronger the gravitational field would be at the North Pole compared with the equator. Give your answer as a percentage of the strength at the equator.

13 Two stars of masses *M* and *m* are in orbit around each other. As shown in the following diagram, they are a distance *R* apart. A spacecraft located at point X experiences zero net gravitational force from these stars. Calculate the value of the ratio $\frac{M}{m}$.



14 Give the most appropriate units for measuring gravitational field strength.

15 There are bodies outside our solar system, such as neutron stars, that produce very large gravitational fields. A typical neutron star can have a mass of 3.0×10^{30} kg and a radius of just 10 km. Calculate the gravitational field strength at the surface of such a star.

16 A newly discovered solid planet located in a distant solar system is found to be distinctly non-spherical in shape. Its polar radius is 5000 km, and its equatorial radius is 6000 km.

The gravitational field strength at the poles is $8.0\,\text{N kg}^{-1}$. How would the gravitational field strength at the poles compare with the strength at the equator?

17 An astronaut travels away from Earth to a region in space where the gravitational force due to Earth is only 1.0% of that at Earth's surface. What distance, in Earth radii, is the astronaut from the centre of the Earth?

18 Which of the following are properties of a geostationary satellite?

A It is in a low orbit.

B It orbits the Earth once every 24 hours.

C It passes over the North Pole.

D It is weightless.

19 Complete the table below, which contains information about the four largest moons of Jupiter:

Name	Radius of orbit ($\times 10^3$ km)	Period of orbit (h)
Io	422	42.5
Europa	671	
Ganymede	1070	
Callisto	1883	

20 Calculate the gravitational field strength experienced by a satellite orbiting at an altitude of 19 000 km.

21 Titan is the largest Moon of Saturn. It orbits Saturn once every 15.9 days at an average distance of 1.22×10^6 km.

a Calculate the orbital speed of Titan.

b Use this data to calculate the mass of Saturn.

22 **a** Determine the gravitational potential energy of a 1.0 kg mass 100 km above the Earth's surface.

b Determine the total energy of a 1.0 kg mass 100 km above the Earth's surface.

c Determine the height above the Earth's surface at which a 1.0 kg mass would have a total energy of -1×10^7 J.

23 A 20 tonne remote-sensing satellite is in a circular orbit around the Earth at an altitude of 600 km. The satellite is moved to a new stable orbit with an altitude of 2600 km. Determine the increase in the gravitational potential energy of the satellite as it moved from its lower orbit to its higher orbit.

24 A 20 kg rock is speeding towards Mercury. The radius of Mercury is 2.4×10^6 m and its mass is 3.3×10^{23} kg. Calculate the:

a gravitational field strength at 3.0×10^6 m from the centre of Mercury

b gravitational potential energy of the rock at 3.0×10^6 m from the centre of Mercury

c decrease in gravitational energy of the rock as it moves to a point that is just 2.5×10^6 m from the centre of Mercury.

25 A wayward satellite of mass 1000 kg is drifting towards the Earth. How much kinetic energy does the satellite gain as it travels from an altitude of 600 km to an altitude of 200 km?

26 Calculate the escape velocity (in km s^{-1}) for a spacecraft leaving Mercury which has a radius of 2440 km and a mass of 3.30×10^{23} kg.

27 After completing the activity on page 115, reflect on the inquiry question: How does the force of gravity determine the motion of planets and satellites?

MODULE 5 • REVIEW

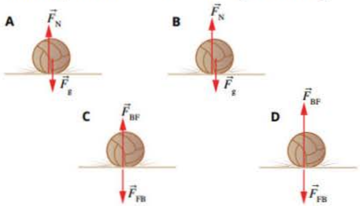
REVIEW QUESTIONS

Advanced mechanics

WS 5.8 WS 5.9 MR 5

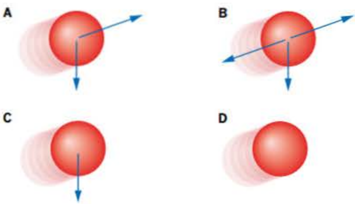
Multiple choice

1 A netball is dropped vertically from a height of 1.5 m onto a horizontal floor. The diagrams below relate to the instant that the ball reaches the floor and is stationary for a short period of time before rebounding. Which of the following correctly represents the action–reaction forces acting between the ball and the floor at this instant? (More than one answer may be correct.).



2 Two friends, Elvis and Kurt, are having a game of catch. Elvis throws a baseball to Kurt, who is standing 8.0 m away. Kurt catches the ball at the same height, 2.0 s after it is thrown. The mass of the baseball is 250 g. Ignore the effects of air resistance.

Which of the following diagrams best shows the forces acting on the ball just after it has left Elvis's hand?



The following information relates to questions 3–6.

Consider an astronaut inside a spacecraft from launch to a stable orbit. Choose your answers to questions 3 to 6 from the following options:

A apparent weightlessness
B weightlessness
C apparent weight
D gravitational force
E none of the above

3 As the astronaut and spacecraft are launched, which of the above will be greater than normal?

4 As the astronaut and spacecraft are launched, which of the above will remain constant?

5 As the astronaut and spacecraft are in a stable orbit above the Earth, which of the above will apply to the astronaut?

6 If the astronaut and spacecraft ventured into deep space, which of the above would apply to the astronaut?

7 Diana rolls a bowling ball down a smooth, straight ramp. Choose the option below that best describes the way the ball will travel.

A with constant speed
B with constant acceleration
C with decreasing speed
D with increasing acceleration

8 After travelling on a Ferris wheel, a boy remarks to a friend that he felt lighter than usual at the top of the ride. Which option explains why he might feel lighter at the top of the ride?

A He lost weight during the ride.
B The strength of the gravitational field was weaker at the top of the ride.
C The normal force there was larger than the gravitational force.
D The normal force there was smaller than the gravitational force.

9 Which of the following is the correct description of the maximum torque on a door?

A The maximum torque will be achieved when the force is at a 45° angle to the door.
B The maximum torque will be achieved when the force is parallel to the door.
C The maximum torque will be achieved when the force is at a 90° angle to the door.
D The maximum torque will be achieved when the force is at a 100° angle to the door.

10 Which of the following options would provide you with the greatest torque when opening a 1 metre-wide door?

A pushing with a force of 33 N at right angles to the door, in the middle of the door
B pushing with a force of 25 N at right angles to the door, 25 cm from the handle edge of the door
C pushing with a force of 50 N at right angles to the door, 25 cm from the hinges of the door
D pushing with a force of 25 N at right angles to the door, at the handle edge of the door

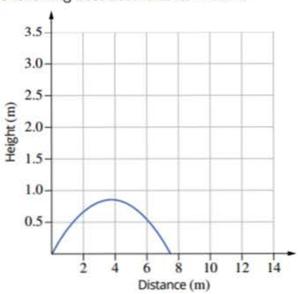
11 Which of the following descriptions of calculating torque is correct?
  A Torque can be calculated using either the perpendicular force or the perpendicular force arm.
  B Torque can only be calculated using the perpendicular force.
  C Torque can only be calculated using the perpendicular force arm.
  D Torque can only be calculated using both the perpendicular force and the perpendicular force arm together.

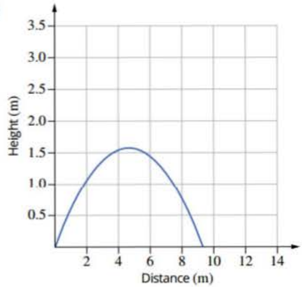
12 Which of the following correctly describes rotational equilibrium?
  A A net torque acts about the reference point and rotation does not occur.
  B A net torque acts about the reference point and rotation occurs.
  C No net torque acts about the reference point and rotation does not occur.
  D No net torque acts about the reference point and rotation occurs.

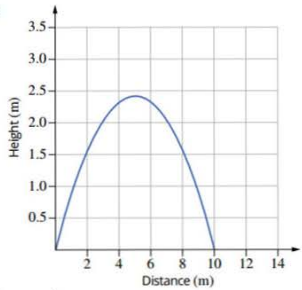
13 When an object is in static equilibrium, it experiences:
  A rotational equilibrium, but not translational equilibrium
  B rotational equilibrium and translational equilibrium
  C neither rotational equilibrium nor translational equilibrium
  D translational equilibrium, but not rotational equilibrium

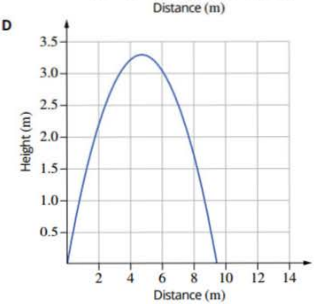
14 The International Space Station orbits the Earth with a period T. A research satellite orbits Neptune, which is approximately 16 times the mass of Earth. What is the period of the research satellite if it has the same orbital radius r as the ISS?
  A $8T$ B $\frac{T}{8}$
  C $4T$ D $\frac{T}{4}$

15 A 45 kg cannon ball is launched at 45° from the horizontal with a launch velocity of $10\,\text{m}\,\text{s}^{-1}$. Which of the following best describes its motion?

A


B


C


D


16 A golf ball is hit and travels 100 m before hitting the ground in 4.0 s. What was the launch angle to two significant figures?
  A 30°
  B 40°
  C 50°
  D 60°

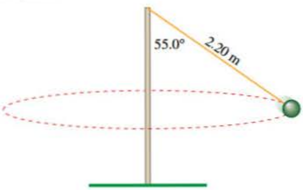
17 A wind turbine has blades 60.0 m in length that rotate so the tips of the blade travel a distance of 130 m in 2.50 seconds. At what angular velocity does the turbine rotate at?

A $62.1° s^{-1}$
B $1.18° s^{-1}$
C $62.1 rad s^{-1}$
D $1.18 rad s^{-1}$

18 A group of students use a spring balance to measure the weight of a 150 g set of slotted masses to be 1.4 N. According to this measurement, what is the gravitational field strength in their classroom?

A $9.8 N kg^{-1}$
B $9.5 N kg^{-1}$
C $9.3 N kg^{-1}$
D $9.1 N kg^{-1}$

19 During a game of Totem Tennis, the ball of mass 100 g is swinging freely in a horizontal circular path shown below. Determine the net force that is acting on the ball at this time.



A 1.4 N towards the right
B 1.4 N towards the left
C 140 N towards the left
D 140 N towards the right

20 The design speed of a 40 m radius velodrome track is $20 m s^{-1}$. What angle is the track banked to?

A 35°
B 40°
C 45°
D 50°

Short answer

21 Use Newton's law of universal gravitation to calculate the size of the force between two masses of 24 kg and 81 kg, with a distance of 0.72 m between their centres.

22 Three children are playing on a seesaw. Susan has a mass of 35 kg and is sitting 3.0 m from the pivot point. Her younger brother James has a mass of 25 kg and is sitting 3.5 m from the pivot on the same side as Susan. Where should their older brother Thomas (mass 42 kg) position himself in order to balance his two younger siblings?

23 An astronaut travels away from Earth to a region in space where the gravitational force due to Earth is only 1.0% of that at Earth's surface. What distance, in Earth radii, is the astronaut from the centre of the Earth?

24 Calculate the acceleration of an object dropped near the surface of Mercury if this planet has a mass of 3.3×10^{23} kg and a radius of 2500 km. Assume that the gravitational acceleration on Mercury can be calculated similarly to that on Earth.

25 A newly discovered solid planet located in a distant solar system is found to be distinctly non-spherical in shape. Its polar radius is 5000 km, and its equatorial radius is 6000 km.

The gravitational field strength at the poles is $8.0 N kg^{-1}$. How would the gravitational field strength at the poles compare with the strength at the equator?

26 Astronaut Alan Shepherd is the only person to have played golf on the moon. The acceleration due to gravity on the moon is approximately $1.6 m s^{-2}$, a lot less than it is on Earth. Assuming the ball travelled a distance of 200 m in 15 s, what was the initial velocity of the golf ball?

27 The ATV2 satellite was launched by the European Space Agency in February 2011 to deliver supplies to the International Space Station. The ATV2 satellite is in a circular orbit of radius 6.73×10^6 m.

The following information may be required to answer these questions.

Mass of ATV2 satellite and cargo = 1.2×10^4 kg
Mass of Earth = 5.98×10^{24} kg
Radius of Earth = 6.37×10^6 m

a What is the weight of the ATV2 and cargo when it is in its orbit?
b Calculate the orbital period of the ATV2 satellite.
c The satellite delivers its cargo to the ISS and now orbits with a mass of just 6.0 tonnes. How does this reduced mass affect the orbital period of the ATV2?

28 Two friends, Julianne and Bryonie, are having a game of catch. Julianne throws a baseball to Bryonie, who is standing 8.0 m away. Bryonie catches the ball at the same height, 2.0 s after it is thrown. The mass of the baseball is 250 g. Ignore the effects of air resistance.

a Determine the value of the maximum height gained by the ball during its flight.
b What was the acceleration of the ball at its maximum height?
c Calculate the speed at which the ball was thrown.

29 A skateboarder of mass 55 kg is practising on a half-pipe of radius 2.0 m. At the lowest point of the half-pipe, the speed of the skater is $6.0\,m\,s^{-1}$.

a What is the acceleration of the skater at this point?

b Calculate the size of the normal force acting on the skater at this point.

c Describe the apparent weight of the skater as they travel through the lowest point in the pipe.

30 A car racing track is banked so that when the cars corner at $40\,m\,s^{-1}$, they experience no sideways frictional forces. The track is circular with a radius of 150 m.

a In the diagram below, the car is travelling at $40\,m\,s^{-1}$. Draw and identify the forces that are acting on the car in the vertical plane at this instant.

b Calculate the angle to the horizontal at which the track is banked.

31 In the Gravitron ride, the patrons enter a cylindrical chamber which rotates rapidly, causing them to be pinned to the vertical walls as the floor drops away. A particular Gravitron ride has a radius of 5.00 m and rotates with a period of 2.50 s. Jodie, of mass 60.0 kg, is on the ride.

a Choose the correct responses in the following statement from the options given in bold:
As the Gravitron spins at a uniform rate and Jodie is pinned to the wall, the horizontal forces acting on her are **balanced/unbalanced** and the vertical forces are **balanced/unbalanced**.

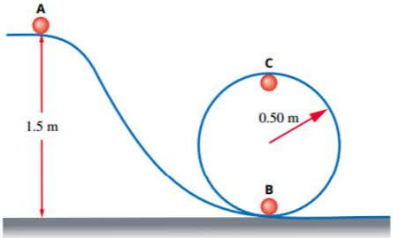
b Calculate the speed of Jodie as she revolves on the ride.

c What is the magnitude of her centripetal acceleration?

d Calculate the magnitude of the normal force that acts on Jodie from the wall of the Gravitron.

e The rate of rotation of the ride is increased so that Jodie completes six revolutions every 10.0 s. What is the frequency of Jodie's motion now?

32 A ball bearing of mass 25 g is rolled along a smooth track in the shape of a loop-the-loop. The ball bearing is given a launch speed at A so that it just maintains contact with the track as it passes through point C. Ignore drag forces when answering these questions.



a Determine the magnitude of the acceleration of the ball bearing as it passes point C.

b How fast is the ball bearing travelling at point C?

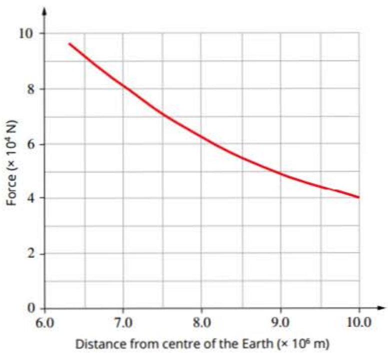
c What is the apparent weight of the ball bearing at point C?

d How fast is the ball bearing travelling at point B?

e Describe the transformations of energy as the ball bearing travels through the system.

f Create a graph of the gravitational potential energy over time if point A is 1.5 m off the ground. You can use arbitrary units for the time scale. Make note of where the ball passes through points A, B and C.

33 A 10 000 kg spacecraft is drifting directly towards the Earth. When it is at an altitude of 600 km, its speed is $1.5\,km\,s^{-1}$. The radius of the Earth is 6400 km. The following graph shows the force on the spacecraft against distance from the Earth.



a How much gravitational potential energy would the spacecraft lose as it falls to a distance of 6500 km?

b Determine the speed of the spacecraft at a distance of 6500 km.

c What is the weight of the spacecraft when it is at an altitude of:

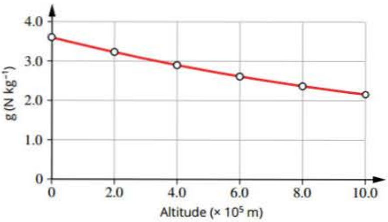
i 3600 km

ii 6.0×10^5 m?

d How does the acceleration of the spacecraft change as it moves from an altitude of 600 km to an altitude of 100 km? Include numerical data in your answer.

e There is a point between the Earth and the Moon where the total gravitational field is zero. The significance of this is that returning lunar missions are able to return to Earth under the influence of the Earth's gravitational field once they pass this point. Given that the mass of Earth is 6.0×10^{24} kg, the mass of the Moon is 7.3×10^{22} kg and the radius of the Moon's orbit is 3.8×10^8 m, calculate the distance of this point from the centre of the Earth.

34 A small asteroid has just smashed into the surface of Mars and a lump of Martian rock of mass 20 kg has been thrown into space with 40 MJ of kinetic energy. A graph of gravitational field–distance from the surface of Mars is shown below.



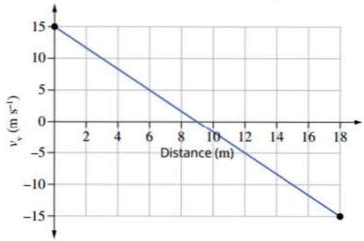
a What is the gravitational force acting on the Martian rock when it is at an altitude of 300 km?

b How much kinetic energy (in MJ) does the rock lose as it travels from the surface of Mars to an altitude of 6.0×10^5 m?

c The rock eventually comes to a stop and starts to fall back towards Mars. Explain how you would determine the altitude at which the rock stopped.

d In your own words, explain the difference between the terms *weight* and *apparent weight*, giving an example of a situation where the magnitudes of these two forces would be different.

35 A ball is kicked towards goal posts a distance of 26 m away. The vertical velocity of the ball is given as a function of the horizontal distance travelled in the figure below.



a What is the horizontal distance when the ball reaches maximum height?

b The ball is in the air for 3.0 s. What is its horizontal velocity?

c What is the ball's launch angle?

d What is the maximum height of the projectile?

e The ball is kicked a second time. The magnitude of the launch velocity is kept the same, but now the launch angle is decreased to 50°. Create a graph of vertical velocity as a function of horizontal distance. Does the ball reach the goal posts?

MODULE 6 Electromagnetism

Discoveries about the interactions that take place between charged particles and electric and magnetic fields not only produced significant advances in physics, but also led to significant technological developments. These developments include the generation and distribution of electricity, and the invention of numerous devices that convert electrical energy into other forms of energy.

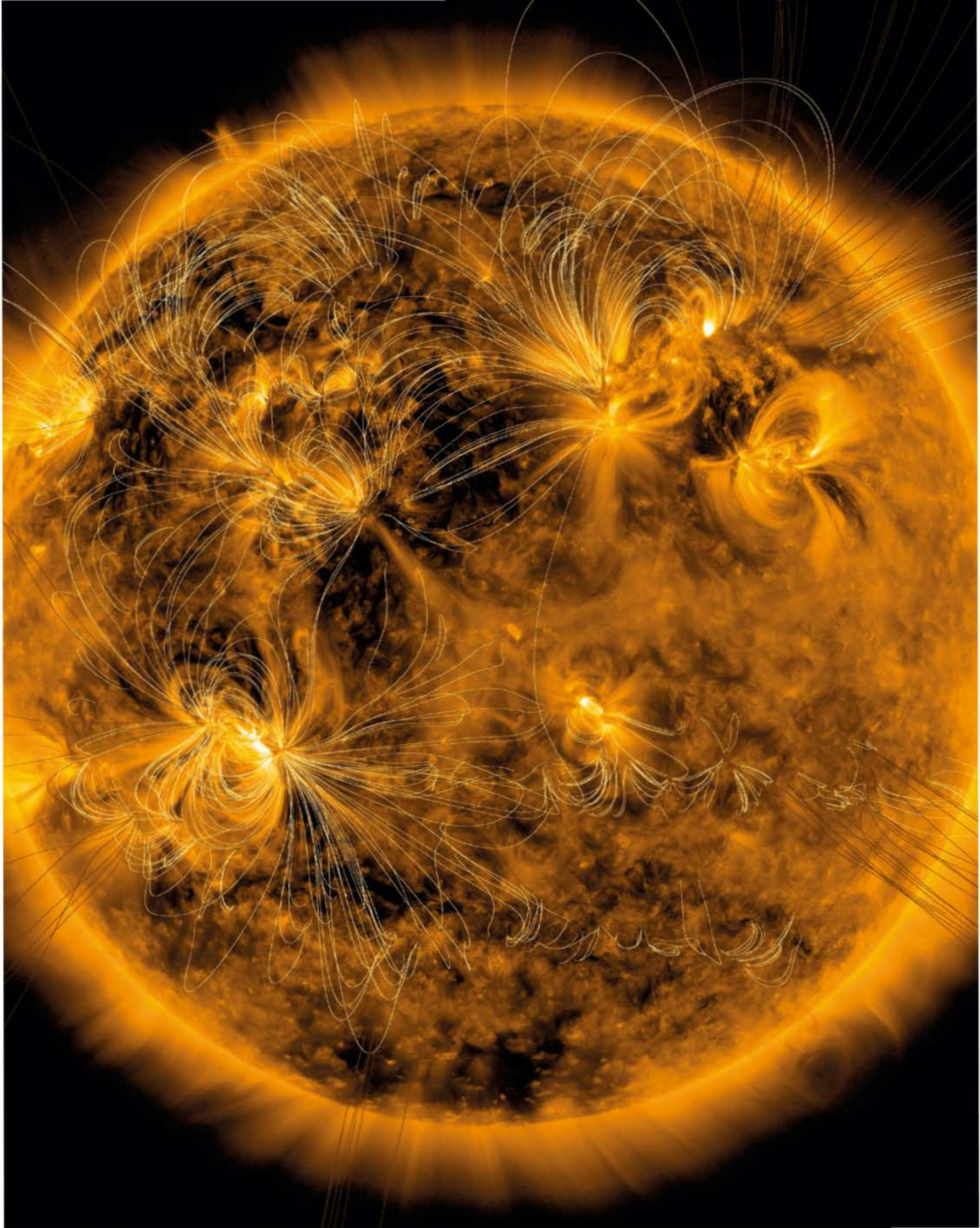
Understanding the similarities and differences in the interactions of single charges in electric and magnetic fields provides you with a conceptual foundation for this module. Phenomena that include the force produced on a current-carrying wire in a magnetic field, the force between current-carrying wires, Faraday's law of electromagnetic induction, the principles of transformers, and the workings of motors and generators can all be understood as instances of forces acting on moving charged particles in magnetic fields.

The law of conservation of energy underpins all of these interactions. The conversion of energy into forms other than the intended form is a problem that constantly drives engineers to improve designs of electromagnetic devices.

Outcomes

By the end of this module you will be able to:

- develop and evaluate questions and hypotheses for scientific investigation PH12-1
- design and evaluate investigations in order to obtain primary and secondary data and information PH12-2
- conduct investigations to collect valid and reliable primary and secondary data and information PH12-3
- select and process appropriate qualitative and quantitative data and information using a range of appropriate media PH12-4
- analyse and evaluate primary and secondary data and information PH12-5
- explain and analyse the electric and magnetic interactions due to charged particles and currents, and evaluate their effect both qualitatively and quantitatively PH12-13



CHAPTER

05 Charged particles, conductors, and electric and magnetic fields

Studying the interaction between charged particles, conductors, and electric and magnetic fields has led to significant advances in technology, including the discovery and commercial realisation of electric power generation and distribution. These discoveries have transformed society.

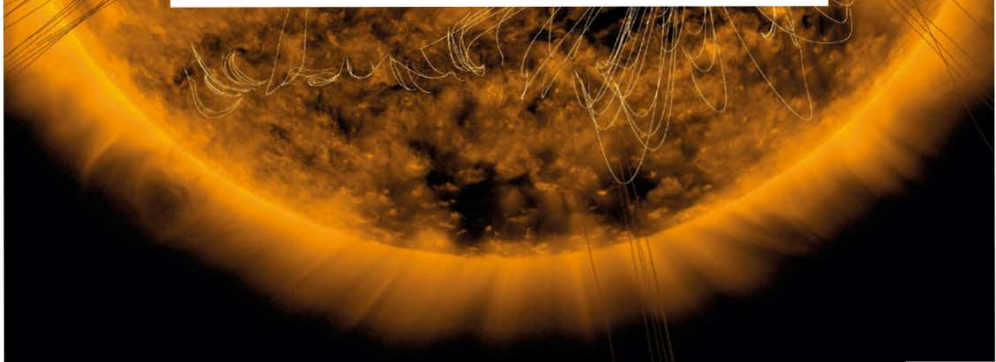
Content

PHYSICS INQUIRY

What happens to stationary and moving charged particles when they interact with an electric or magnetic field?

By the end of this chapter you will be able to:

- investigate and quantitatively derive and analyse the interaction between charged particles and uniform electric fields, including the: (ACSPH083) ICT N
  - electric field between parallel charged plates $\left(E = \frac{V}{d}\right)$
  - acceleration of charged particles by the electric field $\left(\vec{F}_{\text{net}} = m\vec{a}, \vec{F} = q\vec{E}\right)$
  - work done on the charge $\left(W = qV,\ W = qEd,\ K = \frac{1}{2}mv^2\right)$
- model qualitatively and quantitatively the trajectories of charged particles in electric fields and compare them with the trajectories of projectiles in a gravitational field CCT ICT L N
- analyse the interaction between charged particles and uniform magnetic fields, including: (ACSPH083)
  - acceleration, perpendicular to the field, of charged particles
  - the force on the charge $(F = qv_{\perp}B = qvB\sin\theta)$ N
- compare the interaction of charged particles moving in magnetic fields to: CCT
  - the interaction of charged particles with electric fields
  - other examples of uniform circular motion (ACSPH108).



5.1 Particles in electric fields

REVISION

Electric fields

Any charged object has a region of space around it—an **electric field**—where another charged object will experience a force. This is one aspect of the electromagnetic force. Unlike gravity, which only exerts an attractive force, electric fields can exert forces of attraction or repulsion.

An electric field surrounds positive and negative charges, and exerts a force on other charges within the field. The charge produced by electrons and protons is listed in Table 5.1.1. The electric field around charged objects can be represented by field lines.

TABLE 5.1.1 Charge and mass of protons and neutrons

	Charge (C)	Mass (kg)
proton	$+1.602 \times 10^{-19}$	1.675×10^{-27}
electron	-1.602×10^{-19}	9.109×10^{-31}

PHYSICS INQUIRY CCT N

Charged interactions

What happens to stationary and moving charged particles when they interact with an electric or magnetic field?

COLLECT THIS...

- aluminium pie dish
- Styrofoam cup (or other handle made from insulating material)
- tinsel, tied in a loop
- wool fabric
- piece of Styrofoam with a flat side as large as the pie dish

DO THIS...

1. Sticky tape or hot glue the Styrofoam cup to the inside of the pie dish to act as a handle.
2. Rub the flat face of the piece of Styrofoam with the wool fabric until there is a build-up of charge that causes the hairs on your arm to stand up when they are brought near. Do not touch the Styrofoam with your body.
3. Holding the handle, place the metal pie dish on the charged Styrofoam. Then briefly touch the metal pie dish with your finger. After this, do not touch the metal with your body
4. Hold the dish up using the handle, and drop a piece of looped tinsel on the dish from above. The tinsel will fly up once it touches the metal so move your hand out of the way quickly.

RECORD THIS...

Describe the motion of the tinsel, including the average height above the pie dish.

Present a free-body diagram of the tinsel. Use this diagram to calculate the magnitude of the charge on the tinsel and pie dish, assuming they are equal.

REFLECT ON THIS...

What happens to stationary and moving charged particles when they interact with an electric field?

What improvements could be made to the activity?

WS 6.1

There are four fundamental forces in nature that act at a distance. That is, they can exert a force on an object without making any physical contact with it. These are called non-contact forces, and include the strong nuclear force, the weak nuclear force, the electromagnetic force and the gravitational force. In Chapter 4 you explored the motion of objects in a gravitational field; this section explores the way charged particles interact in an electric field.

You will see vector notation used throughout this module. For instance, the electric field strength will be represented as either E or $\vec{E}$ and the force can be shown as either F or $\vec{F}$. It is an important skill for you to determine whether direction needs to be included in your analysis.

ELECTRIC FIELD LINES

An electric field is a vector quantity, which means it has both direction and a magnitude. In order to visualise electric fields around charged objects you can use electric **field lines**. Some field lines are already visible—for example the girl's hair in Figure 5.1.1 is tracing out the path of the field lines. Diagrams of field lines can also be constructed.

Field lines are drawn with arrowheads on them indicating the direction of the force that a small positive test charge would experience if it were placed in the electric field. Therefore, field lines point away from positively charged objects and towards negatively charged objects. Usually, only a few representative lines are drawn.

The density of field lines (how close they are together) is an indication of the relative strength of the electric field. Figure 5.1.2 shows some examples of how to draw electric field lines.

FIGURE 5.1.1 The girl's hair follows the lines of the electric field produced when she became charged while sliding down a plastic slide.

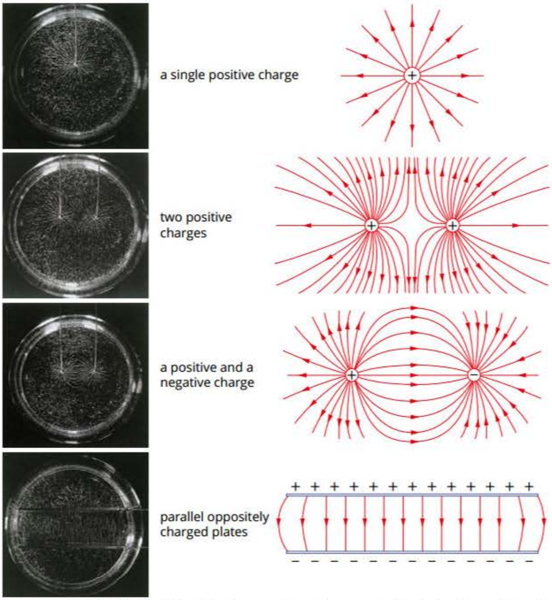


FIGURE 5.1.2 On the left-hand side of the figure, grass seeds suspended in oil align themselves with the electric field. The diagram to the right of each photo shows lines representing the electric field.

FORCES ON FREE CHARGES IN ELECTRIC FIELDS

If a charged particle, such as an electron, were placed within an electric field, it would experience a force. The direction of the field and the sign of the charge allow you to determine the direction of the force.

Figure 5.1.3 shows a positive test charge (proton) and a negative test charge (electron), within a uniform electric field. The direction of an electric field is defined as the direction of the force that a positive charge would experience within the electric field. Therefore a proton will experience a force in the same direction as the field, whereas an electron will experience a force in the opposite direction to the electric field.

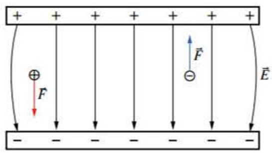


FIGURE 5.1.3 The direction of the electric field ($\vec{E}$) indicates the direction in which a force would act on a positive charge. A negative charge would experience a force in the opposite direction to the field.

Electric field strength can be thought of as the force applied per coulomb of charge. The force experienced by a charged particle due to an electric field can be determined using the equation:

$\vec{F} = q\vec{E}$

where

$\vec{F}$ is the force on the charged particle (N)

q is the charge of the object experiencing the force (C)

$\vec{E}$ is the strength of the electric field ($N\,C^{-1}$)

This equation illustrates that the force experienced by a charge is proportional to the strength of the electric field, $\vec{E}$, and the size of the charge, q. The force on the charged particle will cause the charged particle to accelerate in the field. This means that the particle could increase its velocity, decrease its velocity, or change its direction while in the field.

To calculate the acceleration due to the force experienced, you can use Newton's second law which relates the net force $\vec{F}_{net}$ to the mass m and the acceleration $\vec{a}$:

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

Worked example 5.1.1

THE ACCELERATION OF A CHARGED PARTICLE

Calculate the magnitude of the acceleration experienced by an electron travelling in a uniform electric field of strength $3.1 \times 10^{-5}\ N\,C^{-1}$.
(Use $q_e = -1.602 \times 10^{-19}$ C and $m_e = 9.109 \times 10^{-31}$ kg.)

Thinking	Working
Recall the formula for the force experienced by a charged particle due to an electric field.	$\vec{F} = q\vec{E}$
Substitute the values for electric field strength and the charge of the electron, q_e.	$\vec{F} = q\vec{E}$ $= (-1.602 \times 10^{-19}) \times (3.1 \times 10^{-5})$ $= 4.97 \times 10^{-24}$ N (in the opposite direction to the field)
Recall Newton's second law and solve for a. You are looking for the magnitude of the acceleration so a direction is not needed.	$\vec{F}_{net} = m\vec{a}$ $\vec{a} = \frac{\vec{F}_{net}}{m}$ $= \frac{4.97 \times 10^{-24}}{9.109 \times 10^{-31}}$ $a = 5.5 \times 10^{6}\ m\,s^{-2}$

Worked example: Try yourself 5.1.1

THE ACCELERATION OF A CHARGED PARTICLE

Calculate the magnitude of the acceleration experienced by an electron travelling in a uniform electric field of strength $5 \times 10^{-6}\ N\,C^{-1}$.
(Use $q_e = -1.602 \times 10^{-19}$ C and $m_e = 9.109 \times 10^{-31}$ kg.)

WORK DONE IN UNIFORM ELECTRIC FIELDS

Electrical potential energy is a form of energy that is stored in an electric field. Work is done on the field when a charged particle is forced to move in the electric field. Conversely, when energy is stored in the electric field then work can be done by the field on the charged particle.

Electrical potential (V) is defined as the work required per unit charge to move a positive point charge from infinity to a place in the electric field. The electrical potential at infinity is defined as zero. This definition leads to the equation $V = \frac{W}{q}$.

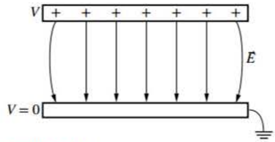


FIGURE 5.1.4 The potential of two plates when one has a positive potential and the other is earthed.

> $W = qV$
>
> where
>
> W is the work done on a positive point charge or on the field (J)
>
> q is the charge of the point charge (C)
>
> V is the electrical potential ($J\,C^{-1}$) or volts (V)

You may come across the equation for the electric field strength in the form of $E = -\frac{V}{d}$. In this version of the equation, the negative value is introduced as the electric field E and is always from the higher potential towards the lower potential, so that the change in potential difference ($V = V_f - V_i$) becomes a negative value. For the equation used in this book, the potential V is given as the magnitude of the potential difference, i.e. a positive value. For example, in a parallel plate experiment one plate is at an electrical potential of 15 V and the other plate is earthed (i.e. 0 V). The potential difference is then equal to $V = 15\,V$.

Consider two parallel plates, as shown in Figure 5.1.4, in which the positive plate is at a potential (V) and the other plate is earthed, which is defined as zero potential. The difference in potential between these two plates is called the electrical **potential difference** (V).

An alternative measure of the electric field strength in volts per metre is calculated using the equation:

> $E = \frac{V}{d}$
>
> where
>
> E is the electric field strength ($V\,m^{-1}$)
>
> V is the potential difference (V)
>
> d is distance (m)

Between any two points in an electric field, E, separated by a distance, d, that is parallel to the field, the potential difference, V, is then defined as the change in the electrical potential between these two points (Figure 5.1.5).

CALCULATING WORK DONE

You can now derive an equation for calculating the work done on a point test charge to move it a distance across a potential difference. Combining the equations $W = qV$ and $V = Ed$ gives:

> $W = qEd$
>
> where
>
> W is the work done on the point charge or on the field (J)
>
> q is the charge of the point charge (C)
>
> E is the magnitude of the electrical field strength ($V\,m^{-1}$ or $N\,C^{-1}$)
>
> d is the distance between points, parallel to the field (m)

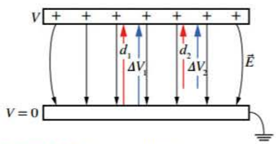


FIGURE 5.1.5 The potential difference between two points in a uniform electric field. The change in potential ΔV is dependent on a given distance d, so that in a uniform electric field, $E = \frac{\Delta V_1}{d_1} = \frac{\Delta V_2}{d_2}$.

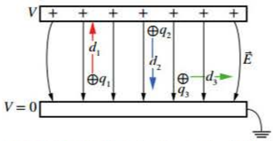


FIGURE 5.1.6 Work is being done on the field by moving q_1, and work is being done by the field on q_2. No work is done on or by q_3 since it is moving perpendicular to the direction of the field.

Work done by or on an electric field

When calculating work done, which changes the electrical potential energy, remember that work can be done either:

- by the electric field on a charged object, or
- on the electric field by forcing the object to move.

You need to examine what's happening in a particular situation to know how the work is being done.

For example, if a charged object is moving in the direction it would naturally tend to go within an electric field, then work is done by the field. So when a positive point charge is moved in the direction of the electric field, the electric field has done work on the point charge; see, for example, q_2 in Figure 5.1.6.

When work is done by a charged object on an electric field, the object is forced to move against the direction it would naturally go. Work has been done on the field by forcing the object to move against the field. For example, if a force causes a positive charge to move towards the positive plate within a uniform electric field, work has been done on the electric field by forcing the object to move. (See q_1 in Figure 5.1.6.)

If a charge doesn't move any distance parallel to the field then no work is done on or by the field. For q_3 in Figure 5.1.6, the charge has moved perpendicularly to the field so no work is done.

Worked example 5.1.2

WORK DONE ON A CHARGE IN A UNIFORM ELECTRIC FIELD

A student sets up a parallel plate arrangement so that one plate is at a potential of 12.0 V and the other earthed plate is positioned 0.50 m away. Calculate the work done to move a proton a distance of 10.0 cm towards the negative plate. ($q_p = +1.602 \times 10^{-19}$ C)
In your answer identify what does the work and what the work is done on.

Thinking	Working
Identify the variables presented in the problem to calculate the electric field strength E.	$V_2 = 12.0$ V $V_1 = 0$ V $V = 12.0 - 0 = 12.0$ V $d = 0.50$ m
Use the equation $E = \frac{V}{d}$ to determine the electric field strength.	$E = \frac{V}{d}$ $= \frac{12.0}{0.50}$ $= 24.0\ \text{V m}^{-1}$
Use the equation $W = qEd$ to determine the work done. Note that d here is the distance that the proton moves.	$W = qEd$ $= 1.602 \times 10^{-19} \times 24.0 \times 0.100$ $= 3.84 \times 10^{-19}$ J
Determine if work is done on the charge by the field or if work is done on the field by the charge.	As the positively charged proton is moving naturally towards the negative plate, work is done on the proton by the field.

Worked example: Try yourself 5.1.2

WORK DONE ON A CHARGE IN A UNIFORM ELECTRIC FIELD

A student sets up a parallel plate arrangement so that one plate is at a potential of 36.0 V and the other earthed plate is positioned 2.00 m away. Calculate the work done to move an electron a distance of 75.0 cm towards the negative plate.
($q_e = -1.602 \times 10^{-19}$ C)
In your answer identify what does the work and what the work is done on.

SKILLBUILDER N

Using units in an equation to check for dimensional consistency

Scientists know that each term in an equation stands for a particular quantity and the position of the term in the equation tells scientists where that quantity should go. The units used to measure that quantity are not used in the calculations. Units are only indicated on the final line of the solved equation. For example, this is the equation for the area (*A*) of a rectangle of length (*L*) and width (*W*):

$$A = L \times W$$

If L has a value of 7 m and W has a value of 4 m, it is written:

$$\begin{aligned} A &= L \times W \\ &= 7 \times 4 \\ &= 28\,\text{m}^2 \end{aligned}$$

You can use units to check the dimensional consistency of the answer. In the example above, the two quantities of *L* (length) and *W* (width) both have to be expressed in consistent units, in this case metres (m), to give an answer that is expressed in square metres ($\text{m} \times \text{m} = \text{m}^2$).

If you had made a mistake, and used the formula A = L + W instead, the answer would be expressed in metres only. This is not the correct unit to express area, so you would know that was wrong.

Similarly, you can equate different units using dimensional analysis. For example, if you combine the equations for the magnitude of the electric field $E = \frac{V}{d}$ and $E = \frac{F}{q}$, you end up with $\frac{V}{d} = \frac{F}{q}$. Looking at the units for each side of this equation, $\text{V}\,\text{m}^{-1}$ must equal $\text{N}\,\text{C}^{-1}$.

PHYSICSFILE CCT

Gravitational force and electric force

Oppositely charged parallel plates can be arranged one above the other, such that the electric field is vertical between the two plates. The direction of the field can then be manipulated to create an upward force on a charged particle in the field.

If the electric force created by the field on the charged object is equal to the gravitational force on (or weight of) the object, then these two forces can add to provide a net force equal to zero. This means that the charged object will either be suspended between the plates, or (by Newton's first law of motion) will be falling or rising at constant velocity.

This phenomenon was used by Robert Millikan and his PhD student Harvey Fletcher in their oil-drop experiment, performed in 1909, to determine the fundamental charge of an electron to within 1% of the currently accepted value. This will be discussed in greater detail in Chapter 14.

COMPARING GRAVITATIONAL AND ELECTRIC FIELDS

Many of the forces affecting us and the world around us can be described as contact forces. There is direct contact as you open a door, kick a ball or rest on a couch. By contrast, the forces of gravity, magnetism and electricity act over a distance without necessarily having any physical contact (Figure 5.1.7). This was a difficult idea for scientists to come to terms with. Newton still had some misgivings, even when publishing his ideas of universal gravitation. The concept of fields, used to explain how and why forces can act over a distance, is a very powerful tool and one that has allowed us to better explain the fundamental forces of gravity and electromagnetism.

Gravity is an incredible force. Permeating the universe, it brings gas clouds together to form planets, stars and galaxies. It causes stars to collapse to black holes, generating gravitational fields strong enough that even light can't escape. And yet the gravitational force of attraction between two electrons is less than 8×10^{-37} N, which is the same as the electrostatic repulsion between the same two electrons.

The relationships developed for gravitational and electric fields over the last two chapters reveal the parallels and differences between related field concepts for gravitational masses and point charges, both of which are essentially **monopoles**. They are summarised in Table 5.1.2.

FIGURE 5.1.7 The magnet has an effect on the paper clips even though they are not in contact. This is because the paper clips are within the magnetic field produced by the magnet.

A magnet must always be a **dipole**, i.e. it must have two poles—a north and a south. Electric and gravitational fields, on the other hand, can be generated by a monopole (meaning one pole), such as having a single point charge or one gravitational mass.

TABLE 5.1.2 Comparison of gravitational and electric fields.

Quantity or description	Gravitational fields	Electrical fields
how field strength varies with distance, *r*, from a monopole	$g = G\frac{M}{r^2}$	$E = k\frac{q}{r^2}$
force between monopoles	$F_g = G\frac{Mm}{r^2}$	$F = k\frac{q_1q_2}{r^2}$
potential energy changes in a uniform field	$\Delta U = mg\Delta h$	$W = qV$
force due to a uniform field	$F_g = mg$	$F = qE$

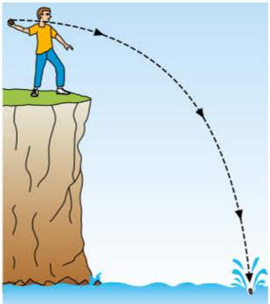


FIGURE 5.1.8 An object thrown horizontally will follow a parabolic path due to the vertical acceleration due to gravity.

PROJECTILE MOTION

In Chapter 2, the projectile motion of objects moving in a gravitational field was discussed. An object with some initial horizontal velocity and a vertical acceleration (due to gravity) will follow a parabolic path (Figure 5.1.8). A similar process will occur for charged particles moving through a potential difference.

Consider an electric field acting on an electron as the result of a pair of oppositely charged parallel plates connected to a DC power supply (Figure 5.1.9 on page 153). The electron is attracted to the positive plate and repelled from the negative plate. An electric field is acting upon any charged particle within this region. This electric field is a vector quantity and may be compared in some ways to the Earth's gravitational field. Recall an electric field has units $N\,C^{-1}$ and is defined as:

$$E = \frac{F}{q}$$

where F is the force (N) experienced by a charged particle due to an electric field (E), and q is the magnitude of the electric charge of a particle in the field, in this case an electron ($q_e = 1.6 \times 10^{-19}$ C).

A charge will then experience a force equal to qE when placed within such an electric field.

Recall that the magnitude of the electric field may also be expressed as:

$$E = \frac{V}{d}$$

where d is the separation of the plates (m) and V is the potential difference (V). Combining these two relationships produces an expression for the force on a charge within a pair of parallel charged plates:

$$\frac{F}{q} = \frac{V}{d}$$

$$F = \frac{qV}{d}$$

In addition, calculations of the energy gained by an electron as it is accelerated towards a charged plate by the electric field can be made. The work done in this case is equivalent to:

$$W = qV$$

This equation can be used to calculate the increase in kinetic energy as an electron accelerates from one plate to another.

If a charge is accelerated from rest from an electron gun, then:

$$\Delta K = W = qV$$

$$\Delta K = \frac{1}{2}mv^2 - \frac{1}{2}mu^2$$

where v is the final speed and u is the initial speed of the charge. If the electron accelerates from rest ($u = 0$), then this can be simplified to:

$$K = \frac{1}{2}mv^2 = qV$$

An **electron gun** uses a heated cathode to produce an electron beam and a series of charged plates to accelerate the beam.

$\frac{1}{2}mv^2 = qV$

This is often referred to as the electron-gun equation.

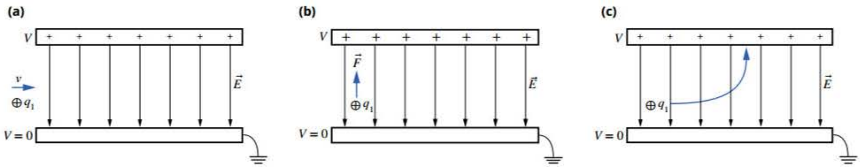


FIGURE 5.1.9 (a) The electron approaches the two oppositely charged parallel plates with a horizontal velocity, v. (b) The electron undergoes a vertical acceleration due to the electrostatic force, $\vec{F}$, that forces the electron towards the positively charged plate. (c) The charge follows a parabolic path, due to its horizontal velocity and its vertical acceleration from the electrostatic force ($\vec{F}$), in the same way that a projectile moves under gravity.

In Chapter 2, it was shown how you can use the equations of motion to predict how a projectile will behave in a gravitational field by breaking the vectors of motion down into their components. The same can be done for a charged particle in an electric field.

GO TO ➤ Section 2.1 page 50

Recall the equations of motion:

$v = u + at$

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

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

$s = vt - \frac{1}{2}at^2$

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

where

v is the final velocity ($m\,s^{-1}$)
u is the initial velocity ($m\,s^{-1}$)
a is the acceleration ($m\,s^{-2}$)
t is the time taken (s)
s is the displacement (m)

SKILLBUILDER N

The quadratic equation

Sometimes you may come across a quadratic expression, such as $ax^2 + bx + c = 0$, where you will need to solve for x. To do this you will need to know the quadratic equation.

The quadratic equation takes the form:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

For example, when using the equations of motion, you might need to find the time taken by using the formula: $s = ut + \frac{1}{2}at^2$. In this example, to solve for t, follow these steps:

1 Rearrange the expression so it is equal to zero:

$$0 = \frac{1}{2}at^2 + ut - s$$

2 Pick out the values for a, b and c by comparing them to the original quadratic expression:

$$a = \frac{1}{2}a,\ b = u \text{ and } c = -s$$

3 Substitute these values into the quadratic formula:

$$t = \frac{-u \pm \sqrt{u^2 - 4 \times \frac{1}{2}a \times -s}}{2 \times \frac{1}{2}a} = \frac{-u \pm \sqrt{u^2 + 2as}}{a}$$

While the quadratic formula can give you two results (due to the ± sign), only take the answer that makes physical sense. In this example, it wouldn't be possible to have a negative answer for time, so only one result will be found.

Worked example 5.1.3

PROJECTILE MOTION OF A CHARGE IN AN ELECTRIC POTENTIAL

Two parallel plates are separated by 1.0 cm. One plate is earthed and the other has a potential of 5.0 kV.

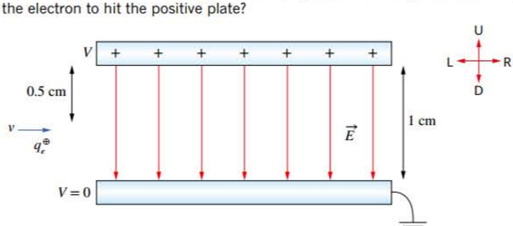
a Determine the final speed of a single electron when accelerating from rest across this potential difference.
($q_e = -1.602 \times 10^{-19}$ C, $m_e = 9.109 \times 10^{-31}$ kg)

Thinking	Working
Write down the variables that are given using appropriate units.	$u = 0\ \text{m s}^{-1}$ $q = 1.602 \times 10^{-19}$ C $m = 9.109 \times 10^{-31}$ kg $V = 5.0 \times 10^{3}$ V $v = ?$
Recall the electron-gun equation.	$\frac{1}{2}mv^2 = qV$
Solve for *v*.	$v^2 = \frac{2qV}{m}$ $= \frac{2 \times 1.602 \times 10^{-19} \times 5.0 \times 10^{3}}{9.1 \times 10^{-31}}$ $= 1.7 \times 10^{15}$ $v = 4.2 \times 10^{7}\ \text{m s}^{-1}$

b Calculate the acceleration of the electron.

Thinking	Working
Find the electric field strength.	$E = \frac{V}{d}$ $= \frac{5.0 \times 10^{3}}{1 \times 10^{-2}}$ $= 5.0 \times 10^{5}\ \text{V m}^{-1}$
Combine Newton's second law with the electrostatic force.	$\vec{F}_{net} = m\vec{a}$ $a = \frac{F}{m}$ $= \frac{qE}{m}$
Solve for *a*.	$a = \frac{-1.602 \times 10^{-19} \times 5.0 \times 10^{5}}{9.1 \times 10^{-31}}$ $= 8.8 \times 10^{16}\ \text{m s}^{-2}$ (towards the charged plate)

c Now assume the initial velocity of the electron was $5.0 \times 10^{7}\ \text{m s}^{-1}$ travelling horizontally as it enters halfway between the plates. How long does it take for the electron to hit the positive plate?



Thinking	Working
Write down the known quantities. While the horizontal velocity is equal to $5 \times 10^7\,\text{m s}^{-1}$, the initial vertical velocity is equal to zero. Take up and right to be positive.	Vertically, $s = 0.5\,\text{cm} = +0.005\,\text{m}$ (half the distance between plates) $u = 0\,\text{m s}^{-1}$ $a = 8.8 \times 10^{16}\,\text{m s}^{-2}$ $t = ?$
Identify the correction equation to use.	$s = ut + \frac{1}{2}at^2$
Substitute the known values and rearrange the expression so it equals zero.	$0.005 = 0 \times t + \frac{1}{2} \times 8.8 \times 10^{16} \times t^2$ $0 = 4.4 \times 10^{16} \times t^2 + 0t - 0.005$
Use the quadratic equation to solve for *t*.	$t = \frac{-0 \pm \sqrt{0^2 - 4 \times 4.4 \times 10^{16} \times -0.005}}{2 \times 4.4 \times 10^{16}}$ $= 3.4 \times 10^{-10}\,\text{s}$

d Assuming the same conditions as for part c, calculate the final horizontal displacement of the electron.

Thinking	Working
Write down the known quantities.	Horizontally, $u = +5.0 \times 10^7\,\text{m s}^{-1}$ $a = 0\,\text{m s}^{-2}$ $t = 3.4 \times 10^{-10}\,\text{s}$ $s = ?$
Identify the correct equation to use.	$s = ut + \frac{1}{2}at^2$
Solve for *s*.	$s = 5.0 \times 10^7 \times 3.4 \times 10^{-10} + \frac{1}{2} \times 0 \times (3.4 \times 10^{-10})^2$ $= 0.17$ $= 1.7\,\text{cm}$ to the right

Worked example: Try yourself 5.1.3

PROJECTILE MOTION OF A CHARGE IN AN ELECTRIC POTENTIAL

Two parallel plates are separated by 2.0 cm. One plate is earthed and the other has a potential of 3.0 kV.

a Determine the final speed of a single electron when accelerating from rest across this potential difference.
($q_e = -1.602 \times 10^{-19}\,\text{C}$, $m_e = 9.1 \times 10^{-31}\,\text{kg}$)

b Calculate the acceleration of the electron.

c Now assume the initial velocity of the electron was $2.0 \times 10^7\,\text{m s}^{-1}$ travelling horizontally. How long does it take for the electron to hit the positive plate? Use the same direction conventions given in Worked example 5.1.3c.

d Assuming the same conditions as for part c, calculate the horizontal displacement of the electron.

5.1 Review

SUMMARY

- An electric field is a region of space around a charged object in which another charged object will experience a force.
- Electric fields are represented using field lines. The spacing between the field lines indicates the strength of the field. The closer together the lines are, the stronger the field.
- Electric field lines point in the direction of the force that a positive charge within the field would experience.
- Between two oppositely charged parallel plates, the field lines are parallel and therefore the field has a uniform strength.
- Electric field strength can be expressed as $E = \frac{F}{q}$ or $E = \frac{V}{d}$.
- A positive charge experiences a force in the direction of the electric field and a negative charge experiences a force in the opposite direction to the field.
- The force on a charged particle can be determined using the equation $\vec{F} = q\vec{E}$.
- Force can be related to the acceleration of a particle using Newton's second law: $\vec{F}_{net} = m\vec{a}$.
- When a charged object is moved against the direction it would naturally move in an electric field, then the charged object does work on the field.
- When a charged object is moved in the direction it would naturally tend to move in an electric field, then the field does work on the charged object.
- The work done on or by an electric field can be calculated using the equations $W = qV$ or $W = qEd$.
- The direction of a field at any point is always the resultant field vector determined by adding the individual field vectors due to each mass, charge or magnetic pole within the affected region.
- In a static (unchanging) field, the strength of the field doesn't change with time.
- The motion of a charged particle in a field can be modelled using the equations of motion. This is analogous to projectile motion within a gravitational field.

KEY QUESTIONS

1 Which of the following options correctly describes an electric field?
  - **A** a region around a charged object that causes a charge on other objects within that region
  - **B** a region around a charged object that causes a force on other objects within that region
  - **C** a region around a charged object that causes a force on other charged objects in that region
  - **D** a region around an object that causes a force on other objects within that region

2 Which of the following options correctly defines the direction of an electric field?
  - **A** away from a negatively charged object
  - **B** away from a positively charged object
  - **C** away from a neutrally charged object
  - **D** towards a positively charged object

3 Identify whether the rules below for drawing electric field lines are true or false:
  - **a** Electric field lines start and end at 90° to the surface, with no gap between the lines and the surface.
  - **b** Field lines can cross; this indicates that the field is in two directions at that point.
  - **c** Electric fields go from negatively charged objects to positively charged objects.
  - **d** Around small charged spheres called point charges you should draw at least eight field lines: top, bottom, left, right and in between each of these.
  - **e** Around point charges the field lines radiate like spokes on a wheel.
  - **f** Between two point charges the direction of the field at any point is the same as the direction of the field due to the closest of the two point charges.
  - **g** Between two oppositely charged parallel plates the field between the plates is evenly spaced and is drawn straight from the negative plate to the positive plate.

4 Calculate the force applied to a balloon carrying a charge of 5.00 mC in a uniform electric field of 2.50 N C^{-1}.

5 Calculate the acceleration of an electron in a uniform electric field of 3.25 N C^{-1}. The mass of an electron is 9.109×10^{-31} kg and its charge is -1.602×10^{-19} C.

6 Calculate the potential difference that exists between two points separated by 30.0 cm, parallel to the field lines, in an electric field of strength 4000 V m^{-1}.

7 Complete the following statement about the field around a monopole from the options given in bold.

The field around a monopole is **linear/radial**, **static/dynamic** and **uniform/non-uniform**.

8 **a** Calculate the acceleration of an electron through two parallel plates with a 5.0 kV potential difference, separated vertically by a distance of 1.2 cm. Use $q_e = -1.6 \times 10^{-19}$ C and $m_e = 9.1 \times 10^{-31}$ kg.
  - **b** Assume the initial velocity of the electron was 1.5×10^{7} m s^{-1} travelling horizontally. How long does it take for the electron to hit the positive plate?

5.2 Particles in magnetic fields

The Australian Centre for Neutron Scattering (ANSTO) is home to the most powerful synchrotron in the Southern Hemisphere (Figure 5.2.1). Looking something like a giant doughnut about 200 m in circumference, it produces beams of electromagnetic radiation, from infrared, through visible light, to 'hard' X-rays.

FIGURE 5.2.1 View of the inside of ANSTO, taken from the mezzanine.

A **synchrotron** is a type of **particle accelerator**. Bunches of electrons are accelerated around a huge evacuated ring to almost the speed of light with energies as high as 3 billion electron-volts (3×10^9 eV). These charges are forced to follow a curved path, due to the **magnetic field** generated by bending magnets. As they accelerate around curves, the electrons give off bursts of radiation. This synchrotron radiation is channelled down tubes called beamlines and utilised by researchers in a range of experimental stations.

THE EFFECT ON A CHARGED PARTICLE IN A MAGNETIC FIELD

In Figure 5.2.2, a beam of electrons is experiencing a force due to a magnetic field. This apparatus is known as a **cathode ray tube** (CRT). The magnitude of the force is proportional to the strength of the magnetic field, B, the velocity of the charge and the angle θ the object is moving with respect to the magnetic field. A charged particle travelling at a steady speed in a magnetic field experiences the force at an angle to its path and will be diverted. This is the theory behind CRT screens. As the direction of the charged particle changes, so does the angle of the force acting on it. In a very large magnetic field the charged particles will move in a circular path. Mass spectrometers and particle accelerators both work on this principle.

This force is referred to as the **Lorentz force**. The force is at a maximum when the charged particle is moving at right angles to the field. There is no force acting when the charged particles are travelling parallel to the magnetic field.

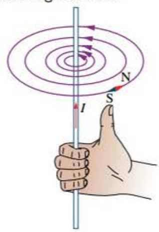
FIGURE 5.2.2 An electron beam being deflected by a magnet.

REVISION

Magnetic fields

Magnetism is a physical phenomenon caused by magnets which results in a magnetic field. Magnetic fields are a vector quantity as they have both a magnitude and a direction. As with electric and gravitational fields, magnetic fields can be represented using field lines. Magnets are always dipolar, so they must have a north and a south pole.

In the 1800s it was discovered that there is a relationship between magnetic fields and electric charge. Hans Christian Ørsted found that an electric current produces a magnetic field. The direction of this field can be found using the right-hand grip rule, where your thumb points in the direction of the conventional current and your fingers point in the direction of the induced magnetic field.



$F = qv_{\perp}B = qvB\sin\theta$

where

F is the force (N)

q is the electric charge on the particle (C)

v is the instantaneous velocity of the particle (m s^{-1})

B is the strength of the magnetic field (T)

θ is the angle the object is moving at with respect to the magnetic field

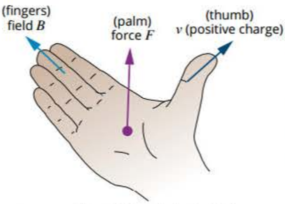


FIGURE 5.2.3 The right-hand rule: Point the thumb of the right hand in the direction of the movement of a positive charge (the direction of conventional current) and the fingers in the direction of the magnetic field. The force on the charge will point outwards from the palm.

Determining the direction of the force

The simple **mnemonic** shown in Figure 5.2.3—the **right-hand rule**—can be used to determine the direction of the force on a charged particle moving in a magnetic field. Using your right hand, with fingers outstretched and flat, point the thumb towards the direction that a positive charge is moving and the outstretched fingers in the direction of the magnetic field. The direction of the resulting force on the charge is the direction in which your palm is facing. The force on a negatively charged particle will therefore be in the *opposite* direction to that on a positively charged particle.

PHYSICSFILE N

The tesla

The unit for the strength of a magnetic field, B, was given the name tesla (T) in honour of Nikola Tesla. Nikola Tesla (1856–1943) was the first person to advocate the use of alternating current (AC) generators for use in town power-supply systems. He was also a prolific inventor of electrical machines of all sorts, including the Tesla coil, a source of high-frequency, high-voltage electricity.

A magnetic field of 1 T is a very strong field. For this reason, a number of smaller units, especially the millitesla (mT), 10^{-3} T, and microtesla (μT), 10^{-6} T, are in common use. Table 5.2.1 shows the strength of some magnets for comparison.

TABLE 5.2.1 Comparison of magnet strengths.

Type of magnet	Strength of magnetic field
very strong electromagnets and 'super magnets'	1 to 20 T
Alnico and ferrite magnets	10^{-2} to 1 T
Earth's surface	5×10^{-5} T

WS 6.2

Worked example 5.2.1

MAGNITUDE OF FORCE ON A POSITIVELY CHARGED PARTICLE

A single positively charged particle with a charge of 1.6×10^{-19} C travels at a velocity of $10\,\text{m}\,\text{s}^{-1}$ at an angle of 30° to a magnetic field of strength 4.0×10^{-5} T.

What is the magnitude of the force the particle will experience from the magnetic field?

Thinking	Working
Establish which quantities are known and which ones are required. All variables are given as scalars as you are looking for the magnitude of the force.	$F = ?$ $q = +1.6 \times 10^{-19}$ C $v = 10\,\text{m}\,\text{s}^{-1}$ $B = 4.0 \times 10^{-5}$ T $\theta = 30°$
Substitute values into the force equation.	$F = qv_{\perp}B = qvB\sin\theta$ $= 1.6 \times 10^{-19} \times 10 \times 4.0 \times 10^{-5} \times \sin 30$
Express the final answer in an appropriate form. Note that only magnitude has been requested so do not include direction.	$F = 3.2 \times 10^{-23}$ N

Worked example: Try yourself 5.2.1

MAGNITUDE OF FORCE ON A POSITIVELY CHARGED PARTICLE

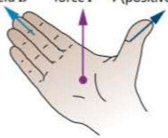
A single positively charged particle with a charge of $+1.6 \times 10^{-19}$ C travels at a velocity of $50\,\text{m s}^{-1}$ perpendicular to a magnetic field of strength 6.0×10^{-5} T.

What is the magnitude of the force the particle will experience from the magnetic field?

Worked example 5.2.2

DIRECTION OF FORCE ON A CHARGED PARTICLE

A single negatively charged particle with a charge of -1.6×10^{-19} C is travelling horizontally out of a computer screen and perpendicular to a magnetic field that runs horizontally from left to right across the screen. In what direction will the force experienced by the charge act?

Thinking	Working
The right-hand rule is used to determine the direction of the force on a positively charged particle.	Align your hand so that your fingers are pointing in the direction of the magnetic field, i.e. left to right and horizontal. If the negatively charged particle is travelling out of the screen, a positively charged particle would be moving in the opposite direction. Align your thumb so it is pointing into the screen, in the direction that a positive charge would travel. Your palm should be facing downwards. That is the direction of the force applied by the magnetic field on the negative charge out of the page. 

Worked example: Try yourself 5.2.2

DIRECTION OF FORCE ON A CHARGED PARTICLE

A single positively charged particle with a charge of $+1.6 \times 10^{-19}$ C is travelling horizontally from left to right across a computer screen and perpendicular to a magnetic field that runs vertically down the screen. In what direction will the force experienced by the charge act?

CIRCULAR MOTION

If a moving charge experiences a force of constant magnitude that remains at right angles to its motion, its direction will be changed but not its speed. In this way, bending magnets within a particle accelerator act to alter the path of the electron beam, rather than to speed the electrons up. As a result, the electrons will follow a curved path of radius r, as shown in Figure 5.2.4.

In this case, the magnitude of the net force acting on the charge is $F_{net} = ma$. This is equivalent to the magnetic force on the charge, so that $qvB = ma$ (recall that the magnitude of the force, F, on a charge, q, moving with speed, v, perpendicular to a magnetic field of strength B is given by $F = qvB \sin 90 = qv_{\perp}B$).

Dots are used to depict a field running directly out of the page, while crosses describe a field running directly into the page.

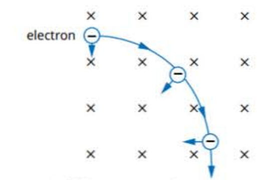


FIGURE 5.2.4 An electron moving in a magnetic field.

GO TO ➤ Section 3.1 page 72

The acceleration in this situation is centripetal (i.e. towards the centre of the circular path) and has magnitude equal to $a_c = \frac{v^2}{r}$. This was shown in Chapter 3.

Substituting this relationship into the previous equation for the force gives:

$$qvB = \frac{mv^2}{r}$$

Rearranging gives an expression that predicts the radius of the path of an electron travelling at right angles to a constant magnetic field:

$$r = \frac{mv}{qB}$$

where

r is the radius of the path (m)

m is the mass of the particle (kg)

v is the speed of the charge ($m\,s^{-1}$)

q is the charge (C)

B is the magnitude of the magnetic field strength (T)

Worked example 5.2.3

CALCULATING SPEED AND PATH RADIUS OF ACCELERATED CHARGED PARTICLES

An electron gun releases electrons which are then accelerated across a potential difference of 32 kV, over a distance of 30 cm between a pair of charged parallel plates. Assume that the mass of an electron is 9.109×10^{-31} kg and the magnitude of the charge on an electron is 1.602×10^{-19} C.

a Calculate the strength of the electric field acting on the electron beam.	
Thinking	**Working**
Ensure that the variables are in their standard units.	$32\,kV = 32 \times 10^3 = 3.2 \times 10^4\,V$ $30\,cm = 0.30\,m$
Apply the correct equation.	$E = \frac{V}{d}$
Solve for *E*. A direction convention is not specified so assume the electric field is in the positive direction.	$E = \frac{3.2 \times 10^4}{0.30}$ $= 1.1 \times 10^5\,V\,m^{-1}$

b Calculate the speed of the electrons as they exit the electron-gun assembly.	
Thinking	**Working**
Apply the correct equation.	$\frac{1}{2}mv^2 = qV$
Rearrange the equation to make *v* the subject.	$v = \sqrt{\frac{2qV}{m}}$
Solve for *v*.	$v = \sqrt{\frac{2 \times 1.602 \times 10^{-19} \times 3.2 \times 10^4}{9.109 \times 10^{-31}}}$ $= 1.1 \times 10^8\,m\,s^{-1}$

c The electrons then travel through a uniform magnetic field perpendicular to their motion. If the field is of strength 0.2 T, calculate the expected radius of the path of the electron beam.	
Thinking	**Working**
Apply the correct equation.	$r = \frac{mv}{qB}$
Solve for *r*.	$r = \frac{9.109 \times 10^{-31} \times 1.1 \times 10^{8}}{1.602 \times 10^{-19} \times 0.2}$ $= 3.1 \times 10^{-3}$ m

Worked example: Try yourself 5.2.3

CALCULATING SPEED AND PATH RADIUS OF ACCELERATED CHARGED PARTICLES

An electron gun releases electrons from its cathode which are accelerated across a potential difference of 25 kV, over a distance of 20 cm between a pair of charged parallel plates. Assume that the mass of an electron is 9.109×10^{-31} kg and the magnitude of the charge on an electron is 1.602×10^{-19} C.

a Calculate the strength of the electric field acting on the electron beam.

b Calculate the speed of the electrons as they exit the electron-gun assembly.

c The electrons then travel through a uniform magnetic field perpendicular to their motion. If the field is of strength 0.3 T, calculate the expected radius of the path of the electron beam.

PHYSICS IN ACTION ICT

Particle accelerators

The ANSTO synchrotron (Figure 5.2.5) accelerates electrons through an equivalent of 3000 million volts (3 GV). At this energy, they travel at 99.99999% of the speed of light. The particles are accelerated by electromagnetic fields, but very long paths are required for the particles to obtain the extremely high speeds needed. To achieve this without the need for tunnels hundreds of kilometres long, particles travel through very strong magnetic fields that cause them to move in a circle. The Australian Synchrotron is 70 m in diameter.

FIGURE 5.2.5 An inside view of the synchrotron at ANSTO.

5.2 Review

SUMMARY

- The force on a charged object within a magnetic field is proportional to the strength of the magnetic field, B, the velocity of the charge, v, the angle the object is moving at with respect to the magnetic field, θ, and the charge on the particle, q, i.e. $F = qv_{\perp}B = qvB\sin\theta$.
- This force is referred to as the Lorentz force.
- The force is:
  - at a maximum when the charged particle is moving at right angles to the magnetic field
  - zero when the charged particle is travelling parallel to the magnetic field.
- The right-hand rule is used to determine the direction of the force on a positive charge moving in a magnetic field, B. The direction of the force on a negatively charged particle is in the opposite direction.
- The magnetic force acts as a centripetal force so that a charged particle moves with circular motion in a magnetic field.
- The radius of the path of a charged particle in a magnetic field can be calculated using the formula $r = \frac{mv}{qB}$.
- Particle accelerators are machines that accelerate charged particles, such as electrons, protons or atomic nuclei, to speeds close to that of light.

KEY QUESTIONS

1 How are particle accelerators able to provide the centripetal acceleration to change the direction of a charged particle using electromagnetic fields?

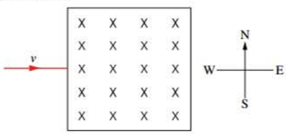
A Charged particles are part of the electromagnetic spectrum.

B Charged particles experience a force from the magnetic field that is proportional to the particle's velocity, constantly accelerating the charged particle.

C The accelerator is curved around the magnetic field.

D Charged particles will always accelerate when placed in a vacuum.

2 An electron with a charge magnitude of 1.6×10^{-19} C is moving eastwards into a magnetic field of strength $B = 1.5 \times 10^{-5}$ T acting into the screen, as shown below. If the magnitude of the initial velocity is $1.0\,\text{m}\,\text{s}^{-1}$, what is the magnitude and direction of the force it initially experiences as it enters the magnetic field?



3 An electron travelling at a speed of $7.0 \times 10^{6}\,\text{m}\,\text{s}^{-1}$ passes through a magnetic field of strength 8.6×10^{-3} T. The electron moves at right angles to the field.

a Calculate the magnitude of the force exerted on the electron by the magnetic field.

b Given that this force directs the electron in a circular path, calculate the radius of its motion.

4 A particle accelerator uses magnetic fields to accelerate electrons to very high speeds. Explain, using appropriate theory and relationships, how the accelerator achieves these high speeds.

5 A single positively charged particle with a charge of $+1.6 \times 10^{-19}$ C is travelling into a computer screen and perpendicular to a magnetic field that runs horizontally from left to right across the screen. In what direction will the force experienced by the charge act?

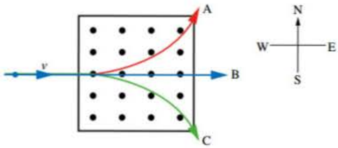
A left to right

B right to left

C vertically up

D vertically down

6 The following diagram shows a particle, with initial velocity v, about to enter a uniform magnetic field, B, directed out of the page.



a If the charge on this particle is positive, what is the direction of the force on this particle just as it enters the field?

b Which path will this particle follow?

c Does the kinetic energy of the particle increase, decrease or remain constant?

d If this particle were negatively charged, what path would it follow?

e What kind of particle could follow path B?

7 A single positively charged particle with a charge of $+1.6 \times 10^{-19}$ C travels at a velocity of $0.5\,\text{m}\,\text{s}^{-1}$ from right to left perpendicular to a magnetic field, B, of strength 2.0×10^{-5} T, running vertically downwards. Choose the option that gives the force that the particle will experience from the magnetic field.

A 1.6×10^{-5} N

B 3.2×10^{-5} N

C 1.6×10^{-19} N

D 1.6×10^{-24} N

8 A single negatively charged particle with a charge of -1.6×10^{-19} travels at a velocity of $1.0\,\text{m}\,\text{s}^{-1}$ from right to left parallel to a magnetic field, B, of strength 3.0×10^{-5} T.

What is the magnitude of the force the particle will experience from the magnetic field?

Chapter review

05

KEY TERMS

cathode ray tube
dipole
electric field
electric field strength
electrical potential
electron gun
field lines
Lorentz force
magnetic field
mnemonic
monopole
particle accelerator
potential difference
right-hand rule
synchrotron

REVIEW QUESTIONS

1 Calculate the force applied to an oil drop carrying a charge of +3.00 mC in a uniform electric field of 7.50 NC^{-1}.

2 Which one or more of the following statements are correct when drawing electric field lines around a charged object?

A Electric field lines go from positively charged objects to positively charged objects.

B Electric field lines go from negatively charged objects to positively charged objects.

C Electric field lines start and end at 90° to the surface, with no gap between the lines and the surface.

D Field lines can cross.

3 Explain the difference between electrical potential and potential difference.

4 Calculate the potential difference that exists between two points separated by 25.0 mm, parallel to the field lines, in an electric field of strength 1000 Vm^{-1}.

5 Between two plates forming a uniform electric field, where will the electrical field strength be at a maximum?

A close to the positive plate

B close to the earthed plate

C at all points between the plates

D at the mid-point between the plates

6 Choose the correct terms from the ones in bold to complete the relationship between work done and potential difference.
When a positively charged particle moves across a potential difference from a positive plate towards an earthed plate, work is done by the **field/charged particle** on the **field/charged particle**.

7 Calculate the work done to move a positively charged particle of 2.5×10^{-18} C a distance of 3.0 mm towards a positive plate in a uniform electric field of 556 NC^{-1}.

8 Calculate the force exerted on an electron ($q_e = 1.6 \times 10^{-19}$ C) travelling at a speed of 7.0×10^6 ms^{-1} at right angles to a uniform magnetic field of strength 8.6×10^{-3} T.

9 A particular electron gun accelerates an electron across a potential difference of 15 kV, a distance of 12 cm between a pair of charged plates. What is the magnitude of the force acting on the electron? (Use $q_e = 1.6 \times 10^{-19}$ C.)

10 For each of the following charged objects in a uniform electric field, determine if work was done on the field or by the field, or if no work is done.

a An electron moves towards a positive plate.

b A positively charged point remains stationary.

c A proton moves towards a positive plate.

d A lithium ion (Li^+) moves parallel to the plates.

e An alpha particle moves away from a negative plate.

f A positron moves away from a positive plate.

11 An alpha particle is located in a parallel plate arrangement that has a uniform electric field of 34.0 Vm^{-1}.

a Calculate the work done to move the alpha particle a distance of 1.00 cm from the earthed plate to the plate with a positive potential. ($q_\alpha = +3.204 \times 10^{-19}$ C)

b For the situation in part (a) decide whether work was done on the field, by the field or if no work was done.

12 A gold(III) ion is accelerated by the electric field created between two parallel plates separated by 0.020 m. The ion carries a charge of $+3q_e$ and has a mass of 3.27×10^{-25} kg. A potential difference of 1000 V is applied across the plates. The work done to move the ion from one plate to the other results in an increase in the kinetic energy of the gold(III) ion. If the ion starts from rest, calculate its final speed. (Use $q_e = -1.602 \times 10^{-19}$ C.)

13 A charged plastic ball of mass 5.00 g is placed in a uniform electric field pointing vertically upwards with a strength of 300.0 NC^{-1}. Calculate the magnitude and sign of the charge required on the ball in order to create a force upwards that exactly equals the weight force of the ball.

14 When electrons are discharged from an electron gun, their motion can be controlled by:

A additional electric fields only

B additional magnetic fields only

C additional electric and magnetic fields

D the motion of the electrons cannot be controlled

15 A single positively charged particle with a charge of $+1.6 \times 10^{-19}$ C travels at a velocity of $30\,\mathrm{m\,s^{-1}}$ at an angle of 30° to a magnetic field, B, of strength 6.0×10^{-5} T.

What is the magnitude of the force the particle will experience from the magnetic field?

16 A single positively charged particle with a charge of $+1.6 \times 10^{-19}$ C travels at a velocity of $60\,\mathrm{m\,s^{-1}}$ at an angle of 50° to a magnetic field, B. The force on the particle is 1.5×10^{-24} N.

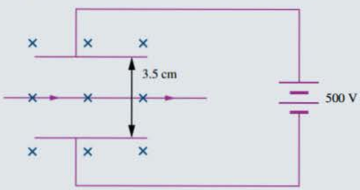
What is the magnitude of the magnetic field acting on the charged particle?

17 The diagram below represents an electron being fired at right angles towards a uniform magnetic field acting out of the page.

a Copy the diagram and mark on it the continued path you would expect the electron to follow.

b Which factors would alter the path radius of the electron as it travels?

18 A stream of electrons travels in a straight line through a uniform magnetic field and between a pair of charged parallel plates, as shown in the diagram.



Calculate the:

a magnitude of the electric field strength between the plates

b speed of the electrons, given that the magnetic field is of strength 1.5×10^{-3} T.

19 A single positively charged particle with a charge of $+1.6 \times 10^{-19}$ C travels at a velocity of $10\,\mathrm{m\,s^{-1}}$ perpendicular to a magnetic field, $\vec{B}$, of strength 3.0×10^{-5} T. What is the magnitude of the force the particle will experience from the magnetic field?

20 After completing the activity on page 146, reflect on the inquiry question: What happens to stationary and moving charged particles when they interact with an electric or magnetic field?

CHAPTER

06 The motor effect

In 1820, Hans Christian Ørsted discovered that an electric current could produce a magnetic field. His work established the initial ideas behind electromagnetism. Since then, our understanding and application of electromagnetism has developed to the extent that much of our modern way of living relies upon it. This includes the interaction and forces between current-carrying conductors, the theory of which is the basis for understanding electric motor operation.

In this chapter, you will investigate the interaction between a current-carrying conductor and a magnetic field, and the effect the magnetic field has on the conductor.

Content

INQUIRY QUESTION

Under what circumstances is a force produced on a current-carrying conductor in a magnetic field?

By the end of this chapter you will be able to:

- investigate qualitatively and quantitatively the interaction between a current-carrying conductor and a uniform magnetic field ($F = lI_{\perp}B = lIB \sin\theta$) to establish: (ACSPH080, ACSPH081) CCT ICT N
  - conditions under which the maximum force is produced
  - the relationship between the directions of the force, magnetic field strength and current
  - conditions under which no force is produced on the conductor
- conduct a quantitative investigation to demonstrate the interaction between two parallel current-carrying wires
- analyse the interaction between two parallel current-carrying wires ($\frac{F}{l} = \frac{\mu_0}{2\pi}\frac{I_1 I_2}{r}$) and determine the relationship between the International System of Units (SI) definition of an ampere and Newton's third law of motion (ACSPH081, ACSPH106). ICT L N

6.1 Force on a conductor

PHYSICS INQUIRY N CCT ICT

Motor effect

Under what circumstances is a force produced on a current-carrying conductor in a magnetic field?

COLLECT THIS...

- AA or AAA battery
- cylindrical rare-earth magnet
- bendable metal wire

DO THIS...

1 Stand the battery on the magnet with the positive end up.

2 Bend the wire to create a circuit. Some example shapes are shown below.

3 Balance the wire on the top of the battery, and ensure the bottom touches the magnet.

RECORD THIS...

Describe the motion of the wire.

Present a top-view diagram of the wire, indicating the current, the magnetic field and the resulting force.

REFLECT ON THIS...

Under what circumstances is a force produced on a current-carrying conductor in a magnetic field?

What variables affect the motion?

What could you do next time to improve your experiment?

An electric current is a flow of electric charges. These may be electrons in a metal wire, electrons and mercury ions in a fluorescent tube, or cations and anions in an electrolytic cell; the nature of the flowing charge that makes up the current does not matter. It is the total rate of flow of charge that is important. A magnetic field is produced around the flow of charge, and a force is experienced within this field. In this section you will investigate the forces that act between a magnetic field and a current-carrying conductor.

THE FORCE ON A CURRENT-CARRYING CONDUCTOR

Recall that a charged particle moving within a magnetic field will experience a force due to the magnetic field (Figure 6.1.1 on page 169). The magnitude of the force is proportional to the strength of the magnetic field, the component of the velocity of the charge that is perpendicular (at right angles) to the magnetic field and the charge on the particle.

When v and B are perpendicular:

$F = qv_{\perp}B$

where

F is the force (N)

q is the electric charge on the particle (C)

v is the component of the instantaneous velocity of the particle that is perpendicular to the magnetic field ($m\,s^{-1}$)

B is the strength of the magnetic field (T)

The force is at a maximum when the charged particle is moving at right angles to the field. There is no force acting when the charged particles are travelling parallel to the magnetic field.

FIGURE 6.1.1 A current-carrying conductor, such as a copper wire connected to a power supply, will feel a force when it is placed within a magnetic field.

Since a conducting wire is essentially a stream of charged particles flowing in one direction, it is not hard to imagine that a conductor carrying a stream of charges within a magnetic field will also experience a force. This is the theory behind the operation of electric motors.

The current in a conductor is dependent on the rate at which charges are moving through the conductor; that is:

$$I = \frac{q}{t}$$

where I is the current (A)

q is the total charge (C)

t is the time taken (s).

For a 1 m length of conductor, the velocity (v) of the charges through the conductor is:

$$v = \frac{s}{t} = \frac{1}{t}$$

And hence

$I = \frac{q}{t} = q \times \frac{1}{t} = qv$, for a 1 m length of conductor.

As $F = qv_{\perp}B$ for a single charge, q, moving perpendicular to a magnetic field, then:

$F = IB$ for a 1 m conductor (as $I = qv$),

so for a conductor of any length,

> **$F = lI_{\perp}B$**
>
> **where**
>
> **F is the force on the conductor perpendicular to the magnetic field (N)**
>
> **l is the length of the conductor (m)**
>
> **$I_{\perp}$ is the current in the conductor perpendicular to the magnetic field (A)**
>
> **B is the strength of the magnetic field (T)**

Just as for a single charge moving in a magnetic field, the force on the conductor is at a maximum when the conductor is at right angles to the field. The force is zero when the conductor is parallel to the magnetic field. In Chapter 5, the right-hand rule (Figure 6.1.2) was introduced. It is used to determine the direction of the force.

GO TO ➤ Section 5.2 page 158

The force experienced by a current-carrying conductor is a vector quantity. The expression noted above applies only to the component of the conductor perpendicular to the magnetic field. To find the force acting on any conductor, or part of a conductor, moving at an angle θ to the magnetic field, you can use the equation:

> **$F = lIB\sin\theta$**
>
> **where**
>
> **F is the force on the conductor (N)**
>
> **l is the length of the conductor (m)**
>
> **I is the current in the conductor (A)**
>
> **B is the strength of the magnetic field (T)**
>
> **θ is the angle between the magnetic field and the conductor**

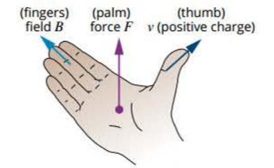


FIGURE 6.1.2 The right-hand rule is used to find the direction of the force. Point the thumb of the right hand in the direction of the movement of a positive charge (the direction of conventional current) and the fingers in the direction of the magnetic field. The force on the charge will point outwards from the palm.

> **At times, you may see the variable for length of a conductor written instead as L or $\vec{l}$.**

Figure 6.1.3 illustrates the relationship of the angle between the magnetic field and current-carrying conductor. This is particularly relevant when applied to practical electric motors. Electric motors are built to ensure that the current-carrying conductor is perpendicular to the magnetic field, so that the motor utilises the full effect of the magnetic force. Figure 6.1.3a illustrates the case where the current-carrying conductor is perpendicular to the magnetic field (θ is 90°), and the full magnitude of the magnetic force is experienced. In Figure 6.1.3b, the current-carrying conductor is parallel to a magnetic field, and the current-carrying conductor experiences no force (as the conductor is at an angle of 0° to the magnetic field).

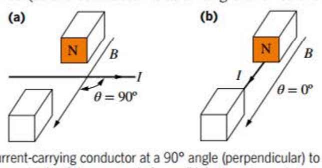
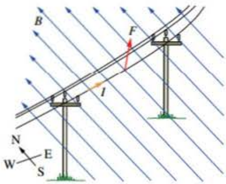


FIGURE 6.1.3 (a) A current-carrying conductor at a 90° angle (perpendicular) to a magnetic field, experiencing the full magnitude of the magnetic force. (b) A current-carrying conductor parallel to a magnetic field. In this situation, the current-carrying conductor experiences no force as the conductor is at an angle of 0° to the magnetic field.

Worked example 6.1.1

MAGNITUDE OF THE FORCE ON A CURRENT-CARRYING WIRE

Determine the magnitude of the force due to the Earth's magnetic field that acts on a suspended power line running east–west near the equator at the moment it carries a current of 100 A from west to east. Assume that the strength of the Earth's magnetic field at this point is 5.0×10^{-5} T.



Thinking	Working
Check the direction of the conductor and determine whether a force will apply. Forces only apply to the component of the wire perpendicular to the magnetic field.	As the current is running west–east and the Earth's magnetic field runs south–north, the current and the field are at right angles and a force will exist.
Establish what quantities are known and what quantities are required. Since the length of the power line hasn't been supplied, consider the force per unit length (i.e. 1 m).	$I = 100$ A $l = 1.0$ m $B = 5.0 \times 10^{-5}$ T $\theta = 90°$ $F = ?$
Substitute values into the force equation and simplify.	$F = IlB\sin\theta$ $= 1.0 \times 100 \times 5.0 \times 10^{-5} \times \sin 90$ $= 5.0 \times 10^{-3}$ N
Express the final answer in an appropriate form with a suitable number of significant figures. Note that only magnitude has been requested, so do not include direction.	$F = 5.0 \times 10^{-3}$ N per metre of power line

Worked example: Try yourself 6.1.1

MAGNITUDE OF THE FORCE ON A CURRENT-CARRYING WIRE

Determine the magnitude of the force due to the Earth's magnetic field that acts on a suspended power line running east–west near the equator at the moment it carries a current of 50 A from west to east. Assume that the strength of the Earth's magnetic field at this point is 5.0×10^{-5} T.

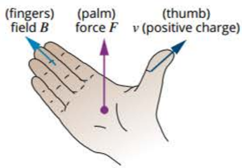
Worked example 6.1.2

FORCE AND DIRECTION ON A CURRENT-CARRYING WIRE

The Amundsen–Scott South Pole Station sits at a point that can be considered to be at the Earth's southern magnetic pole (which behaves like the north pole of a magnet).

Assuming the strength of the Earth's magnetic field at this point is 5.0×10^{-5} T, determine:

a the magnetic force on a 2.0 m length of wire carrying a conventional current of 10.0 A vertically up the exterior wall of one of the buildings	
Thinking	**Working**
Identify the known quantities. The direction of the magnetic field at the southern magnetic pole will be almost vertically upwards.	$I = 10.0$ A $l = 2.0$ m $B = 5.0 \times 10^{-5}$ T $\theta = 0°$ (The section of the wire running up the wall of the building will be parallel to the magnetic field, *B*.) $F = ?$
Substitute values into the appropriate equation and simplify.	$F = IlB \sin\theta$ $= 2.0 \times 10.0 \times 5.0 \times 10^{-5} \times \sin 0$ $= 0$ N Since there is no force, it is not necessary to state a direction.

<table>
<tr><td colspan="2">b the magnetic force on a 2.0 m length of wire carrying a conventional current of 10.0 A running horizontally right to left across the exterior of one of the buildings</td></tr>
<tr><td>Identify the known quantities.
The direction of the magnetic field at the southern magnetic pole will be almost vertically upwards.</td><td>The section of the wire running horizontally through the building will be perpendicular to the magnetic field, B. A force F with a strength equivalent to IlB will apply.</td></tr>
<tr><td>Identify the known quantities.</td><td>$I = 10.0$ A
$l = 2.0$ m
$B = 5.0 \times 10^{-5}$ T
$\theta = 90°$
$F = ?$</td></tr>
<tr><td>Substitute values into the appropriate equation and simplify.</td><td>$F = IlB \sin\theta$
$= 2.0 \times 10.0 \times 5.0 \times 10^{-5} \times \sin 90$
$= 1.00 \times 10^{-3}$ N</td></tr>
<tr><td>Determine the direction of the magnetic force using the right-hand rule.

</td><td>Align your hand so that your fingers are pointing in the direction of the magnetic field, i.e. vertically up.
Align your thumb so it is pointing left, in the direction of the conventional current, i.e. the movement of a positive charge.
Your palm should be facing inwards (towards the building). This is the direction of the force applied by the magnetic field on the wire.</td></tr>
</table>

State the magnetic force in an appropriate form with a suitable number of significant figures. Include the direction to fully describe the vector quantity.	$F = 1.0 \times 10^{-3}$ N inwards

c the magnitude of the magnetic force on a 5.0 m length of wire carrying a conventional current of 5.0 A running at a 45° angle from the side of the building.	
Identify the known quantities.	$I = 5.0$ A $l = 5.0$ m $B = 5.0 \times 10^{-5}$ T $\theta = 45°$ $F = ?$
Substitute values into the appropriate equation and simplify.	$F = IlB \sin\theta$ $= 5.0 \times 5.0 \times 5.0 \times 10^{-5} \times \sin 45$ $= 8.84 \times 10^{-4}$ N
State your answer using an appropriate number of significant figures. The magnitude of the force is a scalar quantity which doesn't require a direction.	$F = 8.8 \times 10^{-4}$ N

Worked example: Try yourself 6.1.2

WS 6.3 PA 6.1

FORCE AND DIRECTION ON A CURRENT-CARRYING WIRE

Santa's house sits at a point that can be considered the Earth's magnetic North Pole (which behaves like the south pole of a magnet).

Assuming the strength of the Earth's magnetic field at this point is 5.0×10^{-5} T, determine:

a the magnetic force on a 1.0 m length of wire carrying a conventional current of 1.0 A vertically up the outside wall of Santa's house

b the magnetic force on a 3.00 m length of wire carrying a conventional current of 15.0 A running horizontally right to left across the outside of Santa's house

c the magnitude of the magnetic force on a 1.5 m length of wire carrying a conventional current of 2.5 A running at a 30° angle from the side of Santa's house.

PHYSICS IN ACTION ICT

Electric motors are everywhere

The discovery of electric motors has revolutionised the world. Electric motors are found in many diverse applications, including household goods and appliances, computer cooling fans, power tools and bicycles (Figure 6.1.4), and even wrist watches. Some larger electric motors can be found in industrial applications and vehicles. The different types of electric motors can consume electric power ranging from tens of milliwatts to tens of megawatts.

Electric motors may be powered by direct current sources (e.g. fixed power supply or batteries), or alternating current sources (e.g. wall outlet mains supply). The principle of operation of electric motors is the same regardless of the size or the application: they utilise the interaction between a magnetic field and a current-carrying conductor to generate a force.

(a)

(b)

FIGURE 6.1.4 Electric motors are found in many diverse applications, such as (a) electric bicycles and (b) power tools.

PHYSICSFILE CCT

Early electric motors

In the early 1800s the principles of electromagnetic induction were discovered, in particular that an electric current produces a magnetic field. In the subsequent years much activity went into further experimentation and work, which finally led to a simple DC motor. While it is difficult to credit the discovery or invention of the first motor with one particular person, Michael Faraday's work was instrumental.

In Faraday's motor (Figure 6.1.5), a magnet was mounted vertically in a pool of mercury. A wire carrying a current hung from a support above, and the mercury provided a path for the current. The magnetic field of the magnet spread outwards from the top of the magnet and so there was a component of this field that was perpendicular to the wire. This produced a horizontal force on the wire that kept it rotating around the magnet.

Use the right-hand rule to convince yourself that if the current flows downwards and the magnetic field points out from the central magnet, the wire will rotate clockwise when viewed from above.

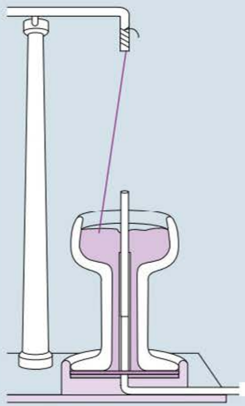


FIGURE 6.1.5 Michael Faraday's electric motor.

6.1 Review

SUMMARY

- The force acting on any conductor, or part of a conductor, moving at an angle θ to the magnetic field, is determined by the equation: $F = Il_{\perp}B = IlB\sin\theta$.
- The force is at a maximum when the charged particle is moving at right angles to the magnetic field.
- The force is zero when the charged particle is travelling parallel to the magnetic field.
- Electric motors are built to ensure that the current-carrying conductor is perpendicular to the magnetic field, to ensure that the motor utilises the full effect of the magnetic force.
- The right-hand rule is used to determine the direction of the force on a positive charge moving in a magnetic field, B. The direction of the force on a negatively charged particle is in the opposite direction.
- The direction of the force is given by the right-hand rule where the force travels out of the palm of the hand, once the thumb and fingers are orientated in the direction of the (conventional) current and magnetic field respectively.

KEY QUESTIONS

1 A single positively charged particle with a charge of $+1.6 \times 10^{-19}$ C is travelling into a computer screen and perpendicular to a magnetic field, B, that runs horizontally from left to right across the screen. In what direction will the force experienced by the charge act?

A left to right
B right to left
C vertically up
D vertically down

2 When is the force on a current-carrying conductor due to a magnetic field the strongest?

A when the conductor is parallel to the magnetic field
B when the conductor is at an angle of 30° to the magnetic field
C when the conductor is at an angle of 45° to the magnetic field
D when the conductor is at an angle of 90° (perpendicular) to the magnetic field

3 An east–west power line of length 100 m is suspended between two towers. Assume that the strength of the magnetic field of the Earth in this region is 5.0×10^{-5} T. Calculate the magnetic force (including direction) on this power line at the moment it carries a current of 80 A from west to east.

4 Determine the force acting on a 1.5 m length of conductor carrying 5.0 A of current that is at an angle of 30° to a magnetic field of strength 5.0×10^{-5} T. The conductor is running along the side of a house, with the current running from right to left. The magnetic field direction is vertically downwards.

5 The diagram below depicts a cross-sectional view of a long, straight current-carrying conductor, located between the poles of a permanent magnet. The magnetic field, B, of the magnet, and the current, I, are perpendicular. Calculate the magnitude and direction of the magnetic force on a 5.0 cm section of the conductor when the current is 2.0 A into the page and B equals 2.0×10^{-3} T.

6 An east–west power line of length 80 m is suspended between two towers. Assume that the strength of the magnetic field of the Earth in this region equals 4.5×10^{-5} T.

a Calculate the magnitude and direction of the magnetic force on this power line at the moment it carries a current of 50 A from east to west.

b Over time, the ground underneath the eastern tower subsides, so that the power line is lower at that tower. Assuming that all other factors are the same, is the magnitude of the magnetic force on the power line greater than before, less than before or the same as before?

6.2 Forces between conductors

In this section you will investigate the effect that two parallel current-carrying conductors have on each other. This interaction is used to derive the SI unit of the ampere.

FORCES BETWEEN TWO CURRENT-CARRYING WIRES

GO TO ➤ Year 11 Section 14.3

Recall from Year 11 that a conductor carrying a current creates a magnetic field surrounding it. According to **Ampere's law**, that magnetic field is defined by:

$$B = \frac{\mu_0 I}{2\pi r}$$

where

B is the magnetic field strength (T)

μ_0 is the magnetic permeability of free space (approximately $1.257 \times 10^{-6}\,\mathrm{N\,A^{-2}}$)

I is the current (A)

r is the radius from the current-carrying conductor to the location where the magnetic field is measured

Assume there are two parallel current-carrying wires, conductor A and conductor B. The magnetic field produced by wire A, seen at wire B, is given by

$$B_A = \frac{\mu_0 I_A}{2\pi r} \qquad \text{(equation 1)}$$

Recall from Section 6.1 that the magnetic force created by a current-carrying conductor is given by

$$F = lIB\sin\theta$$

The force experienced by wire B, due to the magnetic field from wire A, is then given by

$$F_B = lI_B B_A \sin\theta$$

Since the magnetic field and current-carrying conductors are perpendicular to each other, i.e. $\theta = 90°$, then the equation becomes

$$F_B = lI_B B_A \qquad \text{(equation 2)}$$

Substituting equation 1 into equation 2 gives

$$F_B = lI_B \frac{\mu_0 I_A}{2\pi r}$$

This equation can be rearranged to show the force per unit length on conductor B:

$$\frac{F_B}{l} = \frac{\mu_0 I_A I_B}{2\pi r}$$

Newton's third law applied to forces between two current-carrying conductors

Following the same process, it is possible to find a similar set of equations for the magnetic field produced by wire B, seen at wire A. The force per unit length experienced by wire A is then given by:

$$\frac{F_A}{l} = \frac{\mu_0 I_A I_B}{2\pi r}$$

Notice that both forces are equal. This is an example of Newton's third law: when one body exerts a force on another body (an action force), the second body exerts an equal force in the opposite direction on the first (the reaction force). In the case of two current-carrying conductors, the force on wire A from wire B must be equal and opposite to the force on wire B from wire A.

More generally, the equation for the force between two current-carrying conductors can be written as:

> $\frac{F}{l} = \frac{\mu_0 I_1 I_2}{2\pi r}$
>
> where
>
> $\frac{F}{l}$ is the force per unit length between the conductors ($N\,m^{-1}$)
>
> μ_0 is the magnetic permeability of free space (approximately $1.257 \times 10^{-6}\,N\,A^{-2}$)
>
> I_1 is the current conductor 1 (A)
>
> I_2 is the current conductor 2 (A)
>
> r is the distance between the two conductors (m)

> The permeability of free space, μ_0, can also be written more precisely as $4\pi \times 10^{-7}\,N\,A^{-2}$.

If the two parallel conductors carry current in the same direction, the forces attract (Figure 6.2.1a). If the two parallel conductors carry current in opposite directions, the forces will repel (Figure 6.2.1b). This can be proven using the right-hand rule.

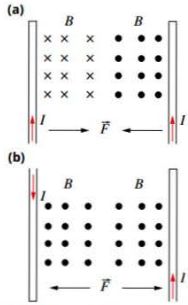


FIGURE 6.2.1 Using the right-hand rule it is possible to show that (a) when the current in the two conductors is in the same direction the force is attractive and (b) when the current is in different directions the force is repulsive.

Worked example 6.2.1

FORCES BETWEEN PARALLEL CURRENT-CARRYING CONDUCTORS

Determine the force per unit length acting between two current-carrying conductors, both carrying 10 A of current in the same direction, spaced 10 cm apart. Use $\mu_0 = 4\pi \times 10^{-7}\,N\,A^{-2}$.

Thinking	Working
Identify the known quantities.	$I_1 = 10\,A$ $I_2 = 10\,A$ $r = 0.10\,m$ $\mu_0 = 4\pi \times 10^{-7}\,N\,A^{-2}$ $F = ?$
Determine if the force is attractive or repulsive.	As both currents are flowing in the same direction, the force between the two conductors would be attractive.
Substitute values into the appropriate equation and simplify.	$\frac{F}{l} = \frac{\mu_0 I_1 I_2}{2\pi r}$ $= \frac{4\pi \times 10^{-7} \times 10 \times 10}{2\pi \times 0.10}$ $= 2.0 \times 10^{-4}\,N\,m^{-1}$ attractive

Worked example: Try yourself 6.2.1

FORCES BETWEEN PARALLEL CURRENT-CARRYING CONDUCTORS

Determine the force per unit length acting between two current-carrying conductors, both carrying 5.0 A of current in the same direction, spaced 20 cm apart. Use $\mu_0 = 4\pi \times 10^{-7}\,N\,A^{-2}$.

PHYSICSFILE ICT

Tiny magnetic fields

Whenever there are moving charges a magnetic field will be produced. While this is more obvious around charge-carrying conductors such as wires, many other objects are able to carry charge and so they can also create a (very tiny) magnetic field around them.

Tap water will have a certain amount of ions in it which will make it a conductor of electricity. A running tap will create a magnetic field of around 1×10^{-7} T at a distance of 20 cm.

Even your brain will create a magnetic field. As electric currents are produced by firing neurons, magnetic fields of around 1×10^{-13} T can be produced on the surface of the brain. Magnetoencephalography (MEG) is an imaging technique (Figure 6.2.2) which is able to map the small magnetic fields the brain produces.

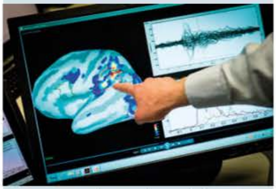


FIGURE 6.2.2 A magnetoencephalography (MEG) image of magnetic fields produced by brain activity.

PA 6.2 PA 6.2

Worked example 6.2.2

FORCES BETWEEN PARALLEL CURRENT-CARRYING CONDUCTORS WITH UNEQUAL CURRENTS

Determine the force per unit length acting between two current-carrying conductors, one carrying 10 A of current, the other carrying 5.0 A of current in the opposite direction, both spaced 10 cm apart. Use $\mu_0 = 4\pi \times 10^{-7}$ N A^{-2}.

Thinking	Working
Identify the known quantities.	$I_1 = 10$ A $I_2 = 5.0$ A $r = 0.10$ m $\mu_0 = 4\pi \times 10^{-7}$ N A^{-2} $F = ?$
Determine if the force is attractive or repulsive.	As the currents are flowing in different directions, the force between the two conductors would be repulsive.
Substitute values into the appropriate equation and simplify.	$\frac{F}{l} = \frac{\mu_0 I_1 I_2}{2\pi r}$ $= \frac{4\pi \times 10^{-7} \times 10 \times 5.0}{2\pi \times 0.10}$ $= 1.0 \times 10^{-4}$ N m^{-1} repulsive

Worked example: Try yourself 6.2.2

FORCES BETWEEN PARALLEL CURRENT-CARRYING CONDUCTORS WITH UNEQUAL CURRENTS

Determine the force per unit length acting between two current-carrying conductors, one carrying 20 A of current, the other carrying 15 A of current in the opposite direction, both spaced 5.0 cm apart. Use $\mu_0 = 4\pi \times 10^{-7}$ N A^{-2}.

DEFINING THE AMPERE

The equation to find the force per unit length between two conductors can be used to define the ampere. The base unit for measuring electric current in the **International System of Units** (SI) is the **ampere**. Its symbol is A, and it is often shortened to amps or amp.

Base measurements such as this are often defined in relation to other data. For instance, the SI definition of the metre is actually calculated as the length light would travel in 1 299 792 458 seconds.

The ampere is defined as being equal to the amount of current needed through two identical parallel conductors of infinite length when they are 1 m apart, in order to produce a force per unit length of 2×10^{-7} N m^{-1}.

This definition can be derived by first identifying the known quantities:

$$\frac{F}{l} = 2 \times 10^{-7}\ \text{N m}^{-1}$$
$$r = 1\ \text{m}$$

Substituting these values into the equation for the force between two conductors and solving for the current, I:

$$\frac{F}{l} = \frac{\mu_0 I_1 I_2}{2\pi r}$$
$$2 \times 10^{-7} = \frac{4\pi \times 10^{-7} \times I^2}{2\pi \times 1}$$
$$I^2 = 1$$
$$I = 1\ \text{A}$$

PHYSICSFILE ICT

Typical current values

The unit of measure for electric current is the ampere. The magnitude of currents found in many everyday devices and applications can vary significantly. It is therefore common to see current values ranging from nanoamperes (nA) through to tens of amperes.

Some examples of typical current values for devices are:

- Current in integrated circuit (microelectronic) components, including individual transistors: 100 nA to 500 nA
- Current to deliver charge to neural tissue in implantable medical devices (e.g. cochlear implant, cardiac pacemaker): 100 μA to 500 μA
- Current consumed by portable electronic chargers or power sources (e.g. laptop or iPad charger): 100 mA to 500 mA
- Current consumed by electronic devices such as TVs or set-top boxes: 500 mA to 2 A
- Current consumed by domestic appliances such as toasters: up to 10 A
- Current consumed by portable heaters: up to 10 A
- Current consumed by a car starter motor: up to 100 A

PHYSICS IN ACTION ICT S

The significance of current consumption

There are many devices and appliances in your house that consume electric power. The current consumption of these devices can vary significantly, which can have a significant impact on your household power consumption (and hence electricity bill).

The unit of measure for electric current is the ampere. Most devices will state their current consumption on a label on the device, or in the instructions. For example, mobile phone or tablet chargers usually consume several hundreds of milliamperes, which is shown on the charger. Some devices state their power consumption in killowatts (kW). For example, a portable heater may consume 2.5 kW. Knowing that the line voltage at your house is 240 V AC, and remembering that power consumption is $P = VI$, a 2.5 kW heater would consume up to 10.4 A of current if running at full capacity.

You can determine the effect some household appliances have on your overall electricity usage by monitoring the power consumption at the electricity meter (Figure 6.2.3). If you use a device that consumes tens of amperes of current, you will see the difference in your energy usage over a short time period (e.g. one day), whereas if you use devices that consume low amounts of current (e.g. milliamperes), it is unlikely that you will see any significant change in energy usage.

FIGURE 6.2.3 An example of an electricity meter, which measures the current consumed by all the circuits in your house. The meter will measure the power consumption per hour (in kilowatt hours), by measuring the current drawn and the supply voltage.

6.2 Review

SUMMARY

- The force per unit length between two parallel current-carrying conductors is given by the equation $\frac{F}{l} = \frac{\mu_0 I_1 I_2}{2\pi r}$.
- If the two parallel conductors carry current in the same direction, the force is attractive.
- If the two parallel conductors carry current in the opposite direction, the force is repulsive.
- The force experienced between the two parallel conductors is equal and opposite, according to Newton's third law.
- The forces between two parallel current-carrying conductors can be used to derive the SI definition of the ampere.

KEY QUESTIONS

1 If two parallel conductors are carrying current in the same direction, is the force between them attractive or repulsive?

2 Which of the following statements correctly describes the change in force between two parallel current-carrying conductors when one current increases by a factor of two, and the other current decreases by a factor of two?

A The force will increase by a factor of two.
B The force will decrease by a factor of two.
C The force will remain the same.
D The force will increase by a factor of four.

3 Calculate the force per unit length between two parallel current-carrying conductors spaced 2.0 cm apart. One conductor is carrying 10 A of current, while the other conductor is carrying 5.0 A of current.

4 Calculate the force per unit length between two parallel current-carrying conductors spaced 10 cm apart, with both conductors carrying 10 A of current flowing in the same direction.

5 Calculate the current being carried by two parallel conductors (both carrying the same current), spaced 20 cm apart, if the force per unit length between the two conductors is $1.0 \times 10^{-4}\,\text{N}\,\text{m}^{-1}$.

Chapter review

KEY TERMS

ampere
Ampere's law
International system of units

REVIEW QUESTIONS

1 Complete the following sentence by selecting the best option.
The magnitude of the magnetic force on a conductor aligned so that the current is running parallel to a magnetic field is:
  A dependent on the size of the current
  B dependent on the size of the magnetic field
  C dependent on the length of the conductor
  D zero
  E a maximum

2 Two parallel conductors are carrying current.
  a If the current in both conductors is in the same direction, is the force repulsive or attractive between them?
  b If the current in both conductors is in opposite directions, is the force repulsive or attractive between them?

3 A pair of straight parallel wires carrying currents of $2I$ and $3I$ are set up a distance d apart. They experience a force F between them. Calculate the force as a function of F if both currents are halved and the distance is doubled.

4 The right-hand rule is used to determine the force on a current-carrying conductor perpendicular to a magnetic field. Identify what part of the hand corresponds to the following physical quantities:
  a magnetic force
  b magnetic field
  c current in the conductor.

5 How much current, I, must be flowing in a wire 3.20 m long if the maximum force on it is 0.800 N and it is placed perpendicular to a uniform magnetic field of 0.0900 T?

6 Calculate the magnitude and direction of the magnetic force on conductors with the following sets of data:
  a $B = 1.0$ mT left, $l = 5.0$ mm, $I = 1.0$ mA up
  b $B = 0.10$ T left, $l = 1.0$ cm, $I = 2.0$ A up

7 Power lines carry an electric current in the Earth's magnetic field. Which would experience the greater magnetic force: a north–south power line or an east–west power line? Explain your answer.

8 If the magnitude of the current in a conductor parallel to a magnetic field doubles, what effect would this have on the force?
  A the force would halve
  B the force would remain the same
  C the force would double
  D the force would quadruple
  E none of the above

9 If you were building an electric motor, how would you position the current-carrying conductor in relation to the magnetic field?

10 An east–west power line of length 200 m is suspended between two towers. Assume that the strength of the magnetic field of the Earth in this region is 8.0×10^{-5} T. Calculate the magnetic force (including direction) on this power line at the moment it carries a current of 100 A from west to east.

11 Determine the magnitude of the force acting on a 3.0 m length of conductor carrying 10 A of current that is at an angle of 30° to a magnetic field of strength 5.0×10^{-5} T.

12 Determine the current in a 10 m conductor that results in a force of 10×10^{-5} N if placed in a magnetic field. The magnetic field strength is 5.0×10^{-5} T, and it runs perpendicular to the conductor.

13 The diagram depicts a cross-sectional view of a long, straight current-carrying conductor, located between the poles of a permanent magnet. The magnetic field, B, of the magnet, and the current, I, are perpendicular.

  a Calculate the magnetic force on a 10.0 cm section of the conductor when the current is 2.0 A into the page and B equals 2.0×10^{-3} T.
  b Calculate the magnetic force on a 10 m section of the conductor when the current is 12.0 A into the page and B equals 10.0×10^{-3} T.
  c Calculate the magnitude of the current if the magnetic force is 5×10^{-2} N acting on a 10 m section of the conductor, when B equals 5.0×10^{-3} T.

14 Calculate the magnitude of the magnetic field if the current is 10 A and the magnetic force is 15×10^{-2} N acting on a 20 m section of the conductor.

15 An east–west power line of length 200 m is suspended between two towers. Assume that the strength of the magnetic field of the Earth in this region equals 7.5×10^{-5} T.

Calculate the magnitude and direction of the magnetic force on this power line at the moment it carries a current of 40 A from east to west.

16 Which of the following statements correctly describes the *change* in force between two parallel current-carrying conductors, when one current decreases by a factor of four, and the other current decreases by a factor of two?

A The force will increase by a factor of eight.
B The force will decrease by a factor of eight.
C The force will remain the same.
D The force will increase by a factor of four.

17 Calculate the force per unit length between two parallel current-carrying conductors spaced 20 cm apart, with one conductor carrying 10 A of current upwards and the other conductor carrying 12 A of current downwards.

18 Calculate the force per unit length between two parallel current-carrying conductors spaced 50 cm apart, with both conductors carrying 15 A of current in the same direction.

19 Calculate the current being carried by two parallel conductors spaced 6 cm apart, if the force per unit length between the two conductors is 7.5×10^{-4} N m^{-1}.

20 Derive the fundamental definition of the ampere using the expression for the forces between two parallel current-carrying conductors.

21 After completing the activity on page 168, reflect on the inquiry question: Under what circumstances is a force produced on a current-carrying conductor in a magnetic field?

CHAPTER 07

Electromagnetic induction

In 1831, Englishman Michael Faraday and American Joseph Henry independently discovered that a changing magnetic flux could induce an electric current in a conductor. This discovery made possible the production of vast quantities of electricity. Today, whether the primary energy source is burning coal, wind, nuclear fission or falling water, most of the world's electrical energy production is the result of electromagnetic induction.

Content

INQUIRY QUESTION

How are electric and magnetic fields related?

By the end of this chapter you will be able to:

- describe how magnetic flux can change, with reference to the relationship $\Phi = B_{\parallel}A = BA\cos\theta$ (ACSPH083, ACSPH107, ACSPH109) ICT N
- analyse qualitatively and quantitatively, with reference to energy transfers and transformations, examples of Faraday's law and Lenz's law $\left(\varepsilon = -N\frac{\Delta\Phi}{\Delta t}\right)$, including but not limited to: (ACSPH081, ACSPH110) ICT N
  - the generation of an electromotive force (emf) and evidence for Lenz's law produced by the relative movement between a magnet, straight conductors, metal plates and solenoids
  - the generation of an emf produced by the relative movement or changes in current in one solenoid in the vicinity of another solenoid
- analyse quantitatively the operation of ideal transformers through the application of: (ACSPH110) ICT N
  - $\frac{V_P}{V_S} = \frac{N_P}{N_S}$
  - $V_P I_P = V_S I_S$
- evaluate qualitatively the limitations of the ideal transformer model and the strategies used to improve transformer efficiency, including but not limited to:
  - incomplete flux linkage
  - resistive heat production and eddy currents
- analyse applications of step-up and step-down transformers, including but not limited to:
  - the distribution of energy using high-voltage transmission lines. CCT

7.1 Magnetic flux

After Hans Christian Ørsted's discovery that an electric current produces a magnetic field, Michael Faraday, an English scientist, was convinced that the reverse should also be true—a magnetic field should be able to produce an electric current.

Faraday wound two coils of wire onto an iron ring (Figure 7.1.1). He connected a battery to one of the coils to create a strong current through it, which therefore created a strong magnetic field. He expected to then detect the creation of an electric current in the second coil. No matter how strong the magnetic field, he could not detect an electric current in the other coil.

One day he noticed that the galvanometer (a type of sensitive ammeter) attached to the second coil flickered when he turned on the current that created the magnetic field. It gave another flicker, in the opposite direction, when he turned the current off. It was not the strength of the magnetic field that mattered, but the change in the magnetic field.

The creation of an electric current in a conductor due to a change in the magnetic field acting on that conductor is now called **electromagnetic induction**.

FIGURE 7.1.1 Michael Faraday's original induction coil. Passing a current through one coil induces a voltage in the second coil by a process called mutual inductance. This coil is now on display at the Royal Institution in London.

PHYSICSFILE ICT

Models and theories

Michael Faraday was not alone in the discovery of electromagnetic induction. Joseph Henry (1797–1878), an American physicist, independently discovered the phenomenon of electromagnetic induction a little ahead of Michael Faraday, but Faraday was the first to publish his results. Henry later improved the design of the electromagnet, using a soft iron core wrapped in many turns of wire. He also designed the first reciprocating electric motor. Henry is credited with first discovering the phenomenon of self-induction, and the unit of inductance is named after him. He also introduced the electric relay, which made the sending of telegrams possible. Henry was the first director of the Smithsonian Institution.

While Faraday will be largely referred to throughout this text, it is worth noting that there can be a number of contributors who together built on the understanding of key ideas. Joseph Henry's contributions should not be forgotten.

CREATING AN ELECTRIC CURRENT

In his attempts to produce an electric current from a magnetic field, Faraday had no success with a constant magnetic field but was able to observe the creation of an electric current whenever there was a change in the magnetic field. This current is produced by an induced emf, ε. Although the term **emf** is derived from the name electromotive force, it is a voltage, or potential difference, rather than a force. Figure 7.1.2 indicates the induction of emf, and therefore current, caused by the perpendicular movement of a conducting wire relative to a magnetic field.

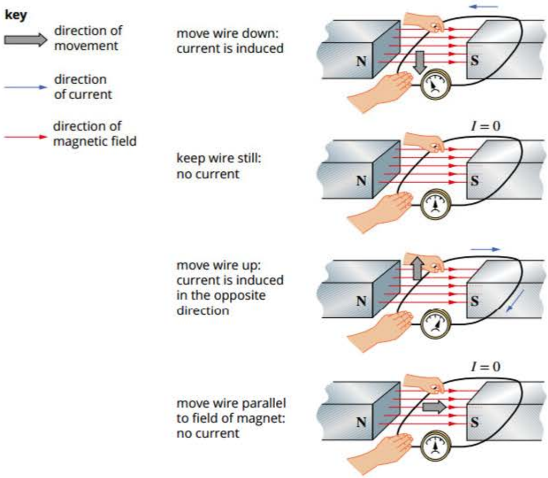


FIGURE 7.1.2 An electromotive force (emf) is induced in a wire when it moves perpendicular to a magnetic field.

MAGNETIC FLUX

To be able to develop ideas about the change in a magnetic field that induces an emf which can then create (or induce) a current, it is useful to be able to describe the 'amount of magnetic field'. This amount of magnetic field is referred to as the **magnetic flux**, a scalar quantity, denoted by the symbol Φ (the Greek letter phi). Faraday pictured a magnetic field as consisting of many lines of force. The density of the lines represents the strength of the magnetic field. Magnetic flux can be related to the total number of these lines that pass within a particular area. A strong magnetic field acting over a small area can produce the same amount of magnetic flux as a weaker field acting over a larger area, as shown in Figure 7.1.3. For this reason, magnetic field strength, B, is also referred to as **magnetic flux density**. B can be thought of as being proportional to the number of magnetic field lines per unit area perpendicular to the magnetic field.

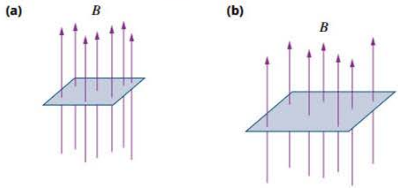


FIGURE 7.1.3 Magnetic flux: (a) a strong magnetic field acting over a small area (b) will have the same magnetic flux as a weaker magnetic field acting over a larger area.

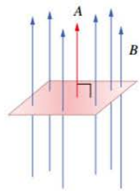


FIGURE 7.1.4 The area vector A (shown in red) is pointed in a direction normal to the plane of the area.

The area variable is actually represented by a vector A, the magnitude of which is equal to the area being examined. The direction of the area vector is normal to the plane of the area. This is shown in Figure 7.1.4.

The magnetic flux will be at a maximum when the area vector is parallel to the magnetic field, and zero when the area vector is perpendicular to the magnetic field.

Based on this, magnetic flux is defined as the product of the strength of the magnetic field, B, and the area of the field, i.e.

$\Phi = B_{\parallel} A$

where

Φ is the magnetic flux. The unit for magnetic flux is the weber, abbreviated to Wb, where $1 \text{ Wb} = 1 \text{ T m}^{-2}$

$B_{\parallel}$ is the strength of the magnetic field parallel to the area vector (T)

A is the area vector (m^2)

Since it is the plane of the area perpendicular to the magnetic field, the angle between the magnetic field and the area through which the field passes will affect the amount of magnetic flux. As the angle changes, the amount of magnetic flux will also change, until it reaches zero when the area under consideration is parallel to the magnetic field. Referring to Figure 7.1.5, then the relationship between the amount of magnetic flux and the angle θ to the field is:

$\Phi = BA \cos \theta$

It is important to note that θ is not the angle between the plane of the area and the magnetic field. Rather, it is the angle between a normal to the area and the direction of the magnetic field; hence the use of $\cos \theta$. When the area is at right angles to the magnetic field, the angle θ between the normal and the field is 0° and $\cos 0 = 1$ (top diagram in Figure 7.1.5). When the area is parallel to the magnetic field, the angle θ between the normal and the field is 90° and $\cos 90 = 0$ (bottom diagram in Figure 7.1.5).

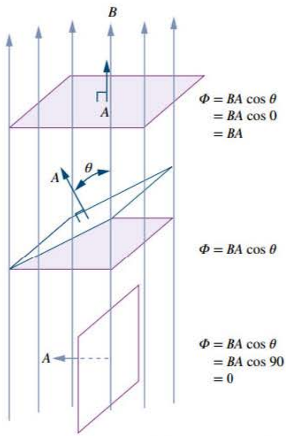


FIGURE 7.1.5 The magnetic flux is the strength of the magnetic field, B, multiplied by the area perpendicular to the magnetic field, given by $A \cos \theta$ and shown as the shaded areas in the above diagrams.

Worked example 7.1.1

MAGNETIC FLUX

A student places a horizontal square coil of wire of side length 5.0 cm into a uniform vertical magnetic field of 0.10 T. How much magnetic flux 'threads' the coil?

Thinking	Working
Calculate the area of the coil perpendicular to the magnetic field.	side length = 5.0 cm = 0.05 m area of the square = $(0.05\text{ m})^2$ $= 0.0025\text{ m}^2$
Calculate the magnetic flux.	$\Phi = B_{\parallel}A$ $= 0.1 \times 0.0025$ $= 0.00025\text{ Wb}$
State the answer in an appropriate form.	$\Phi = 2.5 \times 10^{-4}$ Wb or 0.25 mWb

Worked example: Try yourself 7.1.1

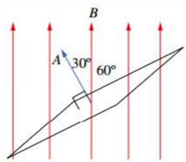
MAGNETIC FLUX

A student places a horizontal square coil of wire of side length 4.0 cm into a uniform vertical magnetic field of 0.050 T. How much magnetic flux 'threads' the coil?

Worked example 7.1.2

MAGNETIC FLUX AT AN ANGLE

A student places a square coil of wire of side length 10.0 cm into a uniform vertical magnetic field of 0.20 T. The plane of the square coil is at an angle of 60° to the magnetic field. How much magnetic flux 'threads' the coil?

Thinking	Working
Calculate the area of the coil.	side length = 10.0 cm = 0.1 m area of the square = $(0.01\text{ m})^2$ $= 0.01\text{ m}^2$
Draw a diagram to calculate the angle θ.	 The plane of the area is 60° to the magnetic field. So the area vector, which is directed normal to the plane, will be at an angle: $\theta = 90 - 60$ $= 30°$
Calculate the magnetic flux.	$\Phi = BA\cos\theta$ $= 0.2 \times 0.01 \times \cos 30$ $= 0.0017\text{ Wb}$
State the answer in an appropriate form.	$\Phi = 1.7 \times 10^{-3}$ Wb or 1.7 mWb

Worked example: Try yourself 7.1.2

MAGNETIC FLUX AT AN ANGLE

A student places a square coil of wire of side length 5.0 cm into a uniform vertical magnetic field of 0.10 T. The plane of the square coil is at an angle of 40° to the magnetic field. How much magnetic flux 'threads' the coil?

Note that in Worked Example 7.1.1 an area of 5 cm × 5 cm = 25 cm^2 was considered, and this corresponds to 0.0025 m^2 or 25×10^{-4} m^2; in other words:

$1\,cm^2 = 1 \times 10^{-4}\,m^2$

THE INDUCED EMF IN A MOVING CONDUCTOR

It was discovered that a change in the magnetic field, when a magnet is moved closer to a conductor, leads to an induced emf that in turn produces an **induced current**. While Faraday largely based his investigations on induced currents in coils, another way of inducing an emf is by moving a straight conductor in a magnetic field. It's not hard to understand why this is the case, when you know that charges moving in a magnetic field will experience a force.

Recall that when a charge, q, moves at a velocity v, perpendicular to a magnetic field B, the charge experiences a force F equal to $qv_{\perp}B$.

Considering the direction of movement shown in Figure 7.1.6, the force on the positive charges within the moving conductor would be along the conductor and out of the page. The force on the negative charges within the conductor would be along the conductor but into the page.

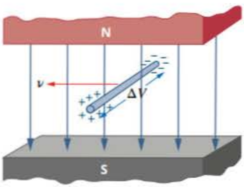


FIGURE 7.1.6 A potential difference, ΔV, will be produced across a straight wire moving to the left in a downward-pointing magnetic field.

As the charges in Figure 7.1.6 move apart due to the force they are experiencing from the magnetic field, one end of the conductor will become more positive, the other will become more negative, and a potential difference, ΔV, or emf will be induced between the ends of the conductor.

7.1 Review

SUMMARY

- An induced emf, ε, is produced by a changing magnetic flux in a process called electromagnetic induction.
- Magnetic flux is defined as the product of the strength of the magnetic field, B, and plane of the area perpendicular to the field lines, i.e. $\Phi = B_{\parallel}A$.
- The amount of magnetic flux varies with the angle of the field to the area under investigation; i.e. $\Phi = BA\cos\theta$. The angle θ is defined as between the magnetic field and the area vector which is directed normal to the plane of the area. The magnetic flux is then a maximum when the area vector is parallel (0°) and zero when the area vector is perpendicular (90°) to the field.
- The unit for magnetic flux is the weber, Wb; $1\,\text{Wb} = 1\,\text{T}\,\text{m}^{-2}$.

KEY QUESTIONS

1 Which of the following scenarios will *not* induce an emf in a long, straight conductor?

A A magnet is stationary alongside the conductor.

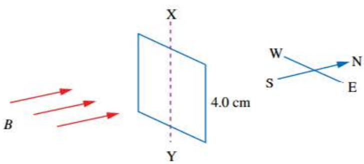
B A magnet is brought near the conductor.

C The conductor is brought into a magnetic field.

D The conductor is rotated within a magnetic field.

2 A student places a 4.0 cm square coil of wire parallel to a uniform vertical magnetic field of 0.050 T. How much magnetic flux 'threads' the coil?

3 A square loop of wire, of side 4.0 cm, is in a region of uniform magnetic field, $B = 2.0 \times 10^{-3}$ T north, as in the diagram below. The loop is free to rotate about a vertical axis XY. When the loop is in its initial position, its plane is perpendicular to the direction of the magnetic field. What is the magnetic flux passing through the loop?



4 The same square loop of wire as in Question 3 is initially perpendicular to the magnetic field. The loop is free to rotate about a vertical axis XY.

Describe what happens to the amount of magnetic flux passing through the loop as the loop is rotated through one complete revolution.

5 A circular coil of wire, of radius 5.0 cm, is perpendicular to a region of uniform magnetic field, $B = 1.6$ mT. What is the magnetic flux passing through the loop?

6 Calculate the magnetic flux through a horizontal square coil of wire of side length 5.0 cm perpendicular to a uniform vertical magnetic field of 0.10 T.

7 Calculate the magnetic flux through a circular coil of wire of radius 3.0 cm. The plane of the coil is at an angle of 50° to the magnetic field of strength 2.5 mT.

7.2 Faraday's and Lenz's laws

PHYSICS INQUIRY N CCT

Electromagnetic induction

How are electric and magnetic fields related?

COLLECT THIS...

- cylindrical rare-earth magnet, longer than the diameter of the tube so that it cannot spin inside the tube
- plastic tubing with an internal diameter large enough for the magnet to fall freely
- a spool of enamel copper wire
- two LEDs, 10mm, at least 5000 mcd, different colours
- soft landing material (pillow, fabric, Styrofoam)

DO THIS...

1. Wind the copper wire around the tubing, creating a 3 cm length of tube with hundreds of loops of wire.
2. Connect the LEDs to the two ends of the wire. Place the LEDs in different orientations. Ensure a good electrical connection by sanding the enamel off the wire. If possible solder the wire and LED legs together.
3. Place the tubing over the soft landing material. Drop the magnet through the tube.

RECORD THIS...

Describe how the energy is transferred and transformed during this activity.

Present a top-view diagram of the tubing, indicating the magnetic flux as the magnet enters the wire section and the direction of the induced current. Draw a similar diagram as the magnet is in the centre of the wire section and when it leaves the wire section.

REFLECT ON THIS...

How are electric and magnetic fields related?

What variables affect the current produced?

Faraday's early experiments largely centred on investigating electromagnetic induction in coils, or multiple loops, of wire. Faraday found that if a magnet is quickly moved into a coil, an emf is induced which causes a current to flow in the coil. If the magnet is removed, then a current flows in the coil in the opposite direction. Alternatively, if the magnet is held steady and the coil is moved in such a way that changes the magnetic flux, then once again an emf is induced and an electric current flows. It doesn't matter whether the coil or the magnet is moved—it is a *change* in flux that is required to induce the emf (Figure 7.2.1). This discovery led Faraday to his law of induction. Faraday's law of induction is the focus of this section.

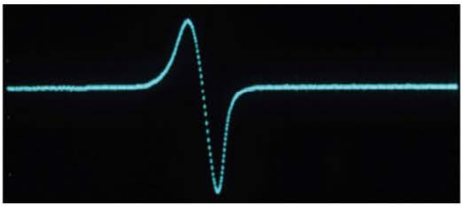


FIGURE 7.2.1 Oscilloscope trace from an electric coil, showing the voltage across the coil as a magnet is dropped through it.

FACTORS AFFECTING INDUCED EMF

Faraday quantitatively investigated the factors affecting the size of the emf induced in a coil. Firstly, emf will be induced by a change in the magnetic field. A simple example of this is to witness the emf induced when a magnet is brought towards or away from a wire coil. The greater the change, the greater the emf.

However, it is not only a change in the strength of a magnetic field, B, that induces an emf. It was noticed that an emf can be induced by changing the area perpendicular to the magnetic field through which the magnetic field lines pass, while keeping B constant. An example of this is the emf induced when a wire coil is rotated in the presence of a fixed magnetic field. This discovery indicates that the requirement for an induced emf is to have a changing magnetic flux, Φ.

Finally, Faraday discovered that the faster the change in magnetic flux, the greater the induced emf. This can be seen in the oscilloscope trace of a magnet falling through a coil as shown in the Figure 7.2.1 on page 189. The magnet is accelerated by gravity as it drops through the coil. Hence, the peak emf induced when the magnet first enters the coil at a relatively lower speed is noticeably less than the peak emf induced when the magnet leaves the coil at a faster speed. Thus, it is the *rate of change* of magnetic flux that determines the induced emf.

FARADAY'S LAW OF INDUCTION

Faraday's investigations led him to conclude that the average emf induced in a conducting loop, in which there is a changing magnetic flux, is proportional to the rate of change of flux.

This is now known as **Faraday's law** of induction and is one of the basic laws of electromagnetism.

Magnetic flux is defined as $\Phi = B_{\parallel}A$.

If the flux through N turns (or loops) of a coil changes from Φ_1 to Φ_2 during a time t, then the average induced emf during this time will be:

$$\varepsilon = -N\frac{(\Phi_2 - \Phi_1)}{t}$$

and if the change in magnetic flux $\Phi_2 - \Phi_1 = \Delta\Phi$, then

$$\varepsilon = -N\frac{\Delta\Phi}{\Delta t}$$

The negative sign is placed there as a reminder of the direction of the induced emf. This is discussed further on in the section. Normally you will be concerned only with the magnitude of the emf, which means you don't consider the negative sign or any negative quantities in a calculation.

If the ends of the coil are connected to an external circuit, then a current, I, will flow. The magnitude of the current is found using Ohm's law, which is:

$$I = \frac{V}{R}$$

where R is the resistance and V is the emf of the coil.

A coil not connected to a circuit will act like a battery not connected to a circuit. There will still be an induced emf but no current will flow.

$\varepsilon = -N\frac{\Delta\Phi}{\Delta t}$

where

ε is the induced emf (V)

N is the number of turns or loops

$\Delta\Phi$ is the change in magnetic flux (Wb)

Δt is the change in time (s)

Worked example 7.2.1

INDUCED EMF IN A COIL

A student winds a coil of area $40\,\text{cm}^2$ with 20 turns. He places it horizontally in a vertical uniform magnetic field of 0.10 T.

a Calculate the magnetic flux perpendicular to the coil.	
Thinking	**Working**
Identify the quantities to calculate the magnetic flux through the coil and convert to SI units where required.	$\Phi = B_{\parallel}A$ $B = 0.10\,\text{T}$ $A = 40\,\text{cm}^2 = 40 \times 10^{-4}\,\text{m}^2$
Calculate the magnetic flux and give your answer with appropriate units.	$\Phi = B_{\parallel}A$ $= 0.10 \times 40 \times 10^{-4}$ $= 4.0 \times 10^{-4}\,\text{Wb}$

b Calculate the magnitude of the average induced emf in the coil when the coil is removed from the magnetic field in a time of 0.5 s.	
Identify the quantities needed to determine the induced emf. Ignore the negative sign.	$\varepsilon = -N\frac{\Delta\Phi}{\Delta t}$ $N = 20$ turns $\Delta\Phi = \Phi_2 - \Phi_1$ $= 0 - 4.0 \times 10^{-4}$ $= 4.0 \times 10^{-4}$ Wb $\Delta t = 0.5$ s
Calculate the magnitude of the average induced emf, ignoring the negative sign that indicates the direction.	$\varepsilon = -N\frac{\Delta\Phi}{\Delta t}$ $= 20 \times \frac{4.0\times10^{-4}}{0.5}$ $= 0.016 = 16$ mV

Worked example: Try yourself 7.2.1

INDUCED EMF IN A COIL

A student winds a coil of area 50 cm^2 with 10 turns. She places it horizontally in a vertical uniform magnetic field of 0.10 T.

a Calculate the magnetic flux perpendicular to the coil.

b Calculate the magnitude of the average induced emf in the coil when the coil is removed from the magnetic field in a time of 1.0 s.

Worked example 7.2.2

NUMBER OF TURNS IN A COIL

A coil of cross-sectional area 1.0×10^{-3} m^2 experiences a change in the strength of a magnetic field from 0 to 0.20 T over 0.50 s. If the magnitude of the average induced emf is measured as 0.10 V, how many turns must be on the coil?

Thinking	Working
Identify the quantities needed to calculate the magnetic flux through the coil when in the presence of the magnetic field and convert to SI units where required.	$\Phi = B_{\parallel}A$ $B = 0.20$ T $A = 1.0 \times 10^{-3}$ m^2
Calculate the magnetic flux when in the presence of the magnetic field.	$\Phi = B_{\parallel}A$ $= 0.20 \times 1.0 \times 10^{-3}$ $= 2.0 \times 10^{-4}$ Wb
Identify the quantities needed to determine the induced emf. Ignore the negative sign.	$\varepsilon = -N\frac{\Delta\Phi}{\Delta t}$ $N = ?$ $\Delta\Phi = \Phi_2 - \Phi_1$ $= 2.0 \times 10^{-4} - 0$ $= 2.0 \times 10^{-4}$ Wb $\Delta t = 0.50$ s $\varepsilon = 0.10$ V
Rearrange Faraday's law and solve for the number of turns on the coil. Ignore the negative sign.	$\varepsilon = -N\frac{\Delta\Phi}{\Delta t}$ $N = -\frac{\varepsilon\Delta t}{\Delta\Phi}$ $= \frac{0.10\times0.50}{2.0\times10^{-4}}$ $= 250$ turns

PHYSICS IN ACTION ICT

Microphones

A microphone is a type of transducer, transforming energy from one form (an audio signal in the form of soundwaves) to another (electric energy/current). Many microphones operate by taking advantage of Faraday's law of induction. The so-called 'dynamic' microphone uses a tiny coil attached to a diaphragm. When soundwaves hit the diaphragm, the diaphragm moves in response to the sound. If the tiny coil is close to a permanent magnet, the movement of the coil in the magnetic field will induce an emf that varies with the original sound. That induced emf will cause a current to flow in the coil due to Faraday's law of induction.

FIGURE 7.2.2 A diver using a metal detector. If a metal object is found underneath the coil of the detector, an emf will be induced which creates a current that will affect the original current. The direction of the induced current is predicted by using Lenz's law.

Worked example: Try yourself 7.2.2

NUMBER OF TURNS IN A COIL

A coil of cross-sectional area $2.0 \times 10^{-3}\,m^2$ experiences a change in the strength of a magnetic field from 0 to 0.20 T over 1.00 s. If the magnitude of the average induced emf is measured as 0.40 V, how many turns must be on the coil?

LENZ'S LAW AND ITS APPLICATIONS

Lenz's law is a common way of understanding how electromagnetic induction obeys the principles of conservation of energy and explains the direction of the induced emf. It is named after Heinrich Lenz, whose research put a definite direction to the current created by the induced emf resulting from a changing magnetic flux.

Understanding the direction of the current resulting from an induced emf and how it is produced has allowed electromagnetic induction to be used in a vast array of devices that have transformed modern society, in particular in electrical generators. A metal detector is another example of a device that uses Lenz's law (Figure 7.2.2).

The direction of an induced emf

Lenz's law states that an induced emf always gives rise to a current whose magnetic field will oppose the original change in flux.

Figure 7.2.3 applies the law to the relative motion between a magnet and a single coil of wire. Moving the magnet towards or away from the coil will induce an emf in the coil, as there is a change in flux. The induced emf will produce a current in the coil, and this induced current will then produce its own magnetic field. It is worth noting that Lenz's law is a necessary consequence of the law of conservation of energy: if Lenz's law were not true then the new magnetic field created by a changing flux would encourage that change, which would have the effect of adding energy to the universe.

Applying Lenz's law, the magnetic field created by the induced current will oppose the change in flux caused by the movement of the magnet. When the north end of a magnet is brought towards the loop from the right, the magnetic flux from right to left through the loop increases. The induced emf produces a current that flows anticlockwise around the loop when viewed from the right. The magnetic field created by this current, shown by the little circles around the wire, is directed from left to right through the loop. It opposes the magnetic field of the approaching magnet.

If the magnet is moved away from the loop, as in part b of Figure 7.2.3, the magnetic flux from right to left through the loop decreases. The induced emf produces a clockwise current when viewed from the right. This creates a magnetic field that is directed from right to left through the loop. This field is in the same direction as the original magnetic field of the retreating magnet. However, note that it is opposing the change in the magnet's flux through the loop by attempting to replace the declining flux.

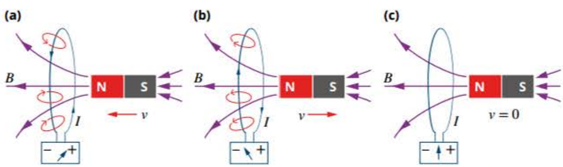


FIGURE 7.2.3 (a) The north end of a magnet is brought towards a coil from right to left, inducing a current that flows anticlockwise. (b) Pulling the north end of the magnet away from the coil from left to right induces a current in a clockwise direction. (c) Holding the magnet still creates no change in flux and hence no induced current.

When the magnet is held stationary, as in part c of Figure 7.2.3, there is no change in flux to oppose and so no current is induced.

The right-hand grip rule and induced current direction

The right-hand grip rule can be used to find the direction of the induced current. Keep in mind that the current must create a magnetic field that opposes the change in flux due to the relative motion of the magnet and conductor. Point your fingers through the loop in the direction of the field that is *opposing* the change and your thumb will then indicate the direction of the conventional current, as shown in Figure 7.2.4.

There are three distinct steps to determine the induced current direction according to Lenz's law:

1. What is the change that is happening?
2. What will *oppose* the change and/or restore the original conditions?
3. What must be the current direction to match this opposition?

These steps will be further examined in Worked example 7.2.3.

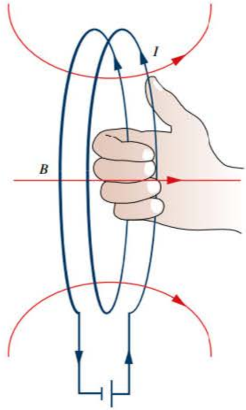
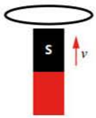


FIGURE 7.2.4 The right-hand grip rule can be used to determine the direction of a magnetic field from a current or vice versa. Your thumb points in the direction of the conventional current in the wire and your curled fingers indicate the direction of the magnetic field through the coil.

Worked example 7.2.3

INDUCED CURRENT IN A COIL FROM A PERMANENT MAGNET

The south pole of a magnet is brought upwards towards a horizontal coil initially held above it. In which direction will the induced current flow in the coil?

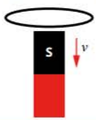


Thinking	Working
Consider the direction of the change in magnetic flux.	The magnetic field direction from the magnet will be downwards towards the south pole. The downward flux from the magnet will increase as the magnet is brought closer to the coil. So the change in flux is increasing downwards.
What will oppose the change in flux?	The induced magnetic field that opposes the change would act upwards.
Determine the direction of the induced current required to oppose the change.	In order to oppose the change, the current direction would be anticlockwise when viewed from above (using the right-hand grip rule).

Worked example: Try yourself 7.2.3

INDUCED CURRENT IN A COIL FROM A PERMANENT MAGNET

The south pole of a magnet is moved downwards away from a horizontal coil held above it. In which direction will the induced current flow in the coil?



Worked example 7.2.4

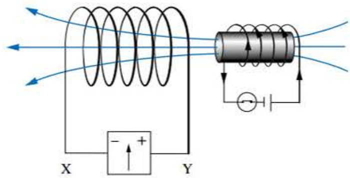
INDUCED CURRENT IN A COIL FROM AN ELECTROMAGNET

Instead of using a permanent magnet to change the flux in the loop as in Worked example 7.2.3, an electromagnet (on the right, in the diagram below) could be used. What is the direction of the current induced in the solenoid when the electromagnet is:

(i) switched on

(ii) left on

(iii) switched off?



Thinking	Working
Consider the direction of the change in magnetic flux for each case.	**(i)** Initially there is no magnetic flux through the solenoid. When the electromagnet is switched on, the electromagnet creates a magnetic field directed to the left. So the change in flux through the solenoid is increasing to the left. **(ii)** While the current in the electromagnet is steady, the magnetic flux through the solenoid is constant and the flux is not changing. **(iii)** In this case, initially there is a magnetic flux through the solenoid from the electromagnet directed to the left. When the electromagnet is switched off, there is no longer a magnetic flux through the solenoid. So the change in flux through the solenoid is decreasing to the left.
What will oppose the change in flux for each case?	**(i)** The magnetic field that opposes the change in flux through the solenoid is directed to the right. **(ii)** There is no change in flux and so there will be no opposition needed and no magnetic field created by the solenoid. **(iii)** The magnetic field that opposes the change in flux through the solenoid is directed to the left.
Determine the direction of the induced current required to oppose the change for each case.	**(i)** In order to oppose the change, the current will flow through the solenoid in the direction from X to Y (or through the meter from Y to X), using the right-hand grip rule. **(ii)** There will be no induced emf or current in the solenoid. **(iii)** In order to oppose the change, the current will flow through the solenoid in the direction from Y to X (or through the meter from X to Y), using the right-hand grip rule.

Worked example: Try yourself 7.2.4

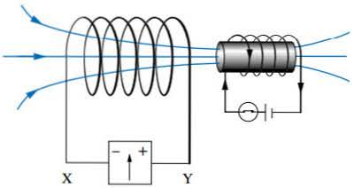
INDUCED CURRENT IN A COIL FROM AN ELECTROMAGNET

What is the direction of the current induced in the solenoid when the electromagnet is:

(i) switched on

(ii) left on

(iii) switched off?



PA 6.3

Induced current by changing area

It's very important to note that an induced emf is created while there is a change in flux, no matter how that change is created. As magnetic flux $\Phi = B_{\parallel}A$, a change can be created by any method that causes a relative change in the strength of the magnetic field, B, and/or the plane of the area perpendicular to the magnetic field. So an induced emf can be created in three ways:

- by changing the strength of the magnetic field
- by changing the area of the coil within the magnetic field
- by changing the orientation of the coil with respect to the direction of the magnetic field.

Figure 7.2.5 illustrates an example of the direction of an induced current that results during a decrease in the area of a coil.

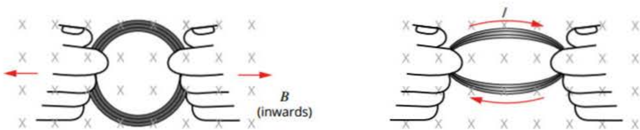


FIGURE 7.2.5 Inducing a current by changing the area of a coil. The amount of flux (the number of field lines) through the coil is reduced and an emf is therefore induced during the time that the change is taking place. The current flows in a direction that creates a field to oppose the reduction in flux into the page.

As the area of the coil decreases due to its changing shape, the flux through the coil (which is directed into the page) also decreases. Applying Lenz's law, the direction of the induced current would oppose this change and will be such that it acts to increase the magnetic flux through the coil into the page. Using the right-hand grip rule, a current would therefore flow in a clockwise direction while the area is changing.

In Figure 7.2.6, the coil is being rotated within the magnetic field. The effect is the same as reducing the area. The amount of flux flowing through the coil is reduced as the coil changes from being perpendicular to the field to being parallel to the field. An induced emf would be created while the coil is being rotated. This becomes particularly important when determining the current direction in a generator.

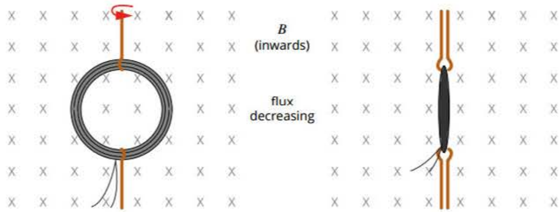
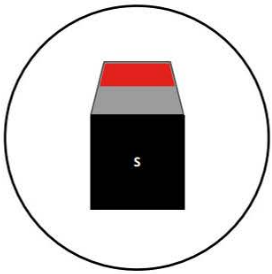


FIGURE 7.2.6 Changing the orientation of a coil within a magnetic field by rotating it reduces the amount of flux through the coil and so induces an emf in the coil while it is being rotated.

Worked example 7.2.5

FURTHER PRACTICE WITH LENZ'S LAW

The north pole of a magnet is moving towards a coil, into the page (the south pole is shown at the top looking down). In what direction will the induced current flow in the coil while the magnet is moving towards the coil?



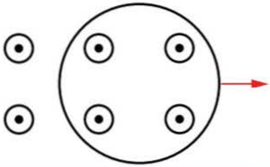
Thinking	Working
Consider the direction of the change in magnetic flux.	The magnetic field direction from the magnet will be away from the north pole, into the page. The flux from the magnet will increase as the magnet is brought closer to the coil. Therefore the change in flux is increasing into the page.
What will oppose the change in flux?	The magnetic field that opposes the change would act out of the page.
Determine the direction of the induced current required to oppose the change.	In order to oppose the change, the current direction would be anticlockwise when viewed from above (using the right-hand grip rule).

Worked example: Try yourself 7.2.5

PA
6.4

FURTHER PRACTICE WITH LENZ'S LAW

A coil is moved to the right and out of a magnetic field that is directed out of the page. In what direction will the induced current flow in the coil while the magnet is moving?



PHYSICSFILE ICT

Eddy currents

Lenz's law is important for many practical applications such as metal detectors, induction stoves and regenerative braking. These all rely on an eddy current, which is a circular electric current induced within a conductor by a changing magnetic field.

Applying Lenz's law, an eddy current will be in a direction that creates a magnetic field that opposes the change in magnetic flux that created it. Thus eddy currents can be used to apply a force that opposes the source of the motion of an external magnetic field. For example, if a metal plate is dragged out of a magnetic field, an eddy current will form within the plate that opposes the change in flux through the area of the plate, and thus opposes the motion of the plate itself due to the interaction of the magnetic fields (Figure 7.2.7).

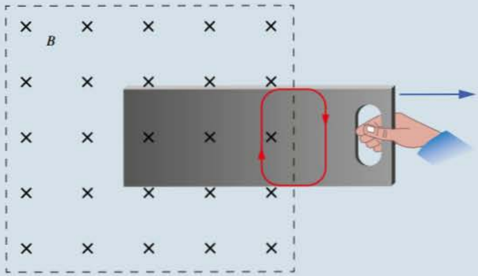


FIGURE 7.2.7 As the metal plate is moved towards the right, out of the magnetic field which is directed into the page, an eddy current forms in a clockwise direction. This eddy current would resist the motion of the plate.

This is the basis of regenerative braking, where the drag of the opposing magnetic field is utilised as a braking force. An eddy current flowing through a conductor with some resistance will also lose energy to the conductor by heating it. This makes eddy currents useful for an induction stovetop, but a potentially major source of energy loss within an AC generator, motor or transformer. Laminated cores with insulating material between the thin layers of iron are used in these applications to reduce the overall conductivity and suppress eddy currents.

The Earth's magnetic field is also a result of eddy currents. The energy that drives the Earth's dynamo comes from the enormous heat produced by radioactive decay deep in the Earth's core. The heat causes huge swirling convection currents of molten iron in the outer core. These convection currents of molten iron act rather like a spinning disk. They are moving in the Earth's magnetic field and so eddy currents are induced in them. It is these eddy currents that produce the Earth's magnetic field.

PHYSICSFILE ICT

Induction stoves

In contrast to a conventional gas or electric stove that heats via radiant heat from a hot source, an induction stove heats via the metal pot in which the food is being cooked. A coil of copper wire is placed within the cooktop (Figure 7.2.8). The AC electricity supply produces a changing magnetic field in the coil. This induces an eddy current in the conductive metal pot. The resistance of the metal in the pot, in which the eddy current flows, transforms electrical energy into heat and cooks the food.

While induction cooktops have only reached the domestic market in relatively recent times, the first patents for induction cookers were issued in the early 1900s. They have significant advantages over traditional electric cooktops in that they allow instant control of cooking power (similar to gas burners), they lose less energy through ambient heat loss and heating time, and they have a lower risk of causing burn injuries. Overall, the heating efficiency of an induction cooktop is around 12% better than traditional electric cooktops and twice that of gas.

FIGURE 7.2.8 The coil of an induction zone within an induction cooktop. The large copper coil creates an alternating magnetic field.

PHYSICS IN ACTION ICT

The Meissner effect

Superconductivity is a phenomenon that occurs when materials are cooled below a critical temperature (usually close to absolute zero), causing the material to have zero electrical resistance. Superconductors prevent magnetic fields from penetrating their interior, so that if a magnet is brought close to a superconductor it will levitate (Figure 7.2.9). This is known as the Meissner effect, named after the German physicists W. Meissner and R. Ochsenfeld in 1933 who discovered this property of superconductors. The Meissner effect is not the same as induced eddy currents. Eddy currents require a changing magnetic flux, hence the magnet would need to move. Yet in the Meissner effect, the magnet is stationary. Instead, this effect is due to quantum mechanical properties of the superconductor.

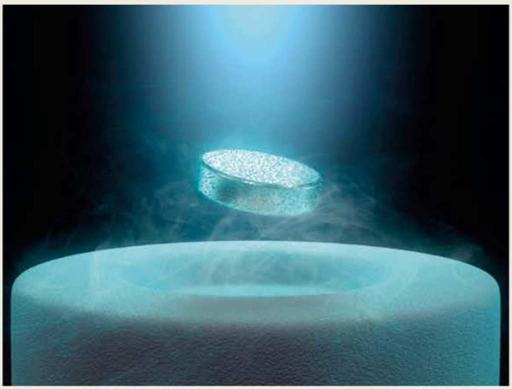


FIGURE 7.2.9 A magnet levitates above a superconductor due to the Meissner effect.

7.2 Review

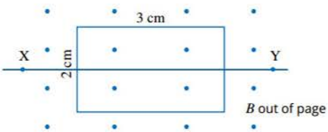
SUMMARY

- The emf induced in a conducting loop in which there is a changing magnetic flux is proportional to the negative rate of change of flux.
- This is described by Faraday's law of induction: $\varepsilon = -N\frac{\Delta\Phi}{\Delta t}$.
- The negative sign in Faraday's law indicates direction. For questions involving only magnitudes, you can ignore the negative sign in your calculations.
- Lenz's law states that an induced emf always gives rise to a current whose magnetic field will oppose the original change in flux.
- There are three distinct steps to determine the induced current direction according to Lenz's law:
  1. What is the change that is happening?
  2. What will oppose the change and/or restore the original conditions?
  3. What must be the current direction to match this opposition?
- An induced emf can be created in three ways:
  - by changing the strength of the magnetic field
  - by changing the area of the coil within the magnetic field
  - by changing the orientation of the coil with respect to the direction of the magnetic field.

KEY QUESTIONS

The following information relates to questions 1–3.

A single rectangular wire loop is located with its plane perpendicular to a uniform magnetic field of 2.0 mT, directed out of the page, as shown below. The loop is free to rotate about a horizontal axis XY.

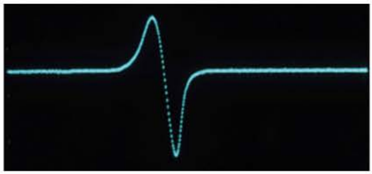


1 How much magnetic flux is threading the loop in this position?

2 The loop is rotated about the axis XY, through an angle of 90°, so that its plane becomes parallel to the magnetic field. How much flux is threading the loop in this new position?

3 If the loop completes one-quarter of a rotation in 40 ms, what is the average induced emf in the loop?

4 When a magnet is dropped through a coil, a voltage sensor will detect an induced voltage in the coil as shown below.



The area under the curve above zero is exactly equal to the area above the curve below zero because:

A The strength of the magnet is the same.

B The area of the coil is the same.

C The strength of the magnet and area of the coil are the same.

D The magnet speeds up as it falls through the coil.

The following information relates to questions 5 and 6.

A coil of 500 turns, each of area 10 cm^2, is wound around a square frame. The plane of the coil is initially parallel to a uniform magnetic field of 80 mT. The coil is then rotated through an angle of 90° so that its plane becomes perpendicular to the field. The rotation is completed in 20 ms.

5 What is the average emf induced in each turn during this time?

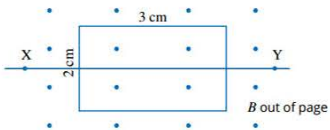
6 What is the effect on the average induced emf due to the multiple coils in Question 5?

7.2 Review *continued*

7 A conducting loop is located in an external magnetic field whose direction (but not necessarily magnitude) remains constant. A current is induced in the loop. Which of the following alternatives best describes the direction of the magnetic field due to the induced current?

- **A** It will always be in the same direction as the external magnetic field.
- **B** It will always be in the opposite direction to the external magnetic field.
- **C** It will be in the same direction as the external magnetic field if the external magnetic field gets weaker, and it will be in the opposite direction to the external magnetic field if the external magnetic field gets stronger.
- **D** The direction can't be determined from the information supplied.

8 A rectangular conducting loop forms the circuit shown below. The plane of the loop is perpendicular to an external magnetic field whose magnitude and direction can be varied. The initial direction of the field is out of the page.



a When the magnetic field is switched off, what will be the direction of the magnetic field due to the induced current?

- **A** out of the page
- **B** into the page
- **C** clockwise
- **D** anticlockwise
- **E** left to right
- **F** right to left

b When the direction of the external magnetic field is reversed, what is the direction of the magnetic field due to the induced current?

- **A** out of the page
- **B** into the page
- **C** clockwise
- **D** anticlockwise
- **E** left to right
- **F** right to left

7.3 Transformers

When Faraday first discovered electromagnetic induction, he had effectively invented the transformer. A **transformer** is a device for increasing and decreasing an alternating current (AC) voltage. Transformers can be found in many electrical devices: they are an essential part of any electrical distribution system and are the focus of this section (Figure 7.3.1).

FIGURE 7.3.1 (a) View of transformers at an electrical substation. The substation takes electricity from the distribution grid and converts it to lower voltages used by industrial or residential equipment. (b) More common are the smaller distribution transformers found on every suburban street. See if you can locate at least one on your street.

THE WORKINGS OF A TRANSFORMER

A transformer works on the principle of a changing magnetic flux inducing an emf. No matter what the size or application, a transformer will consist of two coils known as the primary and secondary coils. The changing flux originates with the alternating current supplied to the primary coil. The changing magnetic flux is directed to the secondary coil where the changing flux will induce an emf in that coil (Figure 7.3.2).

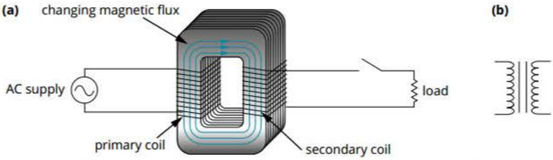


FIGURE 7.3.2 (a) In an ideal transformer, the iron core ensures that all the flux generated in the primary coil also passes through the secondary coil. (b) The symbol used in circuit diagrams for an iron-core transformer.

The two coils can be interwoven using insulated wire or they can be linked by a soft iron core, laminated to minimise eddy current losses. Transformers are designed so that nearly all of the magnetic flux produced by the primary coil will pass through the secondary coil. In an **ideal transformer** the assumption is that this will be 100% efficient and energy losses can be ignored. In a real transformer, this assumption remains a good approximation. Transformers are one of the most efficient devices around, with practical efficiencies often being better than 99%.

AC VERSUS DC

The power distribution system works on alternating current. That may seem odd when many devices run on direct current, but one of the primary reasons is the ease with which alternating current can be transformed from one voltage to another.

A transformer works on the basis of a changing current in the primary coil inducing a changing magnetic flux. This in turn induces a current in the secondary coil. For this to work, the original current must be constantly changing, as it does in an AC supply.

A DC voltage has a constant, unchanging current. With no change in the size of the current, no changing magnetic flux will be created by the primary coil and, hence, no current is induced in the secondary coil. Transformers do not work with the constant current of a DC electrical supply. There will be a very brief induced current when a DC supply is turned on, and a change occurs from zero current to the supply level. There is a similar spike if the DC supply is switched off, but while the DC supply is constant there is no change in magnetic flux to induce a current in the secondary coil.

THE TRANSFORMER EQUATION

When an AC voltage is connected to the primary coil of a transformer, the changing magnetic field will induce an AC voltage of the same frequency as the original supply in the secondary coil. The voltage in the secondary coil will be different and depends upon the number of turns in each coil.

+ ADDITIONAL

Laminations

Eddy currents that are created in the iron core of transformers can generate a considerable amount of heat. Energy that has been lost from the electrical circuit and the transformer as heat may become a fire hazard. To reduce eddy current losses, the transformer core is made of laminations, which are thin plates of iron electrically insulated from each other and placed so that the insulation between the laminations interrupts the eddy currents.

From Faraday's law, the average voltage in the primary coil, V_P, will affect the rate at which the magnetic flux changes:

$$V_P = N_P \frac{\Delta\Phi}{\Delta t}$$

or

$$\frac{\Delta\Phi}{\Delta t} = \frac{V_P}{N_P}$$

where N_p is the number of turns in the primary coil.

The induced voltage in the secondary coil, V_s, will be

$$V_S = N_S \frac{\Delta\Phi}{\Delta t}$$

and

$$\frac{\Delta\Phi}{\Delta t} = \frac{V_S}{N_S}$$

where N_s is the number of turns in the secondary coil.

Assuming that there is little or no loss of flux between the primary and secondary coils, then the flux in each will be the same and

$$\frac{V_P}{N_P} = \frac{V_S}{N_S}$$

or

$$\frac{V_S}{V_P} = \frac{N_S}{N_P}$$

> The transformer equation, relating voltage and number of turns in each coil, is:
> $\frac{V_P}{V_S} = \frac{N_P}{N_S}$ or $\frac{V_S}{V_P} = \frac{N_S}{N_P}$ or $\frac{V_P}{N_P} = \frac{V_S}{N_S}$

> The magnitude of alternating current (AC) voltage or current is expressed as the peak value, peak-to-peak value or RMS (root mean square) value. As AC current or voltage is a time-varying, sinusoidal value, the peak and peak-to-peak values refer to the height of the sinusoidal waveform (from zero to peak, or negative peak to positive peak). The RMS value is effectively the mean (average) value of the AC supply, and is often the value used in measurements.

The transformer equation explains how the secondary (output) voltage is related to the primary input voltage. Either the rms voltage for both or the peak voltage for both can be used.

A **step-up transformer** increases the secondary voltage compared with the primary voltage. The secondary voltage is greater than the primary voltage and the number of turns in the secondary coil is greater than the number of turns in the primary coil, i.e. if $N_S > N_P$ then $V_S > V_P$.

A **step-down transformer** decreases the secondary voltage compared with the primary voltage. The secondary voltage is less than the primary voltage and the number of turns in the secondary coil is less than the number of turns in the primary coil, i.e. if $N_S < N_P$ then $V_S < V_P$.

Worked example 7.3.1

TRANSFORMER EQUATION—VOLTAGE

A transformer is built into a portable radio to reduce the 240 V supply voltage to the required 12 V for the radio. If the number of turns in the secondary coil is 100, what is the number of turns required in the primary coil?

Thinking	Working
State the relevant quantities given in the question. Choose a form of the transformer equation with the unknown quantity in the top left position.	$V_S = 12\text{ V}$ $V_P = 240\text{ V}$ $N_S = 100$ turns $N_P = ?$ $\frac{N_P}{N_S} = \frac{V_P}{V_S}$
Substitute the quantities into the equation, rearrange and solve for N_P.	$\frac{N_P}{100} = \frac{240}{12}$ $N_P = \frac{100 \times 240}{12}$ $= 2000$ turns

Worked example: Try yourself 7.3.1

TRANSFORMER EQUATION—VOLTAGE

A transformer is built into a phone charger to reduce the 240 V supply voltage to the required 6 V for the charger. If the number of turns in the secondary coil is 100, what is the number of turns required in the primary coil?

POWER OUTPUT

Although a transformer very effectively increases or decreases an AC voltage, energy conservation means that the output power cannot be any greater than the input power. Since a well-designed transformer with a laminated core can be more than 99% efficient, the power input can be considered equal to the power output, making it an 'ideal' transformer.

Since power supplied is $P = VI$, then:

$$V_P I_P = V_S I_S$$

The transformer equation can then be written in terms of current, I.

The transformer equation, relating current and the number of turns in each coil:

$\frac{I_P}{I_S} = \frac{N_S}{N_P}$ or $\frac{I_S}{I_P} = \frac{N_P}{N_S}$ or $\frac{I_P}{N_S} = \frac{I_S}{N_P}$

Note carefully that the number-of-turns ratio for currents is the inverse of that for the transformer equation written in terms of voltage.

A transformer will be overloaded if too much current is drawn and the resistive power loss in the wires becomes too great. There will be a point at which the transformer starts to overheat rapidly. For this reason, it is important not to exceed the rated capacity of a transformer.

Worked example 7.3.2

TRANSFORMER EQUATION—CURRENT

A radio with 2000 turns in the primary coil and 100 turns in its secondary coil draws a current of 4.0 A. What is the current in the primary coil?

Thinking	Working
State the relevant quantities given in the question. Choose a form of the transformer equation with the unknown quantity in the top left position.	$I_S = 4.0\text{ A}$ $N_S = 100$ turns $N_P = 2000$ turns $I_P = ?$ $\frac{I_P}{I_S} = \frac{N_S}{N_P}$
Substitute the quantities into the equation, rearrange and solve for I_P.	$\frac{I_P}{4.0} = \frac{100}{2000}$ $I_P = \frac{4.0 \times 100}{2000}$ $= 0.20\text{ A}$

Worked example: Try yourself 7.3.2

TRANSFORMER EQUATION—CURRENT

A phone charger with 4000 turns in the primary coil and 100 turns in its secondary coil draws a current of 0.50 A. What is the current in the primary coil?

PHYSICSFILE ICT S

Standby power

Because very little current will flow in the primary coil of a good transformer to which there is no load connected, the transformer will use little power when not in use. However, this 'standby power' can add up to around 10% of power use. This is why devices such as TVs and computers should be switched completely off when not in use. Over the whole community, standby power amounts to megawatts of wasted power and unnecessary greenhouse emissions! Special switches, such as the 'Ecoswitch' shown below, have been developed that can be connected between the power outlet and the device to make it easier to remember to turn devices completely off when not in use.

FIGURE 7.3.3 Standby switches such as the 'Ecoswitch' make it easier and more convenient to turn devices completely off when not in use, saving up to 10% on power bills.

Worked example 7.3.3

TRANSFORMERS—POWER

The power drawn from the secondary coil of the transformer by a portable radio is 48 W. What power is drawn from the mains supply if the transformer is an ideal transformer?

Thinking	Working
The energy efficiency of a transformer can be assumed to be 100%. The power in the secondary coil will be the same as that in the primary coil.	The power drawn from the mains supply is the power in the primary coil, which will be the same as the power in the secondary coil: $P = 48\,W$.

Worked example: Try yourself 7.3.3

TRANSFORMERS—POWER

The power drawn from the secondary coil of the transformer by a phone charger is 3 W. What power is drawn from the mains supply if the transformer is an ideal transformer?

POWER FOR CITIES: LARGE-SCALE AC SUPPLY

In your school experiments using electrical circuits, it is likely that you have ignored the resistance of the connecting wires because the wires (generally made from copper) are good conductors, and so the resistance is very small over short distances. However, over large distances, even relatively good electrical conductors like copper have a significant resistance.

Modern cities use huge amounts of electrical energy, most of which is supplied from power stations built at a considerable distance from the metropolitan areas. The efficient transmission of the electrical energy with the least amount of power loss over that distance is therefore a very important consideration for electrical engineers, particularly given the vast distances between population centres in Australia.

The power lost in an electrical circuit is given by $\Delta P = \Delta VI$, where ΔV is the voltage drop across the load. Recalling Ohm's law, $\Delta V = IR$, and substituting it into the power equation, the power loss can be expressed in terms of either current and load resistance or voltage drop and load resistance:

$$P_{loss} = \Delta VI = I^2R = \frac{\Delta V^2}{R}$$

By considering the form of the equation including the current carried by the circuit and its electrical resistance ($P_{loss} = I^2R$), it is clear that transmitting large amounts of power using a large current will create very large power losses. If the current in the power lines can be reduced, it will significantly reduce the power loss. Since the power loss is proportional to the square of the current, then if the current is reduced by a factor of 3, for example, the power loss will be reduced by a factor of 3^2 or 9.

The challenge, then, is to transmit the large amounts of power being produced at power stations using a very low current. Transformers are the most common solution to this problem. Using a step-up transformer near the power station, the voltage is increased by a certain factor and, importantly, the current is decreased by the same factor. Due to the $P_{loss} = I^2R$ equation, the power lost during transmission is reduced by the square of that factor.

At this point you might be confused by the alternative equation for power loss: $P_{loss} = \frac{\Delta V^2}{R}$. A simple misunderstanding could make you think that increasing the voltage through the use of a step-up transformer would actually lead to greater power loss, if you use this equation to calculate power loss. However, ΔV represents the voltage drop in a circuit. You must be careful not to confuse the voltage being transmitted along the wires with the voltage drop across the wires. So, even though the voltage being transmitted is increased through the use of a step-up transformer, the voltage drop across the wires would be reduced since $\Delta V = IR$, and thus the power loss would also be reduced.

AC power from the generator is readily stepped up by a transformer to between 240 kV and 500 kV prior to transmission. Once the electrical lines reach the city, the voltage is stepped down in stages at electrical substations for distribution. The power lines in streets will have a voltage of around 2400 V, before being stepped down via small distribution transformers to 240 V for home use.

Worked example 7.3.4

TRANSMISSION-LINE POWER LOSS

300 MW is to be transmitted from the Murray 1 power station in the Snowy Mountains Scheme to Sydney, along a transmission line with a total resistance of 1.0 Ω. What would be the total transmission power loss if the initial voltage along the line was 250 kV?

Thinking	Working
Convert the values to SI units.	$P = 300\,\text{MW} = 300 \times 10^6\,\text{W}$ $V = 250\,\text{kV} = 250 \times 10^3\,\text{V}$
Determine the current in the line based on the required voltage.	$P = VI \therefore I = \frac{P}{V}$ $I = \frac{300 \times 10^6}{250 \times 10^3}$ $= 1200\,\text{A}$
Determine the corresponding power loss.	$P_{\text{loss}} = I^2R$ $= 1200^2 \times 1$ $= 1.4 \times 10^6\,\text{W}$ or $1.4\,\text{MW}$

Worked example: Try yourself 7.3.4

TRANSMISSION-LINE POWER LOSS

300 MW is to be transmitted from the Murray 1 power station in the Snowy Mountains Scheme to Sydney, along a transmission line with a total resistance of 1.0 Ω. What would be the total transmission power loss if the voltage along the line was 500 kV?

Worked example 7.3.5

VOLTAGE DROP ALONG A TRANSMISSION LINE

Power is to be transmitted along a transmission line with a total resistance of 1.0 Ω. The current is 1200 A. What voltage would be needed at the power generation end of the transmission line to achieve a supply voltage of 250 kV? Give your answer to four significant figures.

Thinking	Working
Determine the voltage drop along the transmission line.	$\Delta V = IR$ $= 1200 \times 1.0$ $= 1200\,\text{V}$
Determine the initial supply voltage.	$V_{\text{initial}} = V_{\text{supplied}} + \Delta V$ $= 250 \times 10^3 + 1200$ $= 251200\,\text{V}$ or $251.2\,\text{kV}$

Worked example: Try yourself 7.3.5

VOLTAGE DROP ALONG A TRANSMISSION LINE

WS 6.5 | PA 6.5

Power is to be transmitted along a transmission line with a total resistance of 1.0 Ω. The current is 600 A. What voltage would be needed at the power generation end of the transmission line to achieve a supply voltage of 500 kV? Give your answer to four significant figures.

PHYSICS IN ACTION ICT

The War of Currents

AC and DC power supplies have been in competition for nearly as long as humans have been generating electricity. The heated debates about the benefits and disadvantages of each type of current prompted what has been called the 'War of Currents' in the late 1800s. During this time Thomas Edison, an American inventor and businessman, had created the Edison Electric Light Company that he hoped would supply electricity to large parts of America with his DC generators. Meanwhile, Nikola Tesla, a Serbian-American physicist, had invented the AC induction motor and, with financial support from George Westinghouse, hoped AC would become the dominant power supply. Ultimately, the ease with which AC could be stepped up using transformers for long-distance transmission with minimal power loss (as discussed in detail throughout this chapter) proved to be the prevailing benefit that led to AC winning the 'war'. However, in his attempt to win the competition, Edison attempted to portray the high-voltage AC power as terrifyingly dangerous by using it to electrocute elephants and by inventing the AC-powered electric chair for the American government to execute prisoners on death row.

While AC power is now universal in large-scale power distributions, there is a limit to how high the voltage of an AC system can go and still be efficient. Above approximately 100 kV, corona loss (due to the high-voltage ionising air molecules) begins to occur, and above 500 kV it no longer becomes feasible to transmit electric power due to these effects.

LARGE-SCALE ELECTRICAL DISTRIBUTION SYSTEMS

Large-scale energy transmission is done through an interconnected grid between the power stations and the population centres where the bulk of the electrical energy is used. A wide-area synchronous grid, also known as an interconnection, directly connects a number of generators, delivering AC power with the same relative phase, to a large number of consumers.

No matter the source, the path the electrical power takes to the final consumer is very similar (Figure 7.3.4). Step-up transformers in a large substation near the power station will raise the voltage from that initially generated to 240 000 V (240 kV) or more. The electrical power will then be carried via high-voltage transmission lines to a number of substations near key centres of demand. Substations with step-down transformers then reduce the voltage to more safe levels for distribution underground or via the standard 'electricity pole' you would be familiar with around city and country areas. Each group of 10–15 houses will be supplied by a smaller distribution transformer, mounted on the poles, which reduces the voltage down to the 240 V AC rms voltage that home and business installations are designed to run on (see Figure 7.3.4).

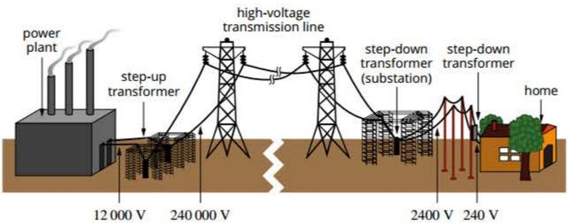


FIGURE 7.3.4 Transmitting electric power from a generator to the home uses AC power, so transformers can be used to minimise power losses through the system.

The use of AC as the standard for distribution allows highly efficient and relatively cheap transformers to convert the initial voltages created at the power station to much higher levels. The same power transmitted at a higher voltage requires less current and therefore less power loss. If it were not for this, the resistance of the transmission wires would need to be significantly reduced, which would require more copper in order to increase their cross-sectional area. This is both expensive and heavy. Less metal makes cables lighter and thinner, and the supporting towers themselves can be comparatively shorter, cheaper and lighter to build.

7.3 Review

SUMMARY

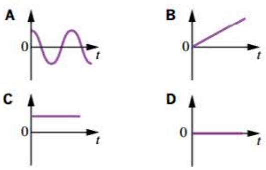
- A transformer works on the principle of a changing magnetic flux inducing an emf. No matter what the size or application, it will consist of two coils known as the primary and secondary coils.
- Ideal transformers are 100% efficient; real transformers are often over 99% efficient, and for this reason power losses within the transformer can be ignored in calculations.
- The transformer equation can be written in different versions but is based on:
  $\frac{V_P}{V_S} = \frac{N_P}{N_S}$
- A *step-up* transformer *increases* the secondary voltage compared with the primary voltage.
- A *step-down* transformer *decreases* the secondary voltage compared with the primary voltage.
- The transformer equation can also be written in terms of current, i.e.:
  $\frac{I_P}{I_S} = \frac{N_S}{N_P}$ or $V_P I_P = V_S I_S$
- Transformers will not work with DC voltage since it has a constant, unchanging current that creates no change in magnetic flux.
- The power supplied in an electrical circuit is given by: $P = VI$
- The power lost in an electrical circuit is given by: $P = I^2R$
- The AC electrical supply from a generator is readily stepped up or down by transformers, hence AC is the preferred form of electrical energy in large-scale transmission systems.

KEY QUESTIONS

1 A non-ideal transformer has a slightly smaller power output from the secondary coil than input to the primary coil. The voltage and current in the primary coil are denoted V_1 and I_1 respectively. The voltage and current in the secondary coil are denoted V_2 and I_2 respectively. Which of the following expressions describes the power output in the secondary coil?

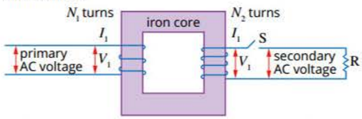
A $V_1 I_1$
B $V_2 I_2$
C $V_1 I_2$
D $I_2^2 R$

2 A voltage sensor is connected to the output of a transformer and a series of different inputs is used. Which of the following graphs is the most likely output displayed on a voltage graph for a steady DC voltage input?



3 A security light is operated from a mains voltage 240 V rms through a step-down transformer with 800 turns on the primary winding. The security light operates normally on an rms voltage of 12 V.
How many turns are on the secondary coil?

4 The figure below depicts an iron-core transformer. An alternating voltage applied to the primary coil produces a changing magnetic flux. The secondary circuit contains a switch, S, in series with a resistor, R. The number of turns in the primary coil is N_1 and in the secondary coil, N_2. The power in the first coil is P_1 and in the second coil, P_2. Assume that this is an ideal transformer.



a Write an equation that defines the relationship between the power in the primary coil, P_1, and the power in the secondary coil, P_2.

b Write an equation that defines the relationship between the current in the secondary coil, I_2, and the current in the primary coil, I_1, in terms of the number of turns in each coil.

5 A solar-powered generator produces 5.0 kW of electrical power at 500 V. This power is transmitted to a distant house via twin cables of total resistance 4.0 Ω. What is the total power loss in the cables?

7.3 Review *continued*

6 A 100 km transmission line made from aluminium cable has a total resistance of 10 Ω. The line carries the electrical power from a 500 MW power station to a substation. If the line is operating at 250 kV, what is the power loss in the line?

7 A power station generates 500 MW of power to be used by a town 100 km away. The power lines between the power station and the town have a total resistance of 2.0 Ω.

  a If the power is transmitted at 100 kV, what current would be required?

  b What voltage would be available at the town? Give your answer in kilovolts (kV).

8 Power loss can be expressed by the formula $P = \frac{\Delta V^2}{R} = I^2R$. Therefore, select which of the following statements is true, and justify why the other response is incorrect:

  A The greater the voltage being transmitted in a transmission line, the greater the power loss.

  B The greater the current in the transmission line, the greater the power loss.

Chapter review

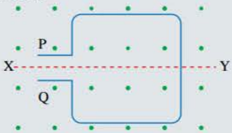
07

KEY TERMS

electromagnetic induction
emf
Faraday's law
ideal transformer
induced current
Lenz's law
magnetic flux
magnetic flux density
step-down transformer
step-up transformer
transformer

REVIEW QUESTIONS

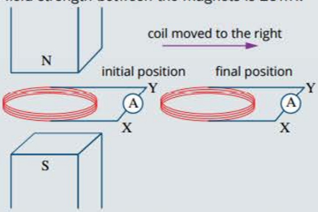
1 A rectangular coil of area $40\,cm^2$ and resistance $1.0\,\Omega$ is located in a uniform magnetic field $B = 8.0 \times 10^{-4}\,T$ which is directed out of the page. The plane of the coil is initially perpendicular to the field as depicted in the diagram below.



a What is the magnitude of the emf induced in the coil when the strength of the magnetic field is doubled in a time of 1.0 ms?

b What is the direction of the current caused by the induced emf in the coil when the strength of the magnetic field is doubled in a time of 1.0 ms?

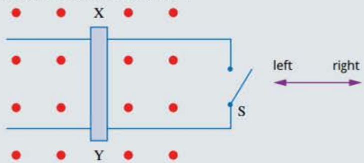
2 During a physics experiment a student pulls a horizontal circular coil from between the poles of two magnets in 0.10 s. The initial position of the coil is entirely in the field, while the final position is free of the field. The coil has 40 turns, each of radius 4.0 cm. The field strength between the magnets is 20 mT.



a What is the magnitude of the average emf induced in the coil as it is moved from its initial position to its final position?

b What is the direction of the current in the coil caused by the induced emf?

3 A copper rod, XY, of length 20 cm is free to move along a set of parallel conducting rails as shown in the following diagram. These rails are connected to a switch, S, which completes a circuit when it is closed. A uniform magnetic field of strength 10 mT, directed out of the page, is established perpendicular to the circuit. S is closed and the rod is moved to the right with a constant speed of $2.0\,m\,s^{-1}$.

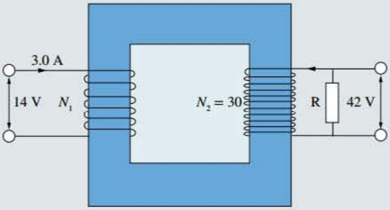


What is the direction of the current through the rod caused by the induced emf?

4 Coils S_1 and S_2 are close together and linked by a soft iron core. The emf in S_1 varies as shown in the graph below. Draw a line graph to show the shape of the variation of the current in S_2.

The following information relates to questions 5 and 6.

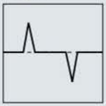
An ideal transformer is operating with an input voltage of 14 V and primary current of 3.0 A. The output voltage is 42 V. There are 30 turns in the secondary winding.



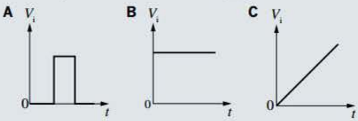
5 What is the output current?

6 How many turns are there in the primary coil?

7 The following diagram shows a graph of induced voltage versus time as it appears on the screen of a cathode ray oscilloscope.



Which of the following input voltages would produce the voltage shown in the CRO display?



The following information refers to questions 8 and 9.

A student builds a simple alternator consisting of a coil containing 500 turns, each of area 10 cm^2, mounted on an axis that can rotate between the poles of a permanent magnet of strength 80 mT. The alternator is rotated at a frequency of 50 Hz.

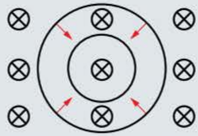
8 Find the average emf of the alternator.

9 Explain what the effect will be on the average emf when the frequency is doubled to 100 Hz.

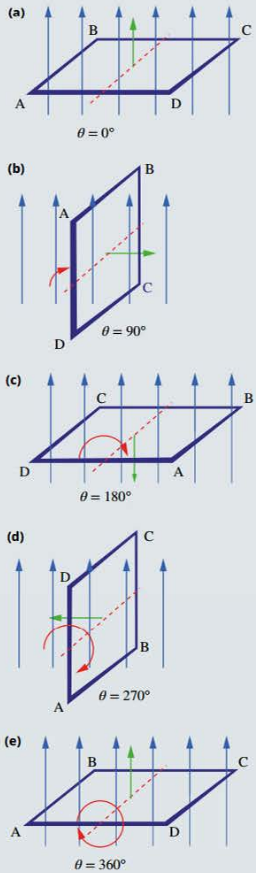
10 A generator is to be installed in a farm shed to provide 240 V power for the farmhouse. A twin-conductor power line with a total resistance of 8 Ω already exists between the shed and house. The farmer has seen a cheap 240 V DC generator advertised and is tempted to buy it.

Identify and explain two significant problems that you foresee with using the 240 V DC generator.

11 A coil in a magnetic field directed into the page is reduced in size. In what direction will the induced current flow in the coil while the coil is being reduced in size?



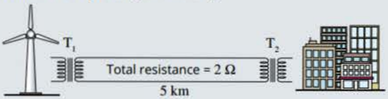
12 A single loop of wire is rotated within a magnetic field, *B*, as shown below.



While the coil is rotating, an emf will be generated as a result of which sides of the coil? Give a reason for your answer.

The following information relates to questions 13–16.

A wind turbine runs a 150 kW generator with an output voltage of 1000 V. The voltage is increased by a transformer T_1 to 10 000 V for transmission to a town 5 km away through power lines with a total resistance of 2 Ω. Another transformer, T_2, at the town reduces the voltage to 250 V. Assume that there is no power loss in the transformers (i.e. they are 'ideal').



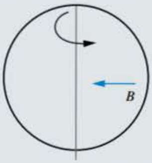
13 What is the current in the power lines?

14 What is the voltage at the input to the town transformer T_2?

15 What is the transmission power loss through the wires?

16 It is suggested that some money could be saved from the scheme by removing the first transformer. Explain, using appropriate calculations, whether this is a good plan.

17 A coil is rotated about its vertical axis such that the left-hand side would be coming out of the page and the right-hand side would be going into it. A magnetic field runs from right to left across the page. In what direction would the induced current in the coil flow?

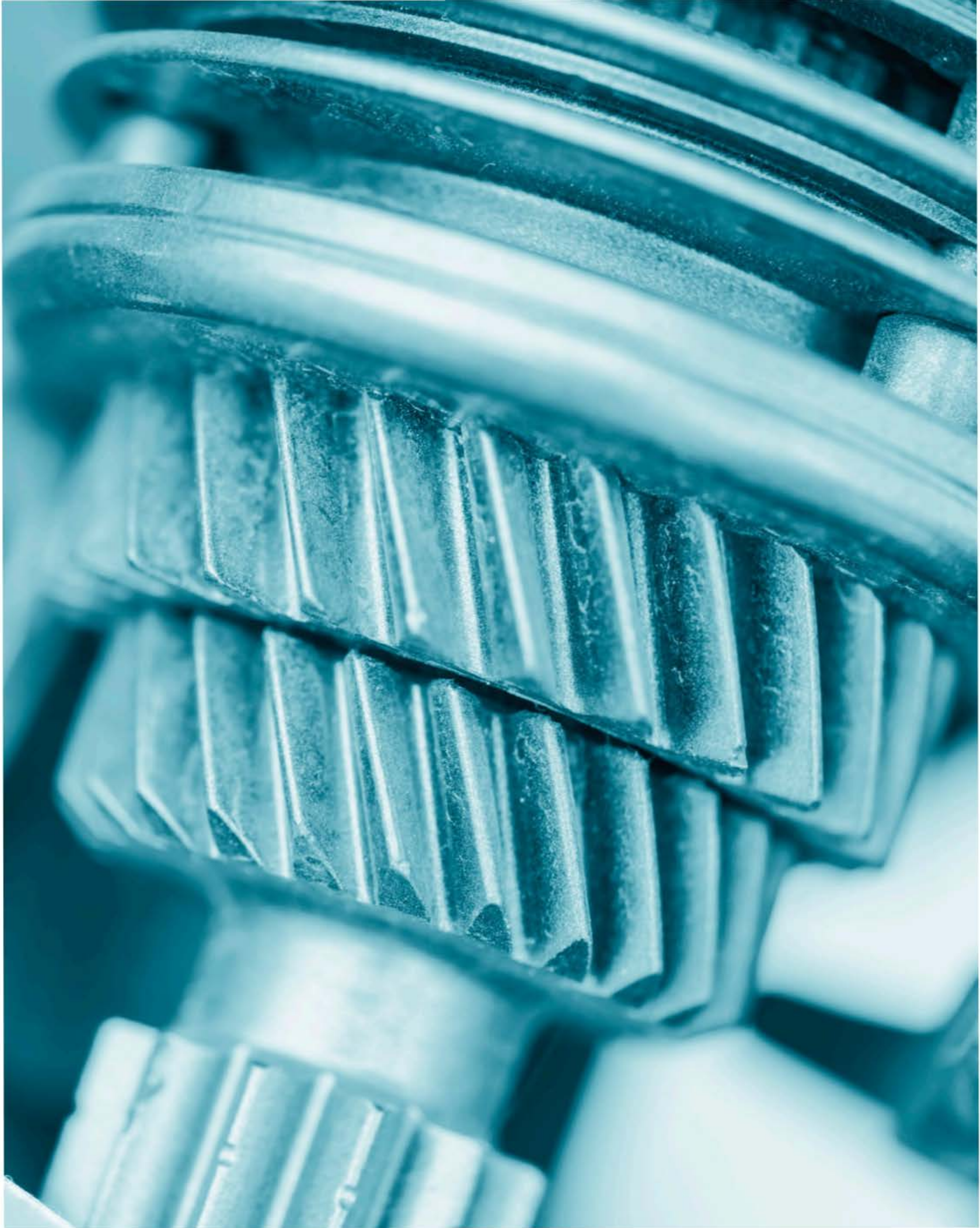


18 A student has a flexible wire coil of variable area of 100 turns and a strong bar magnet, which has been measured to produce a magnetic field of strength $B = 100\,\text{mT}$ a short distance from it. She has been instructed to demonstrate electromagnetic induction by using this equipment to light up an LED rated at 1.0 V. Explain, including appropriate calculations, one method with which she could complete this task.

19 A wire coil consisting of a single turn is placed perpendicular to a magnetic field that experiences a decrease in strength of 0.10 T in 0.050 s. If the emf induced in the coil is 0.020 V, what is the area of the coil?

20 A wire coil consisting of 100 turns with an area of 50 cm^2 is placed inside a vertical magnetic field of strength 0.40 T, and then rotated about a horizontal axis. For each quarter turn, the average emf induced in the coil is 1600 mV. Calculate the time taken for a quarter turn of the coil.

21 After completing the activity on page 189, reflect on the inquiry question: How are electric and magnetic fields related?



CHAPTER 08

Applications of the motor effect

Magnetic and electric fields affect various objects depending upon the nature and distance from the field. There does not need to be direct contact for fields to exert a force. This chapter looks at the application of these magnetic and electric fields. You will use your understanding of fields to explain how DC and AC motors and generators operate, and investigate the incredible impact both motors and generators have had on modern society. This includes the generation of electric energy used to run so many of our daily conveniences, and the motors you find in household appliances, heating and cooling systems, and vehicles. In fact, you probably need not look very far at home, school or in your neighbourhood to see an electric motor, or benefit from the generation of electric energy.

Content

INQUIRY QUESTION

How has the knowledge about the motor effect been applied to technological advances?

By the end of this chapter you will be able to:

- investigate the operation of a simple DC motor to analyse:
  - the functions of its components
  - production of a torque ($\tau = nIA_{\perp}B = nIAB\sin\theta$)
  - effects of back emf (ACSPH108) CCT ICT N
- analyse the operation of simple DC and AC generators and AC induction motors (ACSPH110) CCT
- relate Lenz's law to the law of conservation of energy and apply the law of conservation of energy to:
  - DC motors
  - magnetic braking. CCT

8.1 Motors

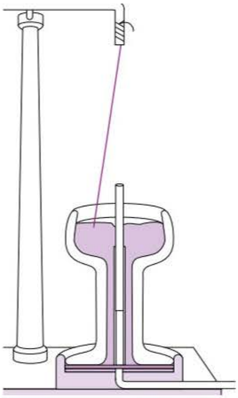


FIGURE 8.1.1 Michael Faraday's electric motor.

GO TO ➤ Section 6.1 page 170

DS 6.1

Physicists have always been interested in the relationship between electricity and magnetism because they wanted to understand the basic workings of the universe. For the world at large, however, this understanding provided a more practical form of excitement. It enabled the generation and use of electricity on a large scale. One of the most obvious applications of the understanding of electromagnetism gained in the 19th century is the electric motor.

DC MOTORS

The main components and the principles have been the same for all DC motors since Michael Faraday built the first one in 1821 (Figure 8.1.1). In Faraday's motor, a magnet was mounted vertically in a pool of mercury. A wire carrying a current hung from a support above (the mercury provided a path for the current). The magnetic field of the magnet spread outwards from the top of the magnet and so there was a component of this field that was perpendicular to the wire. This produced a horizontal force on the wire that kept it rotating around the magnet. Use the right-hand rule to convince yourself that if the current flows down and the magnetic field points out from the central magnet, the wire will rotate clockwise when viewed from above.

In modern **direct current** (DC) motors, a current-carrying coil of wire in a magnetic field experiences a force, F, when it is placed in a magnetic field. In practice, many turns of wire (n) are used and the magnetic field is provided by more than one permanent magnet or by an **electromagnet**.

The formula $F = nlIB$ includes the number of coils of wire, n, which equals 1 for all the examples in this section. Therefore, $F = lIB$ will be used to solve problems throughout this section.

Consider a single square coil of wire, with vertices ABCD, carrying a current, I, in a magnetic field, B, as shown in Figure 8.1.2.

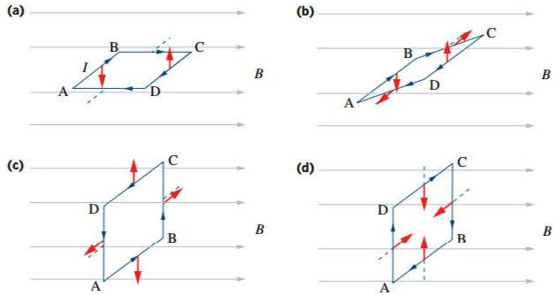


FIGURE 8.1.2 The magnetic force acting on each side of a current-carrying square wire coil in a magnetic field.

Initially the wire coil is aligned horizontally in a magnetic field, B, as in Figure 8.1.2a. Sides AD and BC are parallel to the magnetic field so no magnetic force will act on them. Sides AB and CD are perpendicular to the field so both of these sides will experience a magnetic force. Using the right-hand rule, there is a downward force on AB and an upward force on CD. These two forces will act together on the coil and cause it to rotate anticlockwise. If the coil is free to turn, it will move towards the position shown in Figure 8.1.2b.

In Figure 8.1.2b, there will be a magnetic force acting on every side of the coil. However, the forces acting on sides AD and BC will be equal and opposite in direction. They will tend to stretch the coil outwards but won't affect its rotation. The forces on sides AB and CD will remain and the coil will continue to rotate anticlockwise.

As the coil rotates to the position shown in Figure 8.1.2c, the forces acting on each side are such that they will tend to keep the coil in this position. The force on each side will act outwards from the coil. There are no turning forces at this point, but any further rotation will cause a force in the opposite direction that will cause the coil to rotate clockwise, back to this perpendicular position.

For the coil to continue to rotate anticlockwise at this point, the current direction needs to be reversed. This is shown in Figure 8.1.2d. With the current reversed, all of the forces are reversed, and provided the coil has a little momentum to get it past the perpendicular position, it will continue to rotate anticlockwise.

This ability to reverse the current direction at the point where the coil is perpendicular to the magnetic field is a key design feature in DC motors.

PHYSICSFILE ICT

Michael Faraday

Michael Faraday (1791–1867), depicted in Figure 8.1.3, was an English scientist who worked in the areas of chemistry and physics. He had little formal education. At the age of 14 he became the apprentice to a London bookbinder. During his apprenticeship he read many of the books that came his way. At the age of 21 he became a laboratory assistant to Sir Humphry Davy, who was one of the most prominent scientists of the day. Faraday was a gifted experimenter and after returning from a scientific tour through Europe with Davy, he began to be recognised in his own right for the scientific work he was doing. He was admitted to the Royal Society at age 32. He is credited with the discoveries of benzene, electromagnetic induction and the basis of the modern electric motor. He died in 1867 at Hampton Court. His contributions to science, and in particular his work in the area of electromagnetism, are recognised through the unit of measurement of capacitance known as the farad.

FIGURE 8.1.3 Michael Faraday

Torque

The turning force that the coil experiences in an electric motor is referred to as the **torque** on the coil. Recall from Chapter 3 that torque is the turning effect of any force; for example, when pushing on a swinging door. To achieve the maximum effect, the force should be applied at right angles to the door and at the largest distance possible from the point where the door is hinged. This idea is illustrated in Figure 8.1.4.

GO TO ➤ Section 3.4 page 94

Remember from Chapter 3 that the torque experienced by a system is given by the equation:

$$\tau = r_{\perp}F$$

where $r_{\perp}$ is the perpendicular force arm (in m) and F is the force (in N).

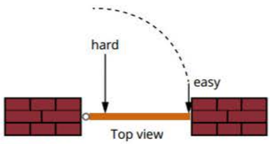


FIGURE 8.1.4 The force required to open a swinging door decreases as the perpendicular distance from the point of rotation increases and the torque, or turning effect, is maximised.

In the case of a coil in a magnetic field, the force experienced is always going to be at right angles to the magnetic field, as can be shown with the right-hand rule. This is shown in Figure 8.1.5 with a top view of a coil turning in a magnetic field. In this diagram, the current is shown to be moving up or down each side of the coil using the dot and cross method. A dot signifies the current is coming out of the page and a cross means the current is running into the page.

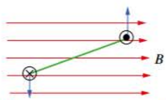


FIGURE 8.1.5 Top view of a coil turning in a magnetic field. The force experienced by each end of the coil is always perpendicular to the field.

A convention used to identify the current direction into or out of a page is to use a dot or a cross. This signifies the current flowing as if it were an arrow: the head of the arrow is the dot; the tail (feather) of the arrow is a cross.

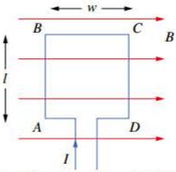


FIGURE 8.1.6 The plane of the coil ABCD is currently parallel to the field B.

GO TO ➤ Section 7.1 page 185

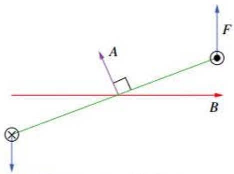


FIGURE 8.1.7 The area used in the torque equation actually represents a vector quantity which is directed normal to the plane of the coil.

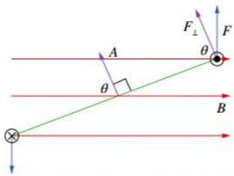


FIGURE 8.1.8 Using trigonometry, it is possible to find the torque experienced by the coil when the area vector A is at some angle to the magnetic field B.

In order to calculate the net torque on a coil, you need to find the torque applied to each side. In Figure 8.1.6, the plane of the coil is parallel to the field, so that the sides AD and BC experience zero force. Sides AB and CD will experience a force of $F_{AB} = lIB$ into the page and $F_{CD} = lIB$ out of the page. Use the right-hand rule to convince yourself of the directions of both of these forces.

The torque applied to the coil from both sides can then be calculated as:

$$\tau_{AB} = \tau_{CD} = r_{\perp}F = \left(\frac{w}{2}\right)lIB$$

So that the net torque becomes

$$\tau_{net} = \tau_{AB} + \tau_{CD} = wlIB$$

As you can see here, this equation multiplies the width of the coil by the length, which is equal to the area of the coil, so that the equation for the net torque becomes:

$$\tau = IAB$$

As in Section 7.1, the area variable in this equation represents a vector, which is directed normal to the plane of the coil (Figure 8.1.7).

The net force experienced by the coil is also dependent on the number of coils or conductors (n). By adding extra coils, the strength of the force increases proportionally (i.e. $F \propto n$), so that the torque must also be directly proportional to the number of coils.

$$\tau = nIAB$$

The maximum torque occurs when the area vector is perpendicular to the magnetic field. Using trigonometry, it is possible to find the torque of the coil at some angle θ. By setting θ to be the angle between the area vector and the magnetic field B (Figure 8.1.8), the force perpendicular to the arms is then given by $F \sin\theta$. This means that the torque equation for a motor can then be expressed by:

> $$\tau = nIA_{\perp}B = nIAB\sin\theta$$
>
> where
>
> τ is the torque (N m)
>
> n is the number of coils
>
> I is the current (A)
>
> $A_{\perp}$ is the area vector perpendicular to the plane of the coil (m^2)
>
> B is the strength of the magnetic field (T)
>
> θ is the angle between the area vector and the magnetic field (°)

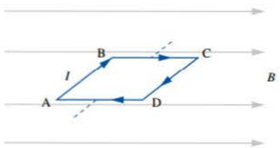
When the area vector is perpendicular to the magnetic field ($\theta = 90°$), the maximum torque is experienced. When the area vector is parallel to the magnetic field ($\theta = 0°$), there is no torque experienced.

This can be explained further with reference to Figure 8.1.2 on page 214. In Figure 8.1.2a, the plane of the current-carrying coil is parallel to the magnetic field, so that $\theta = 90°$. The forces on side AB and side CD create an anticlockwise torque. In Figure 8.1.2c, the forces on side AB and side CD are equal and opposite, therefore there is no net torque. The plane of the current-carrying coil is perpendicular to the magnetic field, so that $\theta = 0°$.

Worked example 8.1.1

TORQUE ON A COIL

A single square wire coil, ABCD, of side length 5.00 cm, is free to rotate within a magnetic field, B, of strength 1.00×10^{-4} T. A current of 1.00 A is flowing through the coil. What is the torque on the coil when the plane is parallel to the magnetic field? In what direction is the coil turning?

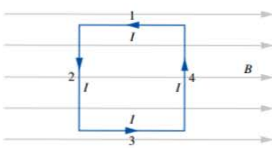


Thinking	Working
Confirm that the coil will experience a force based on the magnetic field and current directions supplied.	The right-hand rule confirms that a downward force applies on side AB. An upward force applies on side CD. The coil will turn anticlockwise. Sides AD and BC lie parallel to the magnetic field and no force will apply.
Identify the variables involved and state them in their standard form. Remember that the angle is between the magnetic field and the area vector which is directed at 90° to the plane of the coil.	$n = 1$ (as there is only one loop in the coil) $B = 1.00 \times 10^{-4}$ T $I = 1.00$ A $A = 0.05 \times 0.05 = 0.0025\,m^2$ $\theta = 90°$ (the plane of the coil is parallel to the field) $\tau = ?$
Calculate the torque on the coil.	$\tau = nIA_{\perp}B = nIAB\sin\theta$ $= 1 \times 1.00 \times 10^{-4} \times 1.00 \times 0.0025 \times \sin 90$ $= 2.50 \times 10^{-7}$ N m The direction is anticlockwise.

Worked example: Try yourself 8.1.1

TORQUE ON A COIL

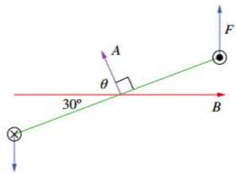
A single square wire coil, with a side length of 4.0 cm, is free to rotate within a magnetic field, B, of strength 1.2×10^{-4} T. A current of 1.5 A is flowing through the coil. What is the torque on the coil when it is parallel to the magnetic field? In what direction is the coil turning?



Worked example 8.1.2

TORQUE ON A COIL AT AN ANGLE

A single square wire coil of side length 10.0 cm is free to rotate within a magnetic field, *B*, of strength 1.00×10^{-4} T. A current of 5.00 A is flowing through the coil. What is the magnitude of the torque on the coil when it is at an angle of 30° to the magnetic field?

Thinking	Working
Determine the angle between the area vector of the coil and the magnetic field. Remember that the angle is between the magnetic field and the area vector which is directed at 90° to the plane of the coil.	 $\theta = 90 - 30$ $= 60°$
Identify the variables involved and state them in their standard form.	$n = 1$ (as there is only one loop in the coil) $B = 1.00 \times 10^{-4}$ T $I = 5.00$ A $A = 0.10 \times 0.10 = 0.01\ \text{m}^2$ $\theta = 60°$ $\tau = ?$
Calculate the magnitude of the torque on the coil.	$\tau = nIA_{\perp}B = nIAB\sin\theta$ $= 1 \times 1.00 \times 10^{-4} \times 5.00 \times 0.01 \times \sin 60$ $\tau = 4.33 \times 10^{-6}$ N m

Worked example: Try yourself 8.1.2

TORQUE ON A COIL AT AN ANGLE

A single square wire coil of side length 5.00 cm is free to rotate within a magnetic field, *B*, of strength 1.00×10^{-4} T. A current of 2.50 A is flowing through the coil. What is the magnitude of the torque on the coil when it is at an angle of 30.0° to the magnetic field?

PRACTICAL DC MOTORS

A basic single-coil electric motor with a simple arrangement to reverse the current direction will work, but it won't turn very smoothly. That's because maximum torque will only apply each half turn or twice for every full rotation. A number of enhancements have been developed over time to make DC motors the highly practical motive force they are today.

The **commutator** is usually made from a split ring of copper or another good conductor on which conducting **brushes** (usually carbon blocks) rub. Each half is connected to one end of the coil of wire. This arrangement of brushes prevents the wire from becoming tangled as the coil rotates. The commutator reverses the current at the point where the coil is perpendicular to the magnetic field, which keeps the coil rotating (Figure 8.1.9).

Practical motors will have many sets of coils of many turns each, spaced at an angle to each other, as shown in Figure 8.1.10.

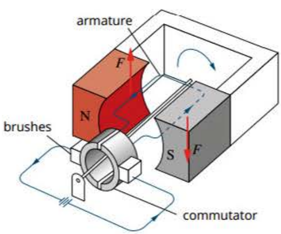


FIGURE 8.1.9 The main parts of a simple but practical single-coil DC electric motor.

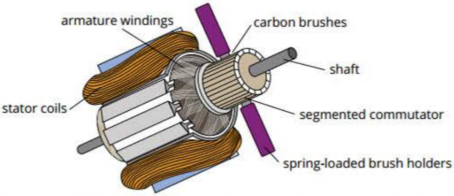


FIGURE 8.1.10 A typical multi-coil DC electric motor, showing the main components. Note that there are many sets of coils offset by an angle from each other. The stator coils produce an electromagnet that provides the magnetic field. The commutator feeds current to the armature coils in the position where maximum torque will be experienced.

The coils are wound around a soft iron core to increase the magnetic field that passes through them. The whole arrangement of core and coils is called an **armature** (as shown in Figure 8.1.10). Permanent magnets are generally used to provide the magnetic field in small motors, but in larger motors electromagnets are used as they can produce larger and stronger fields. These magnets are usually stationary, as distinct from the rotating rotor or armature, and are often referred to as the **stator**. The commutator is arranged to feed current to the particular coil that is in the best position to provide maximum torque. The total torque will be the sum of the torques on all the individual coils.

Generally speaking, the larger the torque in an electric motor, the better. This is achieved by the use of a strong, radial magnetic field; a large number of turns of wire in each coil; a high current; and a large area of coil. All this adds to the cost, so when designing an electric motor, each aspect may be compromised to some extent, depending on the motor's potential use.

PA 6.6

AC INDUCTION MOTORS

The majority of electric motors in use rely on AC current, not DC current. In a DC motor, the magnetic field is stationary (the stator), and the coil that carries the electric current rotates (the rotor). The DC current is connected to the coil on the rotor.

An AC induction motor works by producing a rotating magnetic field. The stator of an AC induction motor is made up of pairs of electromagnets (Figure 8.1.11). The AC current runs through the coils, creating a magnetic field. AC current reverses its polarity at a rate of 50 Hz, so the coils are energised in pairs, producing a magnetic field that rotates around the outside of the motor.

The rotor of the AC induction motor comprises a series of conductors (metal bars). As the magnetic field from the stator is rotating (changing magnetic field), an electric current is induced in the conductors of the rotor. This is due to Faraday's law. The induced electric current produces its own magnetic field. According to Lenz's law, that magnetic field works to oppose the original changing magnetic field, causing the rotor to move in the same direction as the changing magnetic field produced by the stator.

There are many advantages to AC induction motors, one of which is they only have one moving part: the rotor. DC motors have more parts that will wear out and need replacing. A disadvantage to simple AC induction motors is that their speed is fixed at the rate of the AC source (50 Hz).

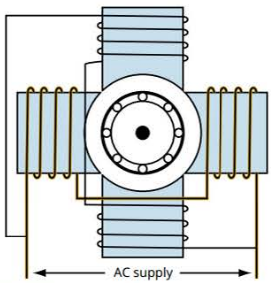


FIGURE 8.1.11 An AC induction motor is made up of pairs of electromagnets.

PHYSICS IN ACTION ICT S CCT

Electric cars

Many early cars (in the nineteenth and early twentieth centuries) ran on electricity, but during the last century the internal combustion engine came to dominate the car market. Recently, in an effort to reduce pollution and save energy, the electric car (Figure 8.1.1) has had a resurgence. One of the central components to the electric car is the energy source—the battery—which provides a DC current. However, the majority of electric cars use AC motors rather than DC motors. This may seem counter-intuitive, but there are many reasons for this choice, including cost, reliability of components, and wear and tear.

FIGURE 8.1.12 In an effort to reduce pollution and save energy, the electric car has made a comeback.

MAGNETIC BRAKING

Conventional braking systems utilise the friction force between two objects pressed together to slow an object down. Examples of this include car or bicycle brakes. It is also possible to use magnetic braking—or eddy current braking—where the electromagnetic force between a magnet and a conductor in motion is used to create the drag force.

GO TO ➤ Section 7.2 page 197

If a conductor moves past a stationary magnet, eddy currents (circular electric currents) will be induced in the conductor by the magnet, according to Faraday's law of induction. Eddy currents were introduced in Chapter 7. Due to Lenz's law, the eddy currents create their own magnetic field that opposes the original magnetic field of the magnet. This in turn creates a drag force between the magnet and the conductor, which slows the conductor down.

The magnetic field in a magnetic brake can be created either by a permanent magnet or an electromagnet. The advantage of using an electromagnet is that the magnet, and hence the braking system, can be turned on and off by adjusting the current in the electromagnet's windings.

A big advantage with magnetic braking is the lack of physical contact between components, which makes for low maintenance and few replacement parts. A disadvantage with magnetic braking is that when there is no motion between the magnet and conductor, there is no static force to maintain the conductor at rest. In this case, the magnetic braking system would need to be supplemented by a friction-based (conventional) braking system.

Magnetic braking is used in several industrial applications, such as high-speed trains (Figure 8.1.13) and to quickly slow down power tools.

FIGURE 8.1.13 High-speed trains use magnetic braking to slow down.

8.1 Review

SUMMARY

- The magnetic force experienced by a current-carrying conductor, moving at an angle θ to the magnetic field, is given by $F = IlB\sin\theta$.
- There is a torque on a coil of wire carrying a current whenever the current is not parallel to the field. Torque is defined as: $\tau = nIA_{\perp}B = nIAB\sin\theta$ where θ is defined as the angle between the area vector normal to the plane of the coil and the magnetic field.
- The wire coil of a simple DC motor keeps rotating because the direction of current, and hence the torque, is reversed each half turn by the commutator.
- In the case of a single square or rectangular coil, the total torque applied to the coil will be twice the torque acting on one side.
- The armature of a practical motor consists of many coils that are fed current by the commutator when they are in the position of maximum torque.
- The total torque in a motor will be the sum of the torques on all the individual coils.
- In an AC induction motor, the stator creates a rotating magnetic field that induces a current in the conductors of the rotor. Through application of Faraday's and Lenz's laws, the rotating magnetic field of the stator effectively pulls the rotor.

KEY QUESTIONS

1 For which of the following situations is torque at a maximum?

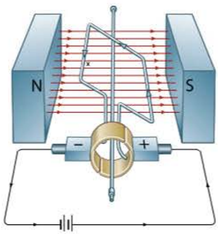
A when the force is applied perpendicular to the axis of rotation

B when the force is applied parallel to the axis of rotation

C when the force is applied at a maximum regardless of direction

D when the force applied is zero

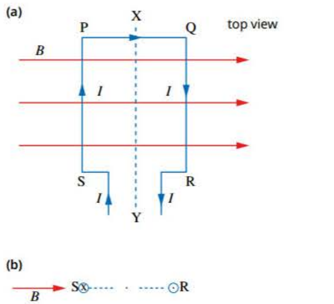
2 The following diagram shows a coil with current *I* inside of a magnetic field. Using the right-hand rule, draw the direction of the force to find the direction of rotation.



3 Part a of the diagram above right depicts a top view of a single current-carrying coil in an external magnetic field *B*.

Part b of the diagram is the corresponding cross-sectional view as seen from point *Y*. The following data applies:

$B = 0.10\,\text{T}$, $PQ = 2.0\,\text{cm}$, $PS = QR = 5.0\,\text{cm}$, $I = 2.0\,\text{A}$.



a What is the magnitude and direction of the magnetic force acting on side PS?

b What is the magnitude and direction of the magnetic force acting on side QR?

c What is the magnitude of the magnetic force acting on side PQ?

d The coil is free to rotate about an axis through XY. In what direction, as seen from Y, would the coil rotate?

e Which of the following does not affect the magnitude of the torque acting on this coil?

A the dimensions of the coil

B the magnetic field strength

C the magnitude of the current through the coil

D the direction of the current through the coil

f What is the total torque acting on the coil when the plane of the coil is parallel to the magnetic field?

4 Briefly describe how an AC induction motor works.

8.2 Generators

PHYSICS INQUIRY

DIY microphone

How has knowledge about the motor effect been applied to technological advances?

COLLECT THIS...

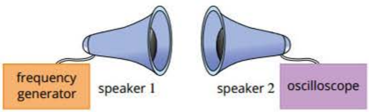
- two identical speakers
- oscilloscope
- frequency generator
- electrical leads

DO THIS...

1 Connect the oscilloscope to one of the speakers.

2 Talk into the speaker, and observe the signal produced by the speaker on the oscilloscope. Try humming different frequency sounds.

3 Set up the second speaker as shown below. If your oscilloscope can display two signals, connect the positive lead from the first speaker into the oscilloscope as well.



4 Turn on the frequency generator to a set frequency. Observe the signal being generated by the second speaker on the oscilloscope. Record the frequency being generated and the frequency displayed on the oscilloscope.

RECORD THIS...

Describe how the energy is transferred and transformed during this activity.

Present your results as a table, including a description of how each speaker is behaving.

REFLECT ON THIS...

How has the knowledge about the motor effect been applied to technological advances?

What are some applications that use the motor effect?

How accurate is the second microphone as a transducer (a device that converts energy from one form to another)?

These days we take the supply of electric power to our homes, schools and businesses for granted. The electric **generator** is probably the most important practical application of Faraday's discovery of electromagnetic induction. The principle of electric power generators is the same whether the result is **alternating current** or direct current. Relative motion between a coil and a magnetic field induces an emf in the coil. In small generators, the coil is rotated within a magnetic field, but in large power stations, car alternators and other industrial-level electric power production, the coils are stationary and an electromagnet rotates inside them.

This might all sound quite similar to the way electric motors work. In fact, it is—a generator is basically just the inverse of a motor.

Induced emf in an alternator or generator

A basic electric generator, or **alternator**, consists of many coils of wire wound on an iron core framework. This is called an armature and it is made to rotate in a magnetic field. The axle is turned by some mechanical means—mechanical energy is being converted to electrical energy—and an emf is induced in the rotating coil.

Consider a single loop of wire in the generator shown in Figure 8.2.1. The loop is rotated clockwise in a uniform magnetic field, B. The amount of flux threading through the loop will vary as it rotates. It is this change in flux that induces the emf.

Lenz's law tells you that as the flux in the loop decreases from position a to b in Figure 8.2.1, the induced current will be in a direction such as to restore a magnetic field in the same direction, relative to the loop, as the external field. The right-hand grip rule can then be used to show that the induced current flows in the direction D → C → B → A.

The direction of the induced current will reverse every time the plane of the loop reaches a point perpendicular to the field. The magnitude of the induced emf will be determined by the rate at which the loop is rotating. It will be a maximum when the rate of change of flux is a maximum. This is when the loop has moved to a position parallel to the magnetic field and the flux through the loop is zero, i.e. the gradient of the flux versus angle graph shown in Figure 8.2.2 is a maximum.

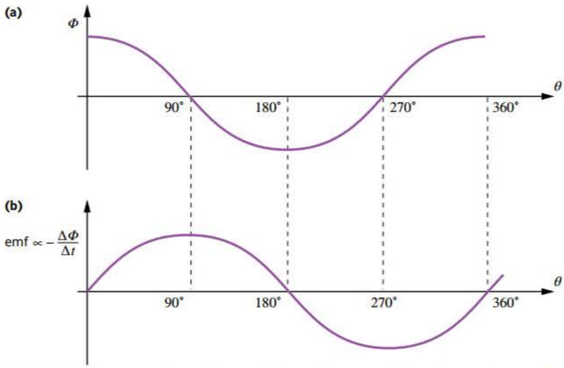


FIGURE 8.2.2 (a) The flux, $\boldsymbol{\Phi}$, through the loop of Figure 8.2.1 as a function of the angle between the field and the normal to the plane of the area, $\boldsymbol{\theta}$. (b) The rate of change of flux and hence emf through the loop as a function of the angle between the field and the normal to the plane of the loop, $\boldsymbol{\theta}$. The loop is rotating at a constant speed.

An alternative way to think about how the emf changes as the loop rotates is to remember that the emf is actually created as the wires AB and CD cut across the magnetic field lines. Maximum emf occurs when these wires cut the magnetic field lines perpendicularly, when θ is 90° or 270°, and zero emf occurs when the motion of these wires is parallel to the field lines, when θ is 0°, 180° or 360°.

AC generators and alternators

A generator's construction is basically the same as a motor. The main components of an AC generator are shown in Figure 8.2.3.

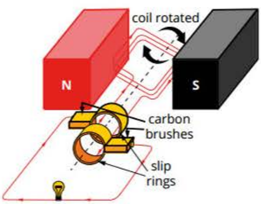


FIGURE 8.2.3 A schematic of an AC generator showing the main features.

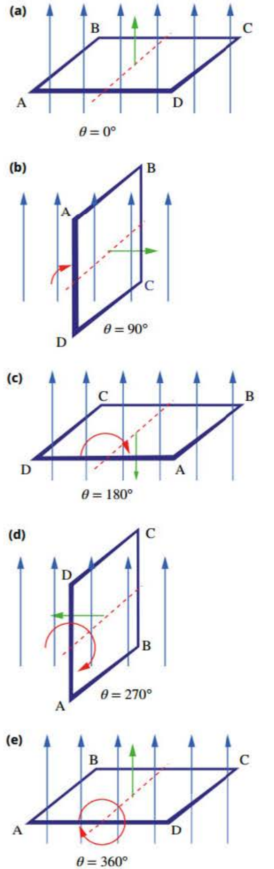


FIGURE 8.2.1 A single loop of a generator rotating in a magnetic field. (a) The plane of the area of the loop is perpendicular to the field B, and the amount of flux $\boldsymbol{\Phi} = B_{\parallel}A$ is at a maximum. (b) The loop has turned one-quarter of a turn and is parallel to the field; $\boldsymbol{\Phi} = 0$. (c) As the loop continues to turn, the flux increases to a maximum but in the opposite sense relative to the loop in (a); $\boldsymbol{\Phi} = -B_{\parallel}A$. (d) The flux then decreases to zero again as the loop is parallel to the field before repeating the cycle again from (e) onwards.

Consider a coil, or armature, with a number of turns, being rotated in a magnetic field, inducing an emf as shown previously in Figure 8.2.2 on page 223. The resultant emf alternates in direction as shown by the graph going above and below the zero emf line in Figure 8.2.2. This type of emf or voltage produces an alternating current (AC) in the coil. How this alternating current in the coil is harnessed determines if the device is an AC alternator or a DC generator.

As was stated earlier, many industrial generators will instead keep the coils still and make the electromagnet rotate within them. The principle of inducing an emf is the same. The coil itself may take a variety of shapes, sizes and positions.

If the output from the coils is transferred to a circuit via continuous **slip rings**, the alternating current in the coil will be maintained at the output. The slip rings also allow the coil to rotate without tangling. Carbon brushes press against the slip rings to allow a constant output to be transferred to a circuit without a fixed point of connection.

PHYSICS IN ACTION ICT

Three-phase generators

Many industrial applications require a more constant maximum voltage than is possible from a single coil. These applications require a three-phase power supply. The coils are arranged such that the emfs vary at the same frequency (Figure 8.2.4a), but with the peaks and troughs of their waveforms offset to provide three complementary currents with a phase separation of one-third of a cycle, or 120° (Figure 8.2.4b). The resulting output of all three phases maintains an emf near the maximum voltage more continuously. Standard electrical supplies include three phases, but most home applications only require a single phase to be connected.

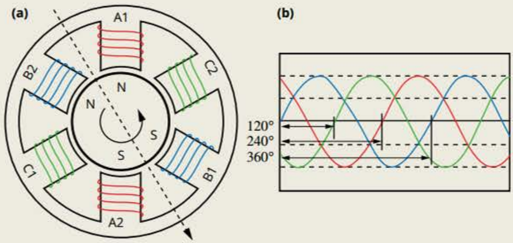


FIGURE 8.2.4 (a) A three-phase power supply has three coils, each producing an output 120° out of phase with the adjoining coil. (b) The resulting output can be combined for a more constant supply voltage.

DC GENERATORS

A DC generator is much like an AC generator or alternator in basic design. The continuous slip rings are replaced by a **split ring commutator**. That is, the ring picking up the output from the coils has two breaks (or splits) in it at opposite sides of the ring. The direction of the output is changed by the commutator every half turn so that the output current is always in the same direction (Figure 8.2.5a). The output will still vary from zero to a maximum every half cycle. The output can be smoothed by placing a capacitor in parallel with the output. More commonly, the use of multiple armature windings and more splits in the split ring commutator can smooth the output by ensuring that the output is always connected to an armature that is in the position for generating maximum emf (Figure 8.2.5b).

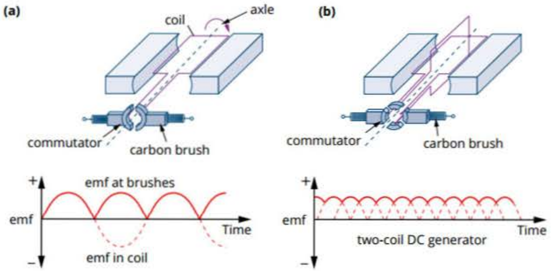


FIGURE 8.2.5 (a) A DC generator has a commutator to reverse the direction of the alternating current every half cycle and so produces a DC output. (b) Multiple armature windings can smooth the output.

In the past, cars used DC generators to power ancillary equipment. More common now is the use of AC generators or alternators, which avoid the problems of wear and sparking across the commutator inherent in the design of DC generators by using a moving electromagnet inside a set of stationary coils to generate current.

WS 6.6

PHYSICSFILE CCT

Back emf in DC motors

The description of the construction and operation of a generator shows that a DC motor and a generator share a lot in common and may even function either way. In fact, every motor can also be used as a generator. The motors of electric trains, for instance, work as generators when a train is slowing down, converting kinetic energy to electrical energy and putting it back into the electrical supply grid. Regenerative brakes in cars work in a similar way. A DC motor will also generate an emf when running normally. This is termed the 'back emf'.

The back emf generated in a DC motor is the result of current produced in response to the rotation of the rotor inside the motor in the presence of an external magnetic field. The back emf, following Lenz's law, opposes the change in magnetic flux that created it, so this induced emf will be in the opposite direction to the emf creating it. The net emf used by the motor is thus always less than the supplied voltage:

$$\varepsilon_{net} = V - \varepsilon_{back}$$

As the motor increases speed, the current induced in it will increase and the back emf will also increase. When a load is applied to the motor, the speed will generally reduce. This will reduce the back emf and increase the current in the motor. If the load brings the motor to a sudden halt—say, an electric drill bit getting stuck—the current may be high enough to burn out the motor and the motor windings. To protect the motor, a resistor is placed in series. It is switched out of the circuit when the current drops below a predetermined level and is switched back into the circuit for protection once the level is exceeded.

ALTERNATING VOLTAGE AND CURRENT

An AC generator produces an alternating current varying sinusoidally over time with the change in magnetic flux. The maximum emf is only achieved for particular points in time. In Australia, mains power oscillates at 50 Hz and reaches a peak voltage of 340 V each cycle or a peak-to-peak voltage of 680V (Figure 8.2.6).

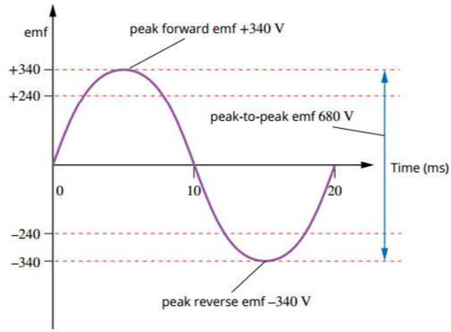


FIGURE 8.2.6 The voltage in Australian power points oscillates between +340 V and −340 V, 50 times each second. The value of a DC supply that would supply the same average power is 240 V.

It is often more useful to know the average power produced in a circuit. The average power can be obtained by using a value for the voltage and current equal to the peak values divided by $\sqrt{2}$. This is referred to as the **root mean square** or rms value.

In effect, the rms values are the values of a DC supply that would be needed to provide the same average power as the AC supply. It is the rms value of the voltage $\left(\frac{340}{\sqrt{2}}=240\right)$ that is normally quoted. This is the effective average value of the voltage and is the value that should be used to find the actual power supplied each cycle by an AC supply. So:

$$V_{rms} = \frac{V_p}{\sqrt{2}}$$

$$I_{rms} = \frac{I_p}{\sqrt{2}}$$

$$P_{rms} = V_{rms} \times I_{rms} = \frac{1}{2}V_pI_p, \text{ and}$$

$$P_p = \sqrt{2}V_{rms} \times \sqrt{2}I_{rms} = 2V_{rms}I_{rms}$$

+ ADDITIONAL

Deriving the root mean square formulae

In an AC circuit, the power produced in a resistor is equal to $\frac{V^2}{R}\sin^2\theta$.

The average power will be given by:

$$\frac{1}{2}\frac{V_p^2}{R}$$

If this same power was to be supplied by a steady (DC) source, the voltage V_{ave} of this source would have to be such that:

$$\frac{V_{ave}^2}{R} = \frac{1}{2}\frac{V_p^2}{R}$$

Simplifying:

$$V_{ave}^2 = \frac{V_p^2}{2}$$

$$V_{ave} = \frac{V_p}{\sqrt{2}}$$

This voltage is known as the root mean square voltage or V_{rms}. It is the value of a steady voltage that would produce the same power as an alternating voltage with a peak value equal to $\sqrt{2}$ times as much (Figure 8.2.7).

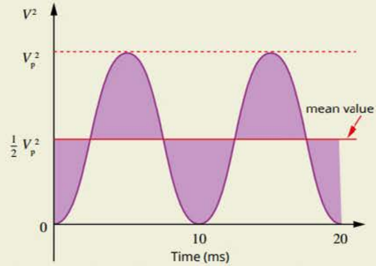


FIGURE 8.2.7 The power transmitted is proportional to the area under a V^2 graph. The power transmitted by an AC circuit (with V_p) is the same as that in a DC circuit with a voltage equal to the square root of $\frac{1}{2}(V_p)^2$; that is $\frac{V_p}{\sqrt{2}}$.

Worked example 8.2.1

PEAK AND RMS AC CURRENT VALUES

A 60 W light globe is connected to a 240 V AC circuit. What is the peak power use of the light globe?

Thinking	Working
Note that the values given in the question represent rms values. Power is given by $P = VI$ so both V and I must be known to calculate the power use. The voltage, V, is given, and the current, I, can be calculated from the rms power supplied.	$P_{rms} = V_{rms}I_{rms}$ $I_{rms} = \frac{P_{rms}}{V_{rms}}$ $= \frac{60}{240} = 0.25$ A
Substitute in the known quantities and solve for peak power, P_p.	$P_p = \sqrt{2}V_{rms} \times \sqrt{2}I_{rms} = 2V_{rms}I_{rms}$ $= 2 \times V_{rms} \times I_{rms}$ $= 2 \times 240 \times 0.25$ $= 120$ W

Worked example: Try yourself 8.2.1

PEAK AND RMS AC CURRENT VALUES

A 1000 W kettle is connected to a 240 V AC power outlet. What is the peak power use of the kettle?

8.2 Review

SUMMARY

- The principle of electric power generators is the same whether the result is alternating current or direct current. Relative motion between a coil and a magnetic field induces an emf in the coil.
- The construction of a generator or an alternator is very similar to that of an electric motor.
- A coil rotated in a magnetic field will produce an alternating induced current in the coil. How that current is harnessed will determine if the device is an AC alternator or a DC generator.
- An AC alternator has slip rings that transfer the alternating nature of the current in the coil to the output. A DC generator has a split-ring commutator to reverse the current direction every half turn so that the output current is always in the same direction.
- The alternating current produced by power stations and supplied to cities varies sinusoidally at a frequency of 50 Hz. The peak value of the voltage of domestic power (V_p) is ±340 V, and the peak-to-peak voltage (V_{p-p}) is 680 V.
- The root mean square voltage, V_{rms}, is the value of an equivalent steady voltage (DC) supply that would provide the same power.
  $V_{rms} = \frac{V_p}{\sqrt{2}}$
- The rms value of domestic mains voltage in Australia is 240 V.
- The average power in a resistive AC circuit is:
  $P = V_{rms}I_{rms}$
  $= \frac{1}{2} \times V_p \times I_p$

KEY QUESTIONS

1 The back emf generated in a DC motor is the result of current produced in response to the rotation of the armature in the motor in the presence of an external magnetic field. As a result of the back emf, what will the net emf used by a DC motor be?

A the same as the supplied voltage

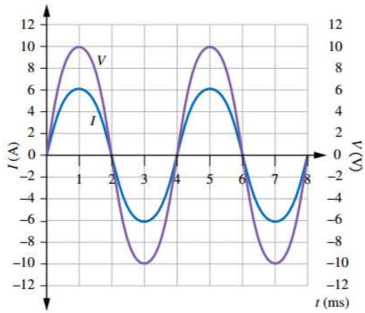
B less than the supplied voltage

C greater than the supplied voltage

D greater or less than the supplied voltage, depending on the speed of the motor

2 An AC supply of frequency 50 Hz is connected to a circuit, resulting in an rms current of 1.0 A being observed. Draw a graph that shows one full period of the variation of current with time for this circuit.

3 A student decides to test the output power of a new amplifier by using a voltage sensor to capture and display the alternating current I and voltage V that it produces. The result is shown in the graph.



What is the rms power rating of the amplifier?

4 An electric toaster designed to operate at a V_{rms} of 240 V has a power rating of 600 W. What is the peak current in the heating element?

5 Describe a benefit of three-phase power generation.

6 A 45 W light globe is connected to a 240 V AC circuit. What is the peak power use of the light globe?

Chapter review

08

KEY TERMS

alternating current
alternator
armature
brushes
commutator
direct current
electromagnet
generator
root mean square
slip rings
split ring commutator
stator
torque

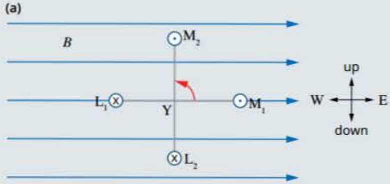
REVIEW QUESTIONS

The following information applies to questions 1–4.

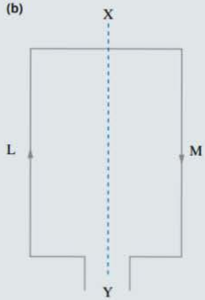
Diagram (a) below shows an end-on view of a current-carrying loop, LM. The loop is free to rotate about a horizontal axis XY. You are looking at the loop from the Y end of the axis. The same loop is seen from the top in figure (b). Initially, arms L and M are horizontal (L_1–M_1). Later they are rotated so that they are vertical (L_2–M_2). The loop is located in an external magnetic field of magnitude *B* directed east (at right angles to the axis of the loop). Note the current directions in part a: out of the page in M and into the page in L.

To answer the following questions use the direction conventions up–down and W–E as shown in part a.

(a)



(b)



1 When LM is aligned horizontally (L_1–M_1), what is the direction of the magnetic force on:
   a side L
   b side M?

2 In what direction, as seen from Y, will the loop rotate?

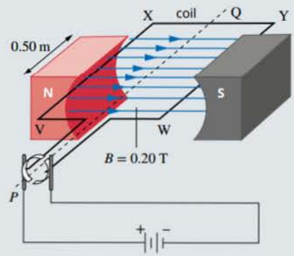
3 When LM is aligned vertically (L_2–M_2), what is the:
   a direction of the magnetic force on side L
   b direction of the magnetic force on side M
   c magnitude of the torque acting on the loop? Give a reason for your answer.

4 When LM is aligned vertically, which one of the following actions will result in a torque acting on the coil that will keep it rotating in an anticlockwise direction? (Assume it still has some momentum when it reaches the vertical position.)
   A decreasing the current through the loop
   B increasing the magnetic field strength
   C reversing the direction of the current through the coil

5 Briefly explain the function of the commutator in an electric motor.

The following information applies to questions 6–8.

The diagram shows a simplified version of a DC motor.



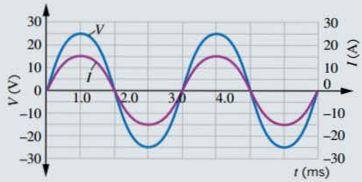
6 Calculate the magnitude of the force on segment WY when a current of 1.0 A flows through the coil.

7 In which direction will the coil begin to rotate? Give your reasoning.

8 Which of the following actions would cause the coil to rotate faster?

A increasing the current
B increasing the magnetic field strength
C increasing the cross-sectional area of the coil
D all of the above

9 A physics student uses a voltage/current sensor to display the current, *I*, through, and the voltage, *V*, across, the output terminals of a small generator. The graph obtained from the display is shown below.



a What is the approximate rms voltage for the signal?
b Calculate the peak power output of the generator.

10 A student decides to test the power output of a new stereo amplifier. The maximum rms power output guaranteed by the manufacturer (assumed accurate) is 60 W. Which set of specifications is consistent with this power output?

Peak voltage (V)	Peak current (A)
A 20	3.0
B 40	6.0
C 40	12.0
D 20	6.0

11 A single square wire coil, with a side length of 10.0 cm, is free to rotate within a magnetic field, *B*, of strength 1.0×10^{-4} T. A current of 2.0 A is flowing through the coil. What is the magnitude of the torque on the coil when the plane of the coil is parallel to the magnetic field?

12 A single square wire coil, with a side length 20.0 cm, is free to rotate within a magnetic field, *B*, of strength 5.00×10^{-4} T. A current of 1.0 A is flowing through the coil. What is the torque on the coil when the plane of the coil is at an angle of 60° to the magnetic field?

13 Five square wire coils of side length 15 cm are free to rotate within a magnetic field, *B*, of strength 2.50×10^{-4} T. A current of 3.00 A is flowing through the coil. What is the torque on the coil when the plane of the coil is at an angle of 45° to the magnetic field?

14 Describe the purpose of a commutator in a DC motor.

15 Discuss the advantages of three-phase voltage generation.

16 A washing machine designed to operate at a V_{rms} of 240 V has a power rating of 2.4 kW. What is the peak current in the heating element?

17 Describe one method to produce a steadier DC output voltage from a DC generator.

18 A 2 kW oil column heater is connected via a wall outlet to a 240 V AC circuit. What is the peak power use of the heater?

19 What is the purpose of brushes in a DC motor?

20 Explain what back emf is, and how it can affect a DC motor.

21 After completing the activity on page 222, reflect on the inquiry question: How has the knowledge about the motor effect been applied to technological advances?

MODULE 6 • REVIEW

REVIEW QUESTIONS

Electromagnetism

WS 6.7 WS 6.8 MR 6

Multiple choice

1 Which of the following is the best description of how a transformer transfers electrical energy from the primary windings to the secondary windings?

A The current through the primary windings produces a constant electric field in the secondary windings.

B The current through the primary windings produces a steady magnetic field in the secondary windings.

C The current through the primary windings produces a changing magnetic field in the secondary windings.

2 When a transformer is plugged in to the 240 V mains but nothing is connected to the secondary coil, very little power is used. What is the best explanation for this?

A The primary and secondary coils are in series and so no current can flow in either if the secondary coil is open.

B There can be no magnetic flux generated in the transformer if the secondary coil has no current in it.

C The magnetic flux generated by the current in the primary produces an emf that opposes the applied voltage.

D The magnetic flux generated by the secondary coil almost balances out that due to the primary coil.

The following graphs A–D and table apply to questions 3–5.

A

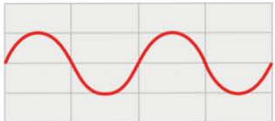
B

C

D

	f (Hz)	*B* (T)	*N*	*A* (cm²)
A	50	0.50	200	100
B	100	0.50	200	100
C	100	1.00	50	100
D	50	0.50	400	100

A simple generator consists of a coil with $N = 100$ turns and an area of $100\,\text{cm}^2$ in a uniform magnetic field of $B = 0.5\,\text{T}$. Originally, it has a frequency of 50 Hz and produces the following voltage output as a function of time.

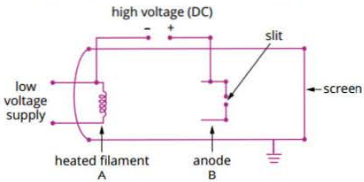


3 Which of the diagrams A–D best describes the display on the CRO when the generator is operating at a frequency of 100 Hz?

4 Which of the specifications in the table could produce a CRO display described by diagram A?

5 Which of the specifications in the table could produce a CRO display illustrated by diagram C?

6 Study the diagram of a simple cathode ray tube.



What is the source of electrons in this device?

A the heated filament at A

B the positive anode at B

C the wires used in the circuit

D the screen used in the circuit

7 Between two plates forming a uniform electric field, where will the electrical field strength be at a minimum?

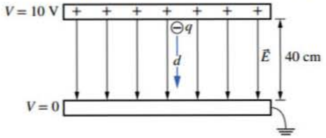
A close to the positive plate

B close to the earthed plate

C at all points between the plates

D at the mid-point between the plates

8 The following diagram shows the force on an electron in an electric field. Calculate the work done on or by the field to move the electron a distance *d* of 5.0 cm.



A 2.0×10^{-19} J work done by the field
B 2.0×10^{-19} J work done on the field
C 3.0×10^{-19} J work done by the field
D 3.0×10^{-19} J work done on the field

9 Which of the following types of fields would you *not* expect to be associated with radial fields?

A gravitational
B electrical
C magnetic
D all of the above

10 If an electron travels through a magnetic field of strength 1.2 T with a speed of 4.2×10^6 m s^{-1}, calculate the radius of the path it will follow.

A 2×10^{-5} m
B 3×10^{-5} m
C 4×10^{-5} m
D 5×10^{-5} m

11 A particular electron gun accelerates an electron across a potential difference of 15 kV, a distance of 12 cm between a pair of charged plates. Calculate the magnitude of the force acting on the electron.

A 1×10^{-13} N
B 2×10^{-13} N
C 1×10^{-14} N
D 2×10^{-14} N

12 A 5.0 Ω coil, of 100 turns and radius 3.0 cm, is placed between the poles of a magnet so that the flux is a maximum through its area. The coil is connected to a sensitive current meter that has an internal resistance of 595 Ω. It is then moved out of the field of the magnet and it is found that an average current of 50 μA flows for 2 s. Had the coil been moved out more quickly so that it was removed in only 0.5 s, what would have been the average current?

A 50 μA
B 100 μA
C 200 μA
D 400 μA

13 A magnet is dropped, with the north end facing down, through a conducting coil inducing a current. Which of the following options in the table describes the direction of the induced current when viewed from above the coil?

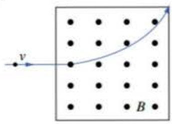
	Magnet entering the coil	Magnet exiting the coil
A	clockwise	anticlockwise
B	anticlockwise	clockwise
C	clockwise	clockwise
D	anticlockwise	anticlockwise

14 As a motor increases speed, what happens to the back emf?

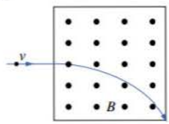
A it increases
B it decreases
C it remains the same

15 Which of the following describes the correct motion of an electron in a magnetic field?

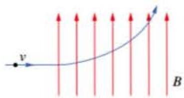
(A)



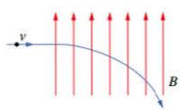
(B)



(C)



(D)



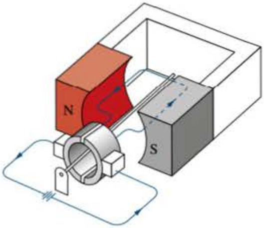
16 Determine the force per unit length acting between two current-carrying conductors, both carrying 5 mA of current in the same direction, spaced 5 cm apart. Use $\mu_0 = 4\pi \times 10^{-7}$ N A^{-2}.

A 10×10^{-10} N m^{-1}, repulsive
B 10×10^{-10} N m^{-1}, attractive
C 10×10^{-11} N m^{-1}, repulsive
D 10×10^{-11} N m^{-1}, attractive

17 Which formula can be used to show the SI definition of the ampere?

A $\frac{F}{l} = \frac{\mu_0}{2\pi}\frac{I_1 I_2}{r}$
B $F = BIl \sin\theta$
C $F = qvB \sin\theta$
D $\varepsilon = -N\frac{\Delta\Phi}{\Delta t}$

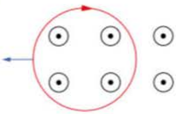
18 A DC motor set-up causes a coil to rotate in a clockwise direction. A commutator is then attached so that the current will change direction; this is shown in the diagram below. What happens to the coil when the commutator is attached?



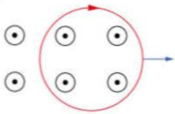
A It will turn one full rotation clockwise before turning back one full rotation anticlockwise.
B It will turn halfway clockwise before turning back halfway anticlockwise.
C It turns continuously clockwise.
D It stops spinning.

19 Which correctly describes the direction of the induced current in a coil moving out of a magnetic field?

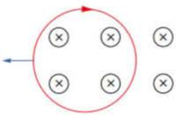
A



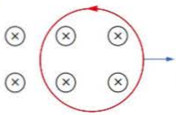
B



C



D

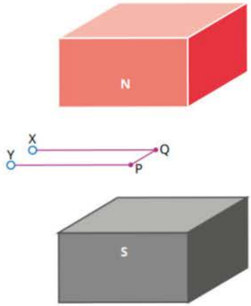


20 A transformer is built to reduce the 240 V supply voltage down to 10 V for an appliance. If the number of turns in the secondary coil is 80, what is the number of turns required in the primary coil?
A 3
B 33
C 192
D 1920

Short answer

21 Calculate the electric field strength and direction at a distance of 3.5 mm directly to the left of a charge of +9.4 μC.

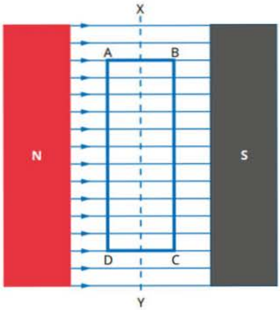
22 The following diagram shows a section of a conducting loop XQPY, part of which is placed between the poles of a magnet whose uniform field strength is 1.0 T. The side PQ has length 5.0 cm. X is connected to the positive terminal of a battery, while Y is connected to the negative terminal. A current of 1.0 A then flows through this loop.



a What is the magnitude of the force on side PQ?
b What is the direction of the force on side PQ?
c What is the magnitude of the force on a 1.0 cm section of side XQ that is located in the magnetic field?
d The direction of the current through the loop is reversed by connecting X to the negative terminal and Y to the positive terminal of the battery. What is the direction of the force on side PQ?

23 In an electron gun, an electron is accelerated by a potential difference of 28 kV.
a At what speed will the electrons exit the assembly?
b If the electron was accelerated a distance of 20 cm between a pair of charged parallel plates, then calculate the size of the electric field strength acting on the electron.

24 A rectangular coil containing 100 turns with dimensions 10 cm × 5 cm is located in a magnetic field $B = 0.25$ T, as shown. It is free to rotate about the axis XY. The coil carries a constant current $I = 200$ mA flowing in the direction ADCB.



What is the magnitude and direction of the magnetic force on the following sides?

a AB
b DC
c AD
d BC

25 This diagram shows a stream of electrons entering a magnetic field. Reproduce the diagram and show the subsequent path of the electrons through the magnetic field.

26 An electron beam travelling through a cathode ray tube is subjected to simultaneous electric and magnetic fields perpendicular to each other. The electrons emerge with no deflection. Given that the potential difference across the parallel plates X and Y is 3 kV, the applied magnetic field is of strength 1.6×10^{-3} T and the velocity of the electrons is 3.25×10^{7} m s^{-1}, calculate the distance between plates X and Y.

27 A rectangular conducting loop of dimensions 100 mm × 50 mm and resistance $R = 2.0\ \Omega$, is located with its plane perpendicular to a uniform magnetic field of strength $B = 1.0$ mT.

a Calculate the magnitude of the magnetic flux Φ threading the loop.

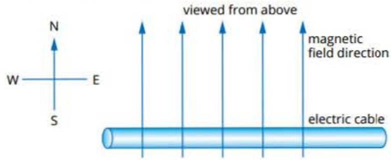
b The loop is rotated through an angle of 90° about an axis, so that its plane is now parallel to B. Determine the magnetic flux Φ threading the loop in the new position.

c The time interval for the rotation $\Delta t = 2.0$ ms. Determine the average emf induced in the loop.

d Determine the value of the average current induced in the loop during the rotation.

e Will the current keep flowing once the rotation is complete and the loop is stationary? Explain your answer.

28 The diagram below shows a horizontal, east–west electric cable, located in a region where the magnetic field of the Earth is horizontal and has a magnitude of 1.0×10^{-5} T. The cable has a mass of 0.05 kg m^{-1}.



a What is the magnitude of the magnetic force on a 1.0 m section of this cable if a 100 A current is flowing through it?

b What is the direction of the current that will produce a force vertically upwards on this cable?

c What magnitude of current would be required to produce zero resultant vertical force on a 1.0 m section of this cable?

d Assume that a 100 A current is flowing through this cable from west to east. What would be the magnitude of the change in magnetic force per metre on this cable if the direction of this current was reversed?

e The cable is no longer at right angles, but makes an angle θ with the direction of the Earth's magnetic field. How would the force on the cable change if the same current passes through it?

29 An ideal transformer is operating with peak input voltage of 600 V and an rms primary current of 2.0 A. The peak output voltage is 3000 V. There are 1000 turns in the secondary winding.

a What is the rms output current?

b What is the output peak-to-peak voltage?

c How many turns are there in the primary winding?

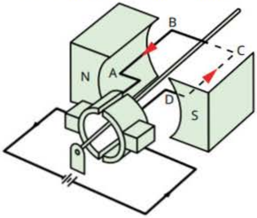
d Determine the rms power consumed in the secondary circuit.

e Calculate the peak power consumed in the secondary circuit.

30 A rectangular loop of 100 turns is suspended in a magnetic field $B = 0.50$ T. The plane of the loop is parallel to the direction of the field. The dimensions of the loop are 20 cm perpendicular to the field lines and 10 cm parallel to them.

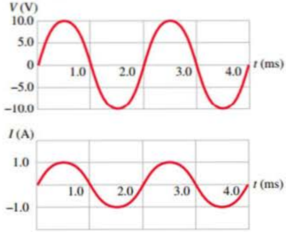
- **a** What amount of flux threads the loop in the position described above?
- **b** How can the amount of flux threading the loop be increased?
- **c** How should the plane of the loop and the magnetic field direction be arranged so the maximum possible flux threads the loop?
- **d** Calculate the maximum possible flux as described in part (c).

31 Consider the electric motor shown below.



- **a** The direction of the current in the coil is shown (from D anticlockwise to A). What is the direction of the forces on sides AB and CD?
- **b** In what position of the coil is the turning effect of the forces greatest?
- **c** At one point in the rotation of the coil the turning effect becomes zero. Explain where this occurs and why the motor actually continues to rotate.
- **d** A rectangular loop of 100 turns is suspended in a magnetic field $B = 0.50$ T. The plane of the loop is parallel to the direction of the field. The dimensions of the loop are 20 cm perpendicular to the field lines and 10 cm parallel to them. It is found that there is a force of 40 N on each of the sides perpendicular to the field. What is the current in each turn of the loop?
- **e** This loop is then replaced by a square loop of 10 cm each side, with twice the current and half the number of turns. What is the force on each of the perpendicular sides now?
- **f** The rectangular loop with the original current from part (d) is returned but a new magnet is found which provides a field strength of 0.80 T. What is the force on the 20 cm side now?

32 The following diagram shows the voltage–time graph and corresponding current–time graph for an alternator that was built by a physics student as part of a research project.

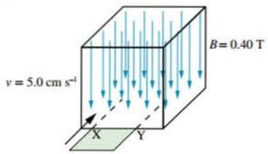


- **a** What is the frequency of the voltage produced by the alternator?
- **b** What is the peak-to-peak output voltage of this alternator?
- **c** What is the rms output voltage of the alternator?
- **d** Calculate the rms output current of the alternator.
- **e** Calculate the rms output power of the alternator.
- **f** What feature distinguishes an alternator from a DC generator?
- **g** How does an alternator operate?

33 A physics student constructs a simple generator consisting of a coil of 400 turns. The coil is mounted on an axis perpendicular to a uniform magnetic field of strength $B = 50$ mT and rotated at a frequency $f = 100$ Hz. It is found that during the rotation, the peak voltage produced is 0.9 V.

- **a** Sketch a graph showing the voltage output of the generator for at least two full rotations of the coil. Include a scale on the time and voltage axes.
- **b** What is the rms voltage generated?
- **c** The student now rotates the coil with a frequency $f = 200$ Hz. How would your answers to part (a) and (b) be affected? Include a new sketch of a graph showing the voltage output of the generator for at least two full rotations of the coil.
- **d** A second generator is rotating at a rate of 3000 revolutions per minute. The magnetic field strength is 0.50 T. The total number of turns in the armature coils is $N = 200$, each of area $A = 100\,\text{cm}^2$. Calculate the frequency of rotation of the generator.
- **e** Calculate the average emf generated during a quarter revolution of the generator coil in part (d).

34 A square conducting loop with sides 20 cm and resistance 0.50 Ω is moving with a constant horizontal velocity of 5.0 cm s^{-1} towards a region of uniform magnetic field of strength 0.40 T directed vertically downwards, as shown in the following diagram. The magnetic field is confined to a cubic region of side 30 cm.



- **a** Describe the direction of the induced current in the side XY of the loop just as it begins to enter the field. Justify your answer.
- **b** Calculate the average emf induced in the loop when it is halfway into the field.
- **c** What current flows in the loop when it is halfway into the field?
- **d** How much electrical power is consumed in the loop when it is halfway into the field?
- **e** What is the source of this power?
- **f** What is the average emf induced in the loop 5 s after it started to enter the cube? Justify your answer.
- **g** What is the direction of the induced current in the side XY just as it begins to emerge from the field? Justify your answer.

35 A farmer has installed a wind generator on a nearby hill, along with a power line consisting of two cables with a combined total resistance of 2.0 Ω. The output of the generator is given as 250 V AC (rms) with a maximum power of 4000 W. She connects up the system and finds that the voltage at the house is indeed 250 V. However, when she turns on various appliances so that the generator is running at its maximum power output of 4000 W, she finds that the voltage supplied at the house is rather low.

- **a** Explain why the voltage dropped when the farmer turned on the appliances in the house.
- **b** Calculate the voltage and power at the house when the appliances are turned on.

She then decides to install ideal transformers at either end of the same power line so that the voltage transmitted from the generator end of the line in this system becomes 5000 V.

- **c** Describe the essential features of the types of transformers that are needed at either end of the power line.

Now assume the generator is operating at full load, i.e. 4000 W.

- **d** What is the current in the power line now when the same appliances are turned on?
- **e** What is the voltage drop along the power line?
- **f** What is the power loss in the power line?
- **g** What voltage is delivered to the house?
- **h** What power is delivered to the house?
- **i** How do the power losses in the system without the transformers compare to the system with the transformers as a percentage of the power generated?
- **j** Explain why the system operated with much lower power losses when the voltage was transmitted at the higher voltage.

MODULE 7 The nature of light

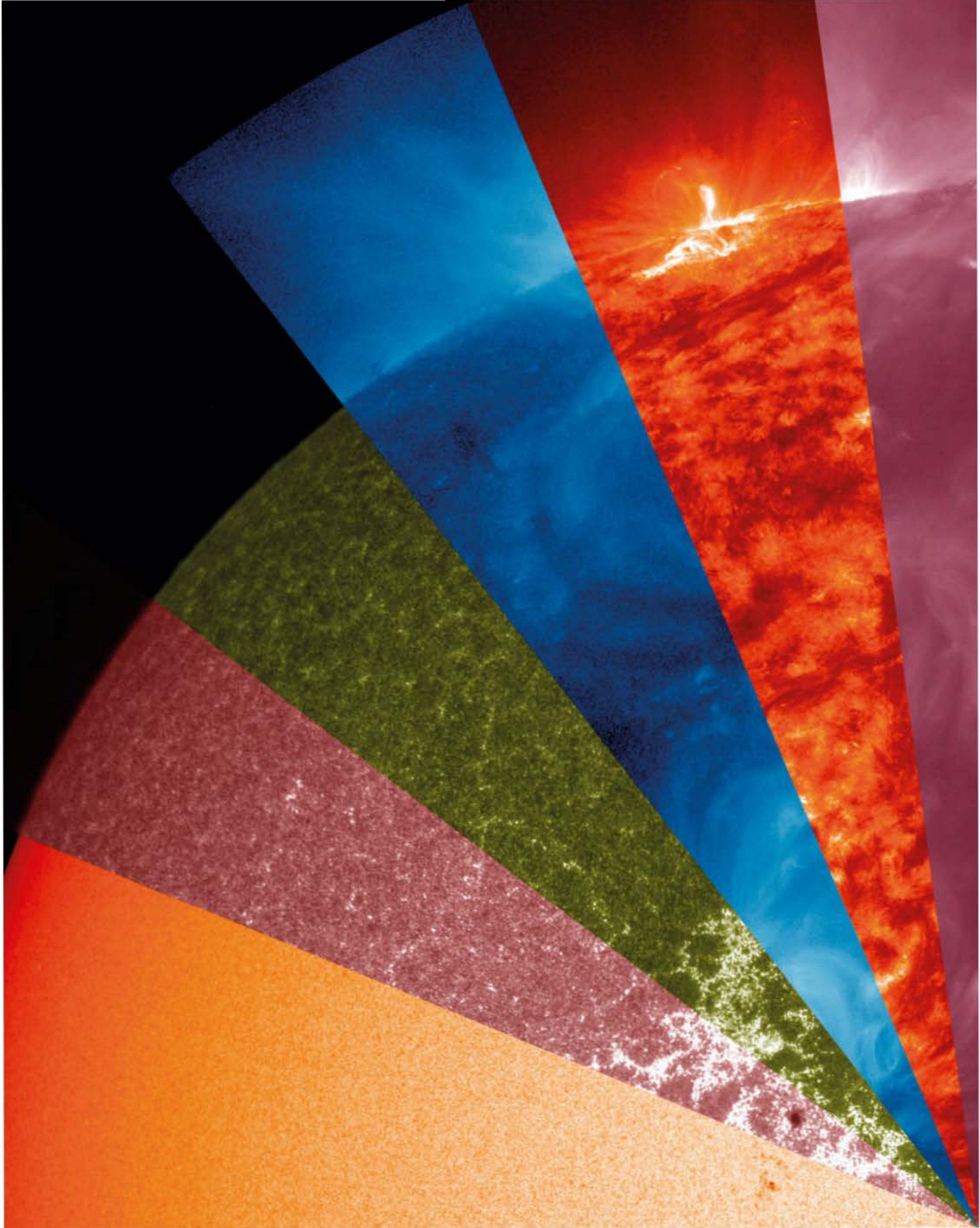
Before the 20th century, physicists, including Newton and Maxwell, developed theories and models about mechanics, electricity and magnetism and the nature of matter. These theories and models had great explanatory power and produced useful predictions. However, the 20th century saw major developments in physics, as existing theories and models were challenged by new observations that could not be explained. These observations led to the development of quantum theory and the theory of relativity. Technologies arising from these theories have shaped the modern world. For example, the independence of the speed of light on the frame of observation or the motion of the source and observer had significant consequences for the measurement of, and concepts about, the nature of time and space.

Throughout this module, you will explore the evidence supporting these physical theories, along with the power of scientific theories to make useful predictions.

Outcomes

By the end of this module you will be able to:

- develop and evaluate questions and hypotheses for scientific investigation PH12-1
- design and evaluate investigations in order to obtain primary and secondary data and information PH12-2
- conduct investigations to collect valid and reliable primary and secondary data and information PH12-3
- select and process appropriate qualitative and quantitative data and information using a range of appropriate media PH12-4
- communicate scientific understanding using suitable language and terminology for a specific audience or purpose PH12-7
- describe and analyse evidence for the properties of light and evaluate the implications of this evidence for modern theories of physics in the contemporary world PH12-14



CHAPTER 09

Electromagnetic spectrum

In this chapter you will develop an understanding of how scientists revealed the nature of light through theory and experiment. You will learn how Maxwell's theory unifies electricity and magnetism to explain light as electromagnetic radiation, and how this theory underlined the importance of the speed of light. Through exploring other frequencies of light you will learn of the wide applications of electromagnetic radiation in spectroscopy, and examine what spectroscopy has revealed about the nature of stars.

Content

INQUIRY QUESTION

What is light?

By the end of this chapter you will be able to:

- investigate Maxwell's contribution to the classical theory of electromagnetism, including:
  - unification of electricity and magnetism
  - prediction of electromagnetic waves
  - prediction of velocity (ACSPH113) CCT ICT
- describe the production and propagation of electromagnetic waves and relate these processes qualitatively to the predictions made by Maxwell's electromagnetic theory (ACSPH112, ACSPH113)
- conduct investigations of historical and contemporary methods used to determine the speed of light and its current relationship to the measurement of time and distance (ACSPH082) CCT ICT
- conduct an investigation to examine a variety of spectra produced by discharge tubes, reflected sunlight or incandescent filaments
- investigate how spectroscopy can be used to provide information about: ICT
  - the identification of elements
- investigate how the spectra of stars can provide information on: CCT ICT
  - surface temperature
  - rotational and translational velocity
  - density
  - chemical composition.

9.1 Electromagnetism

The establishment of the wave model for light raised an important question. Scientists now wanted to know what type of waves were light waves.

Experiments on polarisation provided the important information that light must be a type of transverse wave, since polarisation does not occur for longitudinal waves. Light is also obviously different from mechanical waves because it can pass through the vacuum of space between the Earth and Sun (Figure 9.1.1).

FIGURE 9.1.1 Light cannot be a simple mechanical wave because it can travel through empty space.

In 1820, the Danish physicist Hans Ørsted observed a magnetised compass needle deflected from its alignment to the Earth's magnetic field when a nearby electric circuit was switched on and off. This showed that a wire carrying an electric current generates a magnetic field, and it revealed the first evidence of a relationship between electricity and magnetism. English physicist Michael Faraday expanded this in the 1830s to show that changing magnetic fields produced electric fields. The breakthrough came in 1861 when the Scottish physicist James Clerk Maxwell quantified the relationship through a precise mathematical study of electric and magnetic effects. His work, known as Maxwell's equations, show that electric and magnetic fields move at a speed that closely match experimental estimates of the speed of light. Maxwell went on to develop a comprehensive theory of electromagnetism in which light is a form of **electromagnetic radiation** (EMR).

REVISION

Waves

A wave is a periodic oscillation that allows the transfer of energy from one place to another. The oscillations may be parallel or perpendicular to the direction of travel. Oscillations that are parallel to the direction of travel are known as longitudinal waves; for example, soundwaves. Oscillations that are perpendicular to the direction of travel are known as transverse waves; for example, ocean waves.

Waves can be characterised by the velocity of energy propagation, v; the frequency of oscillations, f; and wavelength, λ. These three variables are related through the wave equation:

$$v = f\lambda$$

Generally, waves require a medium to carry their oscillations. This type of wave is called a mechanical wave, and both sound and ocean waves are examples of mechanical waves. Light (an electromagnetic wave) does not need a medium and can travel through a vacuum.

WS 7.1

THE ELECTROMAGNETIC NATURE OF LIGHT

A point charge generates an electric field. A moving point charge—a current—generates a magnetic field, while a changing magnetic field generates an electromotive force (emf) or voltage.

Maxwell put these ideas together. He proposed that if a changing electric field is produced, for example by moving a charged particle backwards and forwards, then this changing electric field will produce a magnetic field at right angles to the electric field. The changing magnetic field would, in turn, produce a changing electric field and the cycle could be repeated. In effect, this would produce two mutually propagating fields, as shown in Figure 9.1.2. Both the electric and magnetic fields would necessarily oscillate at the same frequency. The electromagnetic radiation would be self-propagating, independent of its generator, and could extend outwards into space as an electromagnetic wave of a fixed frequency.

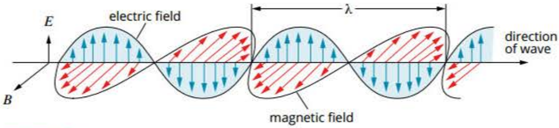


FIGURE 9.1.2 The electric and magnetic fields in electromagnetic radiation are perpendicular to each other and are both perpendicular to the direction of propagation of the radiation. The fields do not 'stick out' of the page; the illustration shows the electric and magnetic fields strengthen, maximise, weaken, vanish, invert and repeat cyclically. The frequency, and therefore the wavelength, λ, of both fields is the same.

Maxwell's theoretical calculations provided a value for the speed at which electromagnetic radiation should propagate through empty space. This closely matched the experimentally determined values for the speed of light so closely that it provided a clue that light has to be a form of EMR. The accepted value for the speed of light today is $299\,792\,458\,\text{m s}^{-1}$. This is such an important constant that it is designated its own symbol, c. In calculations, the speed of light is usually approximated as $c = 3.00 \times 10^8\,\text{m s}^{-1}$.

As EMR travels at a known speed, there is a clear relationship between frequency f and wavelength λ, expressed as $c = f\lambda$, or equivalently $\lambda = \frac{c}{f}$. This is a special form of the wave equation $v = f\lambda$ given at the start of this section.

$c = f\lambda$

where

c is the speed of light (m s^{-1})

f is the frequency of the wave (Hz)

λ is the wavelength of the wave (m)

Worked example 9.1.1

USING THE WAVE EQUATION FOR LIGHT

Calculate the frequency of violet light with a wavelength of 400 nm (i.e. 400×10^{-9} m).

Thinking	Working
Recall the wave equation for light.	$c = f\lambda$
Transpose the equation to make frequency the subject.	$f = \frac{c}{\lambda}$
Substitute in values to determine the frequency of this wavelength of light.	$f = \frac{3.00 \times 10^8}{400 \times 10^{-9}}$ $= 7.50 \times 10^{14}$ Hz

Worked example: Try yourself 9.1.1

USING THE WAVE EQUATION FOR LIGHT

A particular colour of red light has a wavelength of 600 nm. Calculate the frequency of this colour.

During the 1880s Heinrich Hertz performed experiments that demonstrated the existence of 'invisible' electromagnetic waves (radio waves). His basic design involved using a high-voltage source to produce a spark and a detector formed from a loop of copper wire with a small gap. Radio waves produced by the spark were detected when a tiny spark was induced across the gap in the detector. He found that the invisible waves behaved like visible light, and the speed of the waves was the same as the speed of light. Further experiments using reflecting surfaces and other apparatus showed that radio waves can be reflected and refracted, which is consistent with the behaviour of visible light. Hertz's experiments powerfully supported Maxwell's theory.

Maxwell's work represents a pivotal moment in the history of physics. Not only did he provide an explanation of the nature of light, and provide a strong theoretical basis for the wave model of light, he also brought together several formerly distinct areas of study—optics (the study of light), electricity and magnetism.

Maxwell's description of electromagnetic interactions is an example of a classical field theory—like Newton's formulation of gravity. Maxwell's theories involve fields that extend continuously everywhere in space, and those fields determine how matter will interact.

DETERMINING THE SPEED OF LIGHT

Since ancient times, light was known to travel faster than sound—the flash of lightning precedes the sound of thunder. During the 1600s, the Danish scientist Ole Rømer used the then-new invention of the telescope to observe the motion around Jupiter of one of its larger moons, Io. As Io revolved around Jupiter, it was obscured from view for around half of its orbit, eclipsed by Jupiter, and it was possible to measure the time between when Io disappeared and reappeared. Rømer noticed that this time between eclipses varied during the Earth's year, when Earth and Jupiter were closer and when they were further away. He reasoned that Io did not change its orbital velocity, but rather the light from Io took longer to arrive. He deduced that light takes around 22 minutes to cross the diameter of Earth's orbit.

Further calculations with Rømer's data would show that the speed of light would be around 220 000 km s^{-1}. This result was not generally accepted during Rømer's lifetime. Better estimates of the size of Earth's orbit, which required the Earth–Sun distance, would improve this estimation. (James Cook's 1769 voyage to Tahiti to measure the transit of Venus across the Sun, and so measure the Earth–Sun distance, would later help refine these estimates.) Around half a century later, in 1727, the English astronomer James Bradley provided a closer estimate at 290 000 km s^{-1} through an elegant and subtle observation of two stars.

Earth-bound measurements of the speed of light arose in the 1840s. In France, Hippolyte Fizeau shone a bright light through a toothed cogwheel. As the wheel turned, sometimes light would be blocked by the cogs. Fizeau set up a mirror 8 km away from the cog in order to measure the time-of-flight for light to travel the 16 km path (Figure 9.1.3). As the pulse of light also needed to pass again through the teeth on the way back for it to be observed, knowing the speed of rotation of the wheel and knowing the distance between the cogwheel and the mirror, Fizeau was able to calculate the speed of light to be approximately 315 000 km s^{-1}.

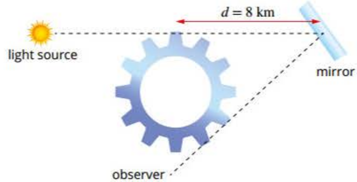


FIGURE 9.1.3 Diagram of Fizeau's apparatus. The first passage through the cogwheel breaks the light beam into pulses with a length dependent on the speed of the wheel's rotation. Whether the light passes the cogwheel a second time depends on the speed of the cogwheel and the speed of light.

Fizeau's collaborator Léon Foucault refined this basic method using a rotating mirror to block the light's path, and by 1862 determined a result of 298 000 km s^{-1}, just over 0.5% off the modern estimate. One of the final time-of-flight measures was during Albert Michelson's experiments in the 1930s, with an accuracy of ±11 km s^{-1}. Note that all of these measurements are for light in air, not light in a vacuum.

During the 1950s, the development of the laser, which provided a light source with a known fixed wavelength, allowed ever more accurate interference-based measurements. By the 1970s the uncertainty in measurement was $299\,792\,456.2 \pm 1.2\,\text{m s}^{-1}$.

Finally, by 1983, the measurements for the speed of light had become so reliable that c became an SI standard unit, fixed to precisely 299 792 458 m s^{-1}. The metre is now defined in terms of c, as the distance light will travel during $\frac{1}{299792458}$ of a second.

PHYSICS **IN ACTION** CCT N

Measuring the speed of light at home

A microwave oven uses EMR at a known frequency to cook food. By determining the wavelength of the microwaves, the speed of light can be easily calculated.

If marshmallows are placed in the microwave oven in which the rotating plate is removed, one layer thick (Figure 9.1.4) and cooked at a low setting, they will begin to melt in different spots. This is because microwave ovens do not cook evenly and the melted areas are at the hottest spots in the microwave. (The uneven cooking is why microwave ovens need a rotating plate.)

If you cook the marshmallows until four or five melted spots appear, by measuring the distance between each melted area you can find the speed of light. The melted spots are the antinodes of the microwave radiation, and the distance between them will correspond to half the wavelength of the microwave: about 6 cm for household microwave ovens. By finding the frequency of the microwave (generally around 2.450 GHz) and multiplying it by the wavelength, it is possible to calculate the speed of light.

FIGURE 9.1.4 Microwave ovens produce electromagnetic radiation with a frequency of 2.45 GHz, which is the resonant frequency of water molecules. Melting marshmallows is one way to determine the positions of the microwave antinodes.

THE ELECTROMAGNETIC SPECTRUM

PA 7.1 PA 7.1

The wavelengths of all the different colours of visible light are between 390 nm and 780 nm. Naturally, physicists were bound to inquire about other wavelengths of electromagnetic radiation. It is now understood that the visible spectrum is just one small part of a much broader set of possible wavelengths known as the **electromagnetic spectrum** (Figure 9.1.5).

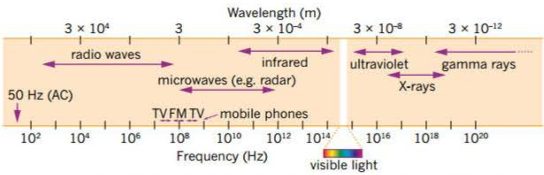


FIGURE 9.1.5 The electromagnetic spectrum. Note that both wavelength and frequency are on logarithmic scales.

Changing the frequency and wavelength of the waves changes the properties of the electromagnetic radiation, and so the electromagnetic spectrum is divided into 'bands', according to how the particular types of EMR are used. The shorter the wavelength of the electromagnetic wave, the greater its penetrating power. This means that waves with extremely short wavelength, such as X-rays, can pass through some materials (e.g. skin), revealing the structures inside (e.g. bone). On the other hand, long wavelength waves such as AM radio waves have such low penetrating power that they cannot even escape Earth's atmosphere, and can be used to 'bounce' radio signals around to the other side of the world. Table 9.1.1 on page 244 compares the characteristics of different waves in the electromagnetic spectrum.

TABLE 9.1.1 A comparison of the different waves in the electromagnetic spectrum and their effects on matter.

Type of wave	Typical wavelength (m)	Typical frequency (Hz)	Comparable size	Effect on matter
AM radio wave	100	3×10^6	sports oval	causes movement of free electrons in a conductor
FM radio or TV wave	3	1×10^8	small car	
microwaves	0.03	1×10^{10}	50c coin	causes molecular rotation
infrared	10^{-5}	3×10^{13}	white blood cell	makes chemical bonds vibrate
visible light	10^{-7}	3×10^{15}	small cell	affects electronic states in atoms or chemical bonds
ultraviolet	10^{-8}	3×10^{16}	large molecule	
X-ray	10^{-10}	3×10^{18}	atom	excites electrons in atomic orbitals
gamma ray	10^{-15}	3×10^{23}	atomic nucleus	causes atomic nuclei disintegration

PHYSICSFILE CCT

Why do we see visible wavelengths?

The Earth's atmosphere blocks many types of EMR (Figure 9.1.6). The highest levels of the atmosphere, the ionosphere, contains charged particles and effectively blocks the high-energy ionising EMR (gamma rays, X-rays). Lower down, molecular ozone, O_3, and nitrogen, N_2, absorbs and blocks about 70% of the ultraviolet EMR. Visible light is transmitted well, as it is not energetic enough to be absorbed. The atmosphere becomes increasingly opaque in the infrared and microwave bands due mainly to absorption by water vapour. At still-lower energies the atmosphere becomes transparent again to shorter-wavelength radio waves, until the lowest energy longer-wavelength EMR cannot penetrate the atmosphere.

During our evolution our eyes have developed photoreceptors ('cones') that employ the visible spectrum. It is not the only option, however. Dogs have two types of photoreceptor, green and blue, which enable them to see blue, green, and yellow. Humans have three types, sensitive to red, green, and blue, which allow us to see colours derived from red, such as orange and purples, that are invisible to dogs. Honeybees also have three types, but the evolution of bees led to their photoreceptors being sensitive to ultraviolet, blue, and green, which makes the pollen of flowers stand out more strongly. Butterflies have five types,

FIGURE 9.1.7 The magnificent mantis shrimp.

and the mantis shrimp (Figure 9.1.7) has sixteen types of photoreceptor, including a type sensitive to polarisation. We see an entire rainbow with just three photoreceptors. What must a mantis shrimp see?

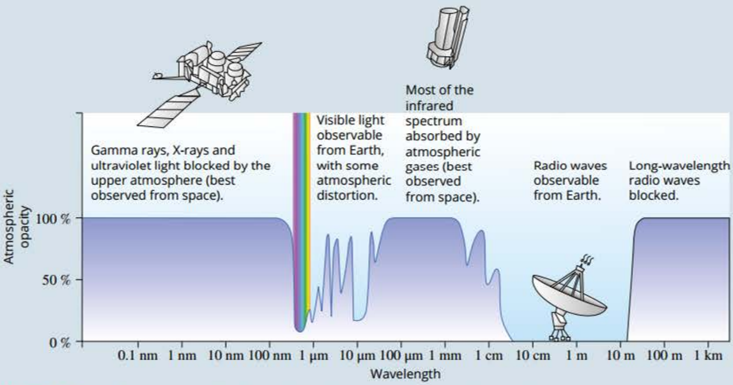


FIGURE 9.1.6 Depending on wavelength, Earth's atmosphere is transparent, translucent or opaque.

Radio waves

One of the most revolutionary applications of EMR is the use of radio waves to transmit information from one point to another over long distances. Radio waves are the longest type of electromagnetic radiation, with wavelengths ranging from one millimetre to hundreds of kilometres. Radio waves are generated when an electrical field changes; conversely, radio waves will change an electrical field.

A radio transmitter converts the signal (e.g. a radio announcer's voice, music or a stream of data) into an alternating current. When this alternating current flows in the transmission antenna, the electrons in the antenna oscillate backwards and forwards. This oscillation of charges in the antenna produces a corresponding electromagnetic wave that radiates outwards in all directions from the antenna.

When the radio wave hits the antenna of a radio receiver, the electrons in the receiver's antenna start to oscillate in the same way as in the transmitting antenna. The radio receiver then reverses the process of the transmitter, converting the alternating current from the reception antenna back into the original signal as seen in Figure 9.1.8.

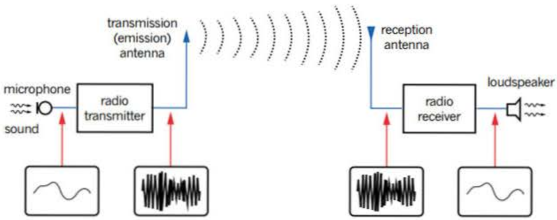


FIGURE 9.1.8 A typical radio transmission system.

The antenna will pick up many radio signals at one time, so the transmitter adds the signal to a carrier wave of a fixed frequency, and the receiver 'tunes into' the same carrier frequency. Many radio stations use the carrier-wave frequency as part of their name; for example, Nova 96.9 transmits using a 96.9 MHz carrier wave.

The carrier wave is altered or 'modulated' by the signal containing the information to be transmitted. An AM radio system uses 'amplitude modulation', which means that the amplitude of the carrier wave is modulated to match the signal. In comparison, FM stands for 'frequency modulation', in which the frequency of the carrier wave is changed to represent the signal. In terms of circuitry, AM systems are much simpler than FM systems, although FM radio waves tend to transmit signals more clearly.

Microwaves

Microwaves have shorter wavelengths and therefore greater penetrating power than radio waves. They can be produced by devices with short antennae and hence are useful in personal communication applications like mobile phones and wireless internet transmission.

Microwaves are also particularly useful in heating and cooking food. All solid objects have a frequency at which they will naturally vibrate. A common example is stringed instruments, such as guitars or violins, which make use of the resonant frequencies of strings under tension. When water molecules are bombarded with radiation with a frequency of 2.45 GHz, they start to rotate quickly. Since this increases the average kinetic energy of the water molecules, the temperature of the water in the substance increases. Effectively, the microwaves cause the water to heat up. This heat then transfers by conduction and convection to the rest of the food.

This is why food sometimes becomes soggy when heated in the microwave: the water molecules heat up faster than the other food molecules around them. It also explains why recipes that do not contain much water, or frozen foods in which the molecules cannot move, cannot be cooked well in a microwave oven.

Infrared

The infrared section of the electromagnetic spectrum lies between microwaves and visible light. Infrared waves are longer than the red waves of the visible spectrum, hence their name ('infra' is from the Latin word for 'below').

Infrared light waves are emitted by objects, to varying degrees, due to their temperature. The warmth that you feel standing next to an electric heater or fire is due to infrared radiation (Figure 9.1.9). The radiant heat the Earth receives from the Sun is transmitted in the form of infrared waves; life on Earth would not be possible without this important form of electromagnetic radiation.

FIGURE 9.1.9 The coals of a fire appear red because they release red light along with infrared radiation, which you experience as heat.

At the boundary between the microwave and infrared bands lies the terahertz band, EMR with wavelengths from 0.1 to 1.0 mm (frequencies of 3–0.3 THz, where 1 THz = 10^{12} Hz). Terahertz EMR does not penetrate liquid water or metal, and is non-ionising, so this radiation is useful in medical imaging and security applications.

Ultraviolet light

As their name suggests, ultraviolet (UV) waves have wavelengths that are shorter than those of violet light and therefore cannot be detected by the human eye. Their shorter wavelengths give UV rays stronger penetrating power than visible light. In fact, UV rays can penetrate human skin and damage the DNA of skin cells, producing harmful skin cancers.

Scientists can make use of UV light to take images. Figure 9.1.10 on page 247 is a UV image of the surface of the Sun taken after a solar flare has occurred. The image has been recoloured so that it highlights areas of different temperature. Here, areas that are coloured white are the hottest. Images like this help scientists to learn about the temperatures of very hot objects. Taking an image of the Sun using visible light would not allow this same distinction.

FIGURE 9.1.10 A recoloured UV image of the surface of the Sun. The white areas show the hottest parts.

X-rays and gamma rays

X-rays and gamma rays have much shorter wavelengths than visible light. This means that these forms of electromagnetic radiation have very high penetrating power. For example, some X-rays can pass through different types of human tissues, which means that they are very useful in medical imaging (Figure 9.1.11).

Unfortunately, this useful penetrating property of X-rays comes with inherent dangers. As X-rays pass through a human cell, they can do damage to the tissue, sometimes killing the cells or damaging the DNA in the cell nucleus, leading to harmful cancers. For this reason, a person's exposure to X-rays must be carefully monitored and limited to avoid harmful side effects.

Similarly, exposure to gamma rays can be very dangerous to human beings. The main natural sources of gamma radiation exposure are the Sun and radioactive isotopes. Fortunately, Earth's atmosphere protects people from most of the Sun's harmful gamma rays, and radioactive isotopes are not commonly found in sufficient quantities to produce harmful doses of radiation.

FIGURE 9.1.11 This X-ray image of a child's hips can be formed because X-rays can pass through human tissue.

PHYSICSFILE N

Violent explosions

Sometimes when a massive star collapses to form a black hole, a large burst of radiation occurs called a gamma-ray burst. These events are extremely energetic, with the radiation being emitted in strong jets (Figure 9.1.12). If the gamma rays were instead released isotropically (meaning in all directions), then this event would release approximately 10^{47} J of energy in less than a minute.

FIGURE 9.1.12 A gamma-ray burst shoots gamma radiation out as jets from a collapsing star.

9.1 Review

SUMMARY

- Although light exhibits many wave properties, it cannot be modelled as a mechanical wave because it can travel through a vacuum.
- The physicist James Maxwell unified the theories of electricity and magnetism to find that light is a form of electromagnetic radiation.
- Electromagnetic waves are transverse waves made up of mutually perpendicular, oscillating electric and magnetic fields.
- Electromagnetic radiation can be used for a variety of purposes depending on the properties of the waves, which are determined by their frequency.
- Oscillating charges produce electromagnetic waves of the same frequency as the oscillation. Electromagnetic waves cause charges to oscillate at the frequency of the wave.
- Light (that is, all electromagnetic radiation) travels through a vacuum at approximately $c = 3.00 \times 10^8\,\text{m s}^{-1}$.
- The wave equation $c = f\lambda$ can be used to calculate the frequency and wavelength of electromagnetic waves.
- Light and matter interact in ways that depend on the energy of the electromagnetic wave.
- The electromagnetic spectrum consists of radio waves, microwaves, infrared, visible light, ultraviolet, X-rays and gamma rays.

KEY QUESTIONS

1 Which of the following does *not* apply to both light waves and mechanical waves?

A wavelengths are measurable
B can travel through a vacuum
C speed of wave is measurable
D waves undergo diffraction

2 In an electromagnetic wave, what is the orientation of the changing electric and magnetic fields?

A at 45° to each other
B parallel to each other
C parallel to each other but in opposite directions
D perpendicular to each other

3 What type of electromagnetic radiation has a wavelength of 200 nm?

A radio waves
B microwaves
C visible light
D ultraviolet light

4 Arrange the types of electromagnetic radiation below in order of decreasing wavelength.
FM radio waves/X-rays/visible light/infrared radiation

5 Calculate the frequencies of the following wavelengths of light. Use $c = 3.00 \times 10^8\,\text{m s}^{-1}$.

a red, wavelength 656 nm
b yellow, wavelength 589 nm

9.2 Spectroscopy

PHYSICS INQUIRY N CCT

Spectrograph

What is light?

COLLECT THIS…

- DVD or CD
- craft knife or other thin, sharp instrument
- strong sticky tape
- cardboard
- scissors

DO THIS…

1. Carefully using the knife, split the DVD into two layers. Discard the top clear layer.
2. Using a piece of strong sticky tape, remove as much of the metal coating as possible.
3. Use the scissors to cut the DVD into quarters.
4. Using the cardboard, cut out a shape to be used as a set of frames for glasses. Use a DVD piece as the glasses lens. Put tape around the edges to make sure no sharp sections remain.
5. Put on your glasses and look at different light sources. Try an LED light, a fluorescent light, an incandescent globe, a bright window and a neon lamp.

RECORD THIS…

Describe how the different light sources look through your glasses.

Present your results by drawing and colouring the spectra from the different light sources.

REFLECT ON THIS…

What is light?

How do different colours of light behave when they go through the DVD lens? (Hint: compare long wavelengths like red light to shorter wavelengths like blue light.)

What do the different spectra tell us about the light source?

The branch of science investigating the spectra produced when matter interacts with or emits electromagnetic radiation is referred to as **spectroscopy**. The spectra are analysed by investigating the intensity of EMR at particular wavelengths. Spectroscopy is widely used in chemistry to detect and characterise atoms and molecules, as well as in astronomy and sensing technologies. Spectroscopy has many specialist fields depending on the nature of energy—EMR, particles (commonly electrons and neutrons), or mechanical methods—and the type of interactions with matter (absorption, emission, scattering, and more). In this section only EMR will be covered. Spectroscopy paved the way to understanding the atomic nature of matter and quantum theory, overturning the wave theory of light. This will be explored more deeply in Chapter 11.

GO TO ➤ Section 11.1 page 287

ABSORPTION SPECTRA

In 1814, the German physicist Joseph von Fraunhofer reported many dark lines appearing in the spectrum of sunlight, as shown in Figure 9.2.1.

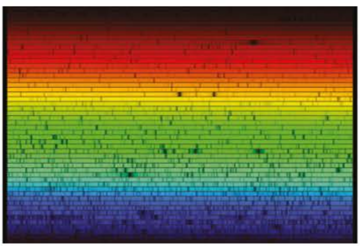


FIGURE 9.2.1 The spectrum of sunlight contains some missing colours known as Fraunhofer lines.

PHYSICSFILE CC ICT

CSIRO

Atomic absorption spectroscopy in its modern form, in which the sample is reduced to its constituent atoms in a flame, was developed by the CSIRO in the 1950s and remains one of the oldest and most widely used forms of spectroscopy. There are many other forms of absorption spectroscopy using different EMR sources. Different wavelengths of interest and detectors can be used in spectroscopic studies of molecules due to the wide range of energies that can interact with molecules. Absorbed EMR can affect molecular rotations, molecular vibrations and molecular electronic states. Table 9.1.1 on page 244 lists some of the frequencies of EMR and its effects on matter.

Essentially Fraunhofer repeated Newton's experiments on dispersion where light is split into its component colours by passing it through a prism. He observed the spectrum (as expected) but also noticed that there were some colours 'missing' from the spectrum. The missing colours appeared as black lines at various points along the spectrum. These apparently missing colours came to be known as Fraunhofer lines.

About 50 years later, scientists including Gustav Kirchhoff and Robert Bunsen (of Bunsen burner fame) recognised that some of these lines corresponded to the colours emitted when certain gases were heated to high temperatures. They deduced that the dark lines were due to these colours (wavelengths or frequencies) being **absorbed** by gases as light made its way through the outer atmosphere of the Sun. This is the Sun's **absorption spectrum**.

Absorption spectra are valuable for scientists who wish to know what elements are present in a sample of gas or in a solution, so their use is not limited to just astronomy. First, light is directed through a cool sample of a gas or through a solution containing an element or compound. Only certain wavelengths (or frequencies) of light will be absorbed by the elements present in the sample, which means that this wavelength will be 'missing' from the spectrum. The wavelengths that are absorbed are unique to each type of atom. For this reason, by analysing which wavelengths are missing—absorption spectroscopy—scientists can determine exactly what elements are present in the sample.

EMISSION SPECTRA

When elements are heated to high temperatures or have an electrical current passed through them, they produce light. Atoms within the material absorb energy and become 'excited' (more on what this means later in Chapter 15). This makes the atom unstable, and eventually it will return to the 'unexcited' or **ground state**. When this happens, the energy that had been absorbed is released. The colour of light that is emitted will depend on the amount of energy it has.

GO TO ➤ Section 15.1 page 394

Since atoms can usually have many different **excited states**, they can produce many different colours. The combination of colours produced by a particular element is distinctive to that element (Figure 9.2.2) and is known as its **emission spectrum**.

FIGURE 9.2.2 The different metals used in fireworks are responsible for the colours in this display. For example, strontium gives red, sodium gives yellow, and copper gives green.

PHYSICSFILE ICT

Measuring spectra

A **spectroscope** (Figure 9.2.3) is a simple device combining a diffraction grating with a viewing telescope, which allows the spectrum of light from a source such as a distant star to be viewed directly. When a graduated scale is added to the spectroscope, the specific wavelengths in the spectrum can be determined. This is known as a spectrometer, and it has become key to analysing the spectra from distant stars. Modern spectrometers use electronic means to record the specific wavelengths of the incident radiation, but they function on the same basic principle.

FIGURE 9.2.3 A model of the spectroscope invented by William Crooke in the nineteenth century.

Incandescent filaments

As mentioned in the previous section, light is emitted when atoms move from an excited state to the ground state. An **incandescent** light bulb, however, produces light by heating a metal filament to a very high temperature (Figure 9.2.4a). This produces electromagnetic radiation at a range of wavelengths; that is, the light from the incandescent globe is a continuous spectrum. Some of the light produced is in the infrared part of the spectrum that is invisible to human beings but is detected as heat. In Figure 9.2.4b, you can see the full spectrum from an incandescent light bulb. The visible spectrum is between about 400 nm and 750 nm, so radiation longer than 750 nm is invisible infrared EMR. Note also that there is a sharp drop of light in the ultraviolet, at wavelengths shorter than 400 nm.

(a)

(b)

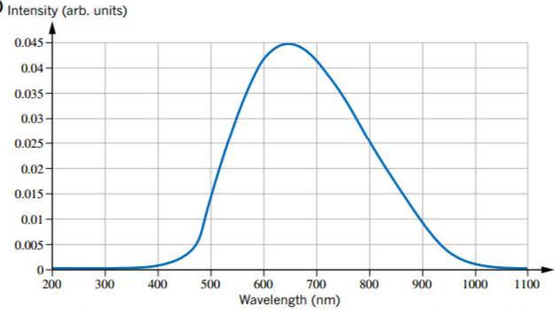


FIGURE 9.2.4 (a) In an incandescent light globe, electricity is passed through a tungsten filament. As the filament heats up, a wide range of wavelengths are emitted. (b) The spectrum of a 230 V, 60 W incandescent light globe.

PA 7.2 PA 7.2

Discharge tubes

Fluorescent lights are a type of **discharge tube** (Figure 9.2.5), which contain a low-pressure gas through which a current is passed. The light emitted by the gas is in the ultraviolet range. The inside of the glass is coated with a material called a phosphor. The ultraviolet light excites the phosphor, which then emits light over the entire visible spectrum. Fluorescent lights emit less energy in the infrared range, so more efficiently convert the electrical energy into light energy and remain cool to touch. They are more expensive to manufacture than an incandescent globe, but this initial cost is recouped with substantially lower running costs. In domestic use an 8 W compact fluorescent globe emits the same light intensity as a 60 W incandescent globe—the same light output for 13% of the energy required.

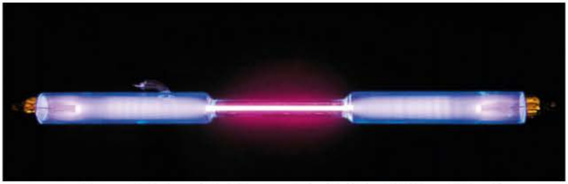


FIGURE 9.2.5 Fluorescent discharge tube. This example is a deuterium (^2H) tube, which emits predominantly in the ultraviolet range.

The spectra of incandescent and fluorescent lights differ. Incandescent sources present a continuous spectrum, with bands greatly broadened on account of the higher temperatures, whereas fluorescent sources are 'spiky' (Figure 9.2.6), with discrete emissions at narrow wavelength bands.

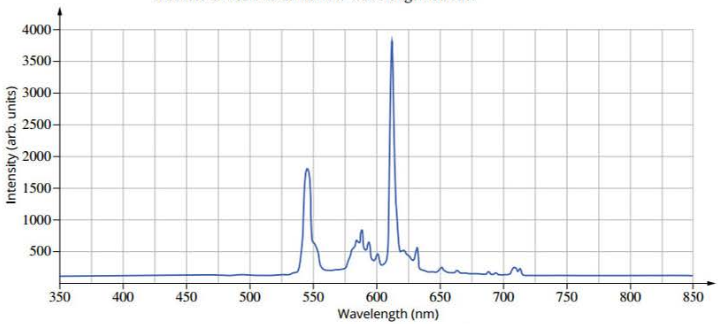


FIGURE 9.2.6 Fluorescent sources produce a 'spiky' spectrum.

Reflected sunlight

Planets and moons do not emit light in the visible range, but rather reflect light from a nearby luminous body like a star. What is not reflected is absorbed, so looking carefully at the reflected sunlight provides information on what is in the planet or moon's atmosphere or on its surface. The process of matching the missing bands with the absorbing substance is essentially the same as determining elements from a flame or in a star. However, a planetary atmosphere is cooler than a star's atmosphere and so contains complicated molecules, rather than simple atoms, which are harder to identify from absorption lines.

Saturn's moon Titan (Figure 9.2.7) is known to have an atmosphere, and measurements of the reflected light in 1944 showed the presence of methane. In 1980 the *Voyager 1* probe and the more recent 2004 *Cassini–Huygens* mission showed the atmosphere is mostly nitrogen, with the methane located high in Titan's atmosphere; lower down can be found trace amounts of ethane, propane, carbon dioxide and carbon monoxide. The hydrocarbons are now known to come from the seas of ethane on the moon. The methane is a mystery, as it should have been converted into other hydrocarbons, and suggests there is a source of methane on Titan that is still unknown. Comet impact as a source of methane can't be the case, as there is too little carbon monoxide.

This thinking is also used to explore for life on other planets. The amount of oxygen in Earth's atmosphere is not in thermodynamic equilibrium with the other gases. There must be a source of oxygen on Earth—we know the source as vegetation—and that as the disequilibrium persists, the source of oxygen must be self-renewing, or, in other words, alive.

By probing the reflected light from distant bodies, astrobiologists hope to find other atmospheres not in chemical equilibrium, and perhaps this might be a signature for life somewhere else in the universe. Perhaps there is methane-based life on the chilly (−179°C) surface of Titan?

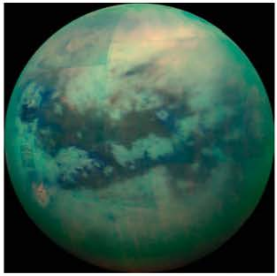


FIGURE 9.2.7 An infrared view of Saturn's moon Titan. In the visible spectrum the moon would only show a hazy orange; the collected infrared light reveals the deeper layers of the atmosphere and surface.

OBSERVING STARS

An important way that spectroscopy is used is in the observation of stars. Astrophysical observations use the full range of the electromagnetic spectrum to reveal the nature of the universe. The observations can be conveniently classed by wavelength range (Figure 9.2.8).

Radio astronomy studies emission from cold objects such as interstellar gas and the cosmic microwave background radiation residual from the big bang.

Submillimetre astronomy is used to look at molecular clouds and galaxy formation.

Infrared astronomy is used to study star clusters, proto-stars and planets.

Optical (visible) astronomy is the oldest form, and reveals the chemical composition and temperature of stars.

Ultraviolet astronomy is used to study hotter objects such as young stars or stars approaching nova.

X-ray astronomy studies very hot objects at temperatures of millions of degrees.

Gamma-ray astronomy studies nuclear processes like supernova explosions and matter–antimatter interaction.

Radio telescopes, such as the one at Parkes in NSW, are used to collect the longer wavelength radio and microwave radiation. Because infrared radiation is absorbed by the Earth's atmosphere, satellites equipped with special infrared telescopes which orbit above Earth's atmosphere are also used to photograph the sky at those wavelengths. While the atmosphere is transparent to wavelengths which are in the visible range or shorter than the visible wavelengths, the atmosphere still causes some distortion of these wavelengths. This distortion ultimately limits the power of Earth-based telescopes to detect these wavelengths accurately. For this reason, space-based telescopes such as the Hubble Space Telescope have proved to be of immense value to astronomers in detecting shorter ultraviolet and X-ray wavelengths. The properties and classification of stars will be covered in more detail in Chapter 13.

FIGURE 9.2.8 The Sun observed at different wavelengths within the electromagnetic spectrum. If you look at the Sun through a camera, it will look like a yellow circle because you are viewing visible light only. Viewing wavelengths outside the visible spectrum can reveal extra information about the Sun's surface and its atmosphere.

GO TO ➤ Section 13.2 page 355

Surface temperature

Electromagnetic radiation is emitted by all objects and systems whose temperature is above absolute zero (0 K). The actual wavelength or frequency of the emitted radiation depends almost entirely on the internal energy of the object and not on the characteristics of the material itself. The wavelength at which the peak intensity occurs, λ_{max}, gives an indication of the temperature of the surface of the object, T. This relationship is known as Wien's law. The relationship between wavelength and temperature using Wien's law is discussed in more detail in Chapter 11. At this stage, just remember that the peak wavelength of the emitted light is inversely proportional to the temperature of the object.

GO TO ➤ Section 11.1 page 284

Wien's law

$\lambda_{max} \propto \frac{1}{T}$

Chemical composition

As stars are hot objects, they emit light across a range of wavelengths (that is, a spectrum). If the light from the star then passes through a cloud of gas, the gas will absorb certain wavelengths, leaving dark bands in the spectrum corresponding to the absorbed wavelengths. This process is shown in Figure 9.2.9. The dark bands are shown in Figure 9.2.9b. The absorbed light (in the cloud of gas) temporarily puts the atoms in the gas into an excited state. As the atoms return to the ground state, light is released. The re-emitted light radiates out in all directions, so if seen from the direction shown in Figure 9.2.9c, it appears as bright bands. The bright bands form what is known as an emission spectrum. The bright bands of the emission spectrum will therefore correspond to the dark bands of the absorption spectrum.

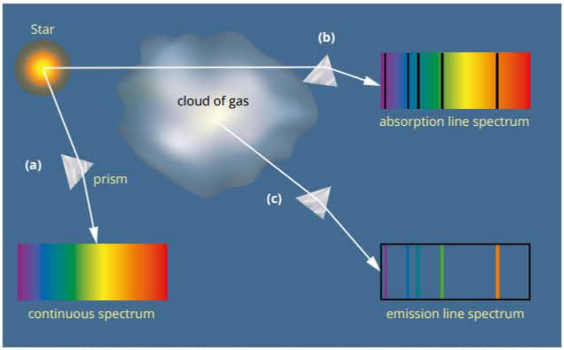


FIGURE 9.2.9 The hot object at the top emits a continuous range of wavelengths with the intensity dependent upon its temperature. (a) The spectrum of wavelengths can be seen when viewed through a prism. (b) If the light passes through a gas, some wavelengths are absorbed, giving rise to an absorption spectrum. (c) Re-emitted light from the gas produces an emission spectrum.

The measurement of a star's visible spectrum gives clues to its chemical composition. The characteristic yellow of the sodium lamp arises from emissions of 589.0 nm and 589.6 nm wavelengths. Equivalently, detection of bands at those frequencies in a star's light is evidence that there is sodium in the star's atmosphere. Through analysis of the bands, astrophysicists determine the chemical composition of stars.

The lines emitted or absorbed are characteristic of each of the 92 chemical elements. For example, sodium always emits or absorbs two very characteristic lines in the yellow region. This explains why a barbecue flame turns yellow when you throw salt on the food—salt is the compound sodium chloride. If these two lines are seen in a spectrum it means that sodium is present. Whether it is present as part of a compound (in the case of sodium chloride) or as a vaporised gas makes no difference—the lines always have the same position in the spectrum. So, by examining the spectrum of a star, it is possible to tell what elements are present at its surface or in its environment.

The Sun's absorption spectrum (shown at the top of Figure 9.2.10) has been examined in great detail and many elements have been identified. In 1868, some lines were observed that did not seem to correspond to any element known on Earth. At the time this element appeared to be exclusive to the Sun and was called 'helium' (from the Greek word ***helios***, meaning 'sun').

The absorption spectra from other stars are quite similar to the solar spectrum in most respects. This confirms that the basic chemistry and physics understood on Earth seems to apply throughout the universe. No strange new lines have ever been discovered in any other stellar spectra. There are subtle differences, however, and it is these differences that tell astronomers an enormous amount about other stars. Figure 9.2.10 shows the absorption spectra for a variety of astronomical objects, including the Sun.

In the case of our Sun, it is 71% hydrogen, 27% helium, and less than 2% other elements by mass. Over 65 elements have been detected in the solar spectrum, of which the top ten by composition is listed in Table 9.2.1. Other stars differ slightly but studies since the 1920s have shown that other stars and nebulae have similar compositions—approximately 90% hydrogen, 10% helium and tiny traces of everything else.

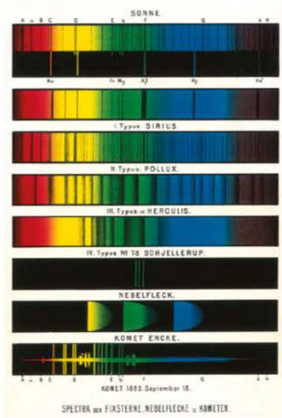


FIGURE 9.2.10 A historical 1884 diagram of the spectra of the Sun (shown as 'Sonne' in German) and other astronomical objects showing the absorption spectra known as Fraunhofer lines. The absorption lines are indicative of the presence of particular elements in the atmosphere of the stars.

TABLE 9.2.1 The composition of the Sun as determined by spectroscopy.

Element	Abundance (by number of atoms)	Abundance (by mass)
hydrogen	91.2%	71.0%
helium	8.7%	27.1%
oxygen	0.078%	0.97%
carbon	0.043%	0.40%
nitrogen	0.0088%	0.096%
silicon	0.0045%	0.099%
magnesium	0.0038%	0.076%
neon	0.0035%	0.058%
iron	0.0030%	0.14%
sulfur	0.0015%	0.040%

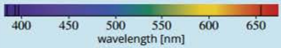
Temperature

You have already seen that the spread of colour through the continuous spectrum is an indication of the surface temperature of the star. The greater the proportion of blue and UV, the hotter the star.

In the 1890s, a scheme for classifying the spectra of stars was developed. It arranged the spectra in an order determined by the presence or absence of certain lines associated with hydrogen and labelled them alphabetically with the letters from A to O. It seemed to be a logical arrangement at the time, but later it became apparent that there was a better way to arrange them.

The new method involved placing similar spectra next to one another and finding a pattern which smoothly changed from one class of spectrum to the next over the whole range. As a result, some of the original classifications were dropped and the rest were re-ordered into a new sequence: OBAFGKM (Table 9.2.2). Later modifications divided each class into smaller subdivisions by adding the digits 0–9 after the letters. Our Sun, with a surface temperature of around 5800 K, is in class G2.

TABLE 9.2.2 The OBAFGKM classification of the spectra of stars.

Spectral type	Approximate temperature (K)	Main characteristics	Spectrum	Colour
O	50000–28000	relatively few lines; the lines of ionised helium		blue
B	28000–10000	the lines of neutral helium		blue–white
A	10000–7500	very strong hydrogen lines		white
F	7500–6000	strong hydrogen lines; ionised calcium lines; numerous metal lines		yellow–white
G	6000–4900	strong ionised calcium lines; numerous strong lines of ionised and neutral iron and other metals		yellow
K	4900–3500	strong lines of neutral metals		orange
M	3500–2000	titanium oxide streaks		red

The real significance of this pattern was not known at the time. It was later found that the changes in the spectra corresponded to differing temperatures, with the O stars being the hottest and the M stars being comparatively cooler. It was shown that the changes between the classes were associated with the fact that different atoms become ionised (lost electrons) at different temperatures. At cooler temperatures the light may not have enough energy to excite the atoms sufficiently to create some of the lines. So it was realised that lines may appear over a certain temperature range but not at higher or lower temperatures.

This discovery provided a new way to determine the temperature of a star which was not affected by colour change as light travelled through interstellar gas. The difference between the two methods gave astronomers useful information about the interstellar gases that might be present between the star and the Earth.

A careful analysis of a spectrum reveals much important information about the star:

- its temperature
- the elements that are present and the state in which those elements exist (solid, liquid, gas)
- the pressure of the gas emitting the light
- any magnetic fields present.

PHYSICS IN ACTION ICT WE

Sarah Martell, astronomical spectroscopist

Dr Sarah Martell (Figure 9.2.11) uses spectroscopy to study the origins, orbits and chemical compositions of the stars in the Milky Way. She is a pioneer in the field of 'chemical tagging', which uses the compositions of stars to identify stars that formed together and to explore the conditions in which they formed. This area of research tells us how elements have been produced and recycled through multiple generations of stars across the lifetime of the Milky Way.

As a student, Dr Martell was intrigued by the many ways that mathematics is expressed in the real world, including through physics and music. At university, she became fascinated with how much can be learned indirectly about astronomical objects—it's not possible to take samples of them for laboratory study, but astronomers can measure the masses of planets, map the internal structure of stars, and even infer the history of entire galaxies using just the light they emit.

Dr Martell is a Senior Lecturer in the School of Physics at the University of New South Wales. She is a leader in the GALAH (Galactic Archaeology with HERMES) survey, a major international project mapping the compositions of stars in the Milky Way from Siding Spring Observatory in northwestern New South Wales. Large-scale studies of the Milky Way are a major theme in modern astronomical research, as part of a drive to understand galaxy formation and growth through in-depth study of the Milky Way and its nearby neighbours.

FIGURE 9.2.11 Sarah Martell with the Antu telescope, one of four that make up the Very Large Telescope facility operated by the European Southern Observatory in the Atacama desert in Chile.

Rotational and translational velocity

The position of known absorption or emission lines are frequently shifted from where they would be expected to be (Figure 9.2.12). The visible spectrum of hydrogen has four wavelengths, at 410, 434, 486 and 656 nm, but an observed star may show wavelengths at 414, 438, 491 and 663 nm. This is not a new type of hydrogen, but rather evidence that the star is moving away from us. The wavelengths are slightly longer—they have been **redshifted**. (Wavelengths can also be **blueshifted**, if a star is approaching and the observed wavelengths shorten.) This effect is often referred to as the Doppler effect (see the Revision box). The extent of the redshift reveals how fast the star is moving.

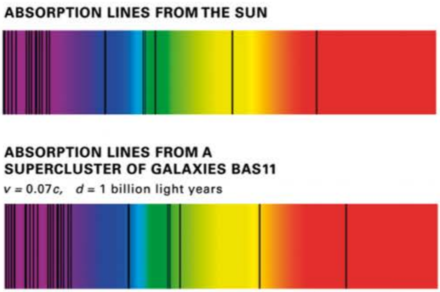


FIGURE 9.2.12 In the lower spectrum the Fraunhofer lines are redshifted (moved to the right), indicating the absorbing object is moving away from the observer.

REVISION

The Doppler effect

The Doppler effect is a phenomenon of waves that is observed whenever there is relative movement between the source of the waves and an observer. It causes an apparent increase in frequency when the relative movement is towards the observer (i.e. the distance between observer and wave source is decreasing) and an apparent decrease in frequency when the relative movement is away from the observer (i.e. the distance between observer and wave source is increasing). It can be observed for any type of wave, and it has been particularly useful in astronomy for understanding the expanding universe.

The rotation of a star broadens the observed atomic absorption bands (Figure 9.2.13). Due to rotation, light emitted from the side of the star rotating away from the observer is redshifted and light from the side of the star rotating towards the observer is equally blueshifted, meaning the observed band smears out over a range of wavelengths. The degree of broadening in the star's spectrum reveals the rate of rotation. The star Archenar has an equatorial velocity of over $250\,km\,s^{-1}$ ($900\,000\,km\,h^{-1}$) and a rotational period of 2.1 days, and the spin gives the star an equatorial diameter 56% larger than its pole-to pole diameter. Regulus A rotates even more quickly; if Regulus A rotated just 15% faster, the rotation would tear the star apart. The faster a star rotates, the wider the absorption bands will appear in its spectrum.

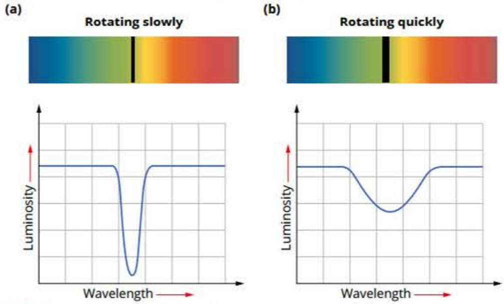


FIGURE 9.2.13 A star's rotation influences the width of its absorption bands. The faster the star spins, the wider the bands.

Density

The radius of a star can also be determined with the help of its spectrum. If the surface temperature of a star is known, a basic law of physics (known as the Stefan–Boltzmann law) can be used to determine the amount of energy given off each second by each unit area of the surface. Knowing this and the total energy given off by the star enables the total surface area and hence the radius r of the star to be calculated.

From observing the brightness and temperature of stars, astronomers can estimate a star's mass m. Assuming the star is approximately spherical, the volume V of the star can be calculated ($V = \frac{4}{3}\pi r^3$), and finally from its volume and mass the star's mean density can be calculated (density $= \frac{\text{mass}}{\text{volume}}$). The mean density of various stars varies greatly. Table 9.2.3 lists the density of some stars; the uncertainties in defining the precise edge of a ball of plasma means there is some variation in the calculated densities.

TABLE 9.2.3 Mean densities of some star types. The densities of air, water and the Earth are provided as useful comparisons.

Stars and other objects	Mean density ($kg\,m^{-3}$)
giant star (K-type)	~0.0001
supergiant star (O-type)	5
the Sun (G-type star)	1408 (overall) 150 000 (core)
white dwarf star	2.1×10^9
neutron star	1×10^{18}
interstellar medium	1×10^{-19}
air	1
water	1000
the Earth	5515 (overall) 13 000 (core)
supermassive black hole	9×10^5

+ ADDITIONAL

Age of a star

Observing many stars and interpreting the spectroscopic measurements has allowed astrophysicists to develop a life cycle for stars. By determining the types of atom present in the star and the star's luminosity, scientists can work out where the star is in its life span and so measure the star's age.

Deep in a nebula, gas—usually hydrogen and a little helium—falls together (collapses) under gravity to a rotating ball of plasma. When the core of the proto-star compresses to a sufficient density and temperature (around 10^7 K), the hydrogen begins to undergo nuclear fusion (a process discussed in more detail in Chapter 16) to form helium, which releases energy. This creates an outward pressure which balances the gravitational collapse. The star is now burning its hydrogen fuel at a steady rate. For a star with the mass of the Sun (known as 1 solar mass, or $1M_{\odot}$), this will take around 10 billion years; the Sun is currently around 5 billion years old. More-massive stars will burn the hydrogen more quickly, for as briefly as a few million years, while lighter stars burn the hydrogen more slowly. The consumption of hydrogen will take around 90% of the star's lifetime. During this time, the star gradually increases in luminosity. For example, the Sun is estimated to be 40% brighter than when it started fusing hydrogen. During this time the proportion of helium in the star increases. The process of building up heavier and more complex elements in the star is known as stellar nucleosynthesis. In the Sun, about 85% of the energy comes from hydrogen–helium fusion, and the remaining 15% from reactions that produce lithium and beryllium (elements 3 and 4).

When the hydrogen is largely consumed, what happens next depends on the star's mass.

- Low mass stars (less than $0.1\,M_{\odot}$) compress and collapse over several years as the hydrogen fuses to helium. These stars cannot fuse helium into heavier elements. Low mass stars live much longer than medium and high mass stars.
- Stars $0.5–1.5\,M_{\odot}$ become red giants. The helium core and hydrogen shell of the star expand. In the shell, hydrogen fuses to helium. In the core, helium can fuse to form elements including carbon, nitrogen, and oxygen. Heavier stars can form neon and magnesium. The shell increases and the star expands and cools. The star cannot compress the core much further, and heavier elements cannot form. Eventually the core will collapse to a white dwarf and the shell of the star will be ejected to form a nebula of metal-enriched gas around the white dwarf.
- For heavier stars, above $10M_{\odot}$, the nucleosysnthesis in the core proceeds further to form sulfur, silicon, aluminium and sodium. These stars are too heavy to limit the core collapse to a white dwarf. They will undergo an explosion—a nova—which will destroy the star and throw out metal-enriched gas.
- Still-heavier stars can undergo nucleosynthesis all the way to iron. Beyond iron, the nucleosynthetic fusion process requires energy instead of releasing it, and the internal pressure of the star cannot stop the core collapse. Heavier elements will be formed over a very short period of explosive nucleosynthesis as the dying star collapses, before the star ends its life in a supernova.

Using these ideas, the measurement of the star's luminosity, its temperature from its spectrum and the spectroscopic detection of elements present from the Fraunhofer lines can be combined to indicate where a star is in its life span and its likely fate.

Gases thrown out from a nova or supernova can be recycled elsewhere into a new star. The presence of carbon in your body and metals such as gold in the Earth show that our own Sun is at least a second-generation star.

Five billion years from now, the Sun will have bloated to a red giant (Figure 9.2.14). On Earth, the atmosphere will have burnt off and the oceans evaporated, leaving salt-encrusted rocks. The Sun, having exhausted the hydrogen in its core, will then begin burning hydrogen in a shell surrounding the core, causing the Sun to expand. The expanding Sun will engulf Mercury and Venus, and possibly the Earth. If the Earth survives this stage of the dying Sun, it will freeze when the Sun collapses into a white dwarf star.

FIGURE 9.2.14 Five billion years from now, the Sun will have expanded to a red giant.

9.2 Review

SUMMARY

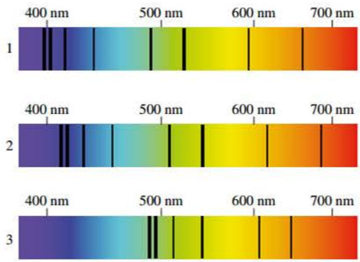
- The production of spectra suggests an internal structure to the atom. A line emission spectrum is produced by energised atoms, whereas an absorption spectrum is created when white light passes through a cold gas.
- The spectrum for any particular element is unique to that element.
- Interpreting the visible spectrum of stars reveals their temperature, density, chemical composition, and rotational and translational velocity.

KEY QUESTIONS

1 When does an element such as sodium produce an emission spectrum?

2 The predicted spectrum from an astronomical object is shown at the top of the figure below. The spectrum at the bottom of the figure is what astronomers actually observe. Explain what is occurring to cause this difference.

3 The following three stars are moving away from the observer. Order the stars in their velocity from fastest to slowest.



Chapter review

KEY TERMS

absorb
absorption spectrum
blueshift
discharge tube
emission spectrum
electromagnetic radiation
electromagnetic spectrum
excited state
ground state
incandescent
redshift
spectroscope
spectroscopy

REVIEW QUESTIONS

1 Although the currently accepted value for the speed of light is 299 792 458 $m s^{-1}$, this is often approximated as $c = 3.00 \times 10^8 m s^{-1}$. Calculate the percentage difference introduced by this approximation.

2 Which statement correctly describes how an electromagnetic wave is generated?

A A changing current creates a constant magnetic field which creates an opposing current.

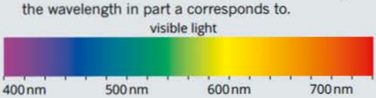
B A changing current creates a changing magnetic field which induces an electric field.

C A changing magnetic field creates an induced constant current which then creates an opposing magnetic field.

D A moving charge creates a constant current, which then creates a changing magnetic field.

3 **a** Calculate the wavelength (in nm) of light with a frequency of 5.0×10^{14} Hz.

b Use the figure below to identify what colour of light the wavelength in part a corresponds to.



4 Calculate the wavelength of a UHF (ultra-high frequency) television signal with a frequency of 7.0×10^7 Hz.

5 Calculate the frequency of an X-ray with a wavelength of 200 pm. (Note: $1 pm = 10^{-12} m$.)

6 Calculate the wavelength of the electromagnetic waves produced by a heavy-duty 915 MHz microwave oven.

7 Describe the relationship between the colours seen in the emission and absorption spectra of hydrogen.

8 Calculate the frequencies of the following wavelengths of light. Use $c = 3.00 \times 10^8 m s^{-1}$.

a blue, wavelength 486 nm

b violet, wavelength 397 nm

9 Why can't radio antennae be made of wood?

10 The absorption spectra of two stars are analysed to find the difference in their surface temperatures. Star A has a peak wavelength at 450 nm and star B has a peak wavelength at 400 nm. Which of the two stars is hotter? Calculate a ratio to describe how much hotter one star is than the other.

11 Why are ultraviolet and X-rays dangerous to life in a way that radio waves are not?

12 The James Webb Space Telescope is a space-based infrared telescope. Give two reasons for the greater validity of data from this telescope, compared to ground-based observations.

13 You conduct an experiment to measure the speed of light using a 2.45 GHz microwave, similar to the activity described in the Physics in Action on page 243. After measuring the distance between melted sections of marshmallows you produce the following data: 5.9 cm, 5.8 cm, 6.0 cm, 5.9 cm; and calculate a value for c. Does this agree with the theoretical value for the speed of light? Discuss.

14 Explain how emission spectra and absorption spectra are produced. Extension: Research how these spectra are produced in stars.

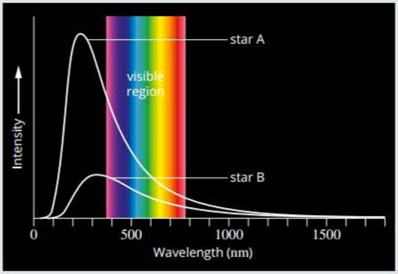
15 Arrange the following stars from coolest to hottest: yellow star, orange star, blue star, red star.

16 The emission spectrum of a distant star shows lines for hydrogen and helium but no other elements. What does that indicate about the age of the star?

17 How is light analysed in order to determine the standard properties for stars of Spectral Class A shown in the table below?

Spectral class	Surface temperature	Colour	Mass ($M_{\odot}$)
A	7500–10000 K	white	2.0–3.0

18 The following graph shows the wavelengths of light for two distant stars as a function of light intensity. Which of these has a greater temperature? Explain your answer.



19 After completing the activity on page 249, reflect on the inquiry question: What is light? You may also wish to do some further research regarding how the eye detects colour in order to fully understand your observations.

CHAPTER 10

Light: wave model

One of the great scientific achievements of the 19th century was the development of a comprehensive wave model for light. This model was able to explain a large number of wave properties including reflection, refraction, dispersion, diffraction, interference and polarisation. This also led to a deeper understanding of phenomena such as heat and radio transmissions.

This chapter follows the historical changes in the understanding of the nature of light. In particular, it explores diffraction, interference and polarisation, which are all properties of light that caused 19th-century scientists to abandon early ray and particle models for a wave model for light.

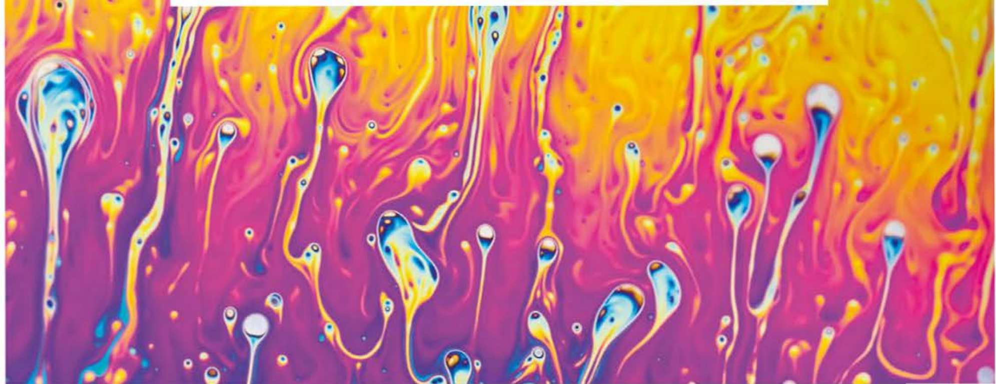
Content

INQUIRY QUESTION

What evidence supports the classical wave model of light and what predictions can be made using this model?

By the end of this chapter you will be able to:

- conduct investigations to analyse qualitatively the diffraction of light (ACSPH048, ACSPH076) ICT
- conduct investigations to analyse quantitatively the interference of light using double-slit apparatus and diffraction gratings ($d \sin\theta = m\lambda$) (ACSPH116, ACSPH117, ACSPH140) ICT N
- analyse the experimental evidence that supported the models of light that were proposed by Newton and Huygens (ACSPH050, ACSPH118, ACSPH123) CCT
- conduct investigations quantitatively using the relationship of Malus' law ($I = I_{max} \cos^2\theta$) for plane polarisation of light, to evaluate the significance of polarisation in developing a model for light (ACSPH050, ACSPH076, ACSPH120). ICT N



10.1 Diffraction and interference

In the late 17th century a debate raged among scientists about the nature of light.

The famous English scientist Sir Isaac Newton (Figure 10.1.1a) explained light in terms of particles or 'corpuscles', with each different colour of the spectrum representing a different type of particle. A Dutch scientist named Christiaan Huygens (Figure 10.1.1b) proposed an alternative model that described light as a type of wave, similar to the water waves observed in the ocean.

FIGURE 10.1.1 (a) Isaac Newton and (b) Christiaan Huygens were two scientists who disagreed about the fundamental nature of light.

A key point of difference between the two theories was that Newton's 'corpuscular' theory suggested that light would speed up as it travelled through a solid material such as glass. In comparison, the wave theory predicted that light would be slower in glass than in air.

Unfortunately, at that time it was impossible to measure the speed of light accurately, so the question could not be resolved scientifically. Newton's esteemed reputation meant that for many years his corpuscular theory was considered correct.

It was not until the early 19th century that experiments first convincingly demonstrated the wave properties of light.

Today, a modern understanding of light draws on aspects of both theories and is, perhaps, more complex than either Newton or Huygens could ever have imagined.

HUYGENS' PRINCIPLE

The theoretical basis for wave propagation in two dimensions was first explained by Christiaan Huygens. Huygens' principle states that each point on a wavefront can be considered a source of secondary wavelets (i.e. small waves).

Consider the plane wave shown in Figure 10.1.2. Each point on the initial wavefront can be treated as if it is a point source producing circular waves, some of which are shown in green. After one period, these circular waves will have advanced by a distance equal to one wavelength. Huygens proved mathematically that when the amplitudes of each of the individual circular waves are added, the result is another plane wave as shown by the new wavefront.

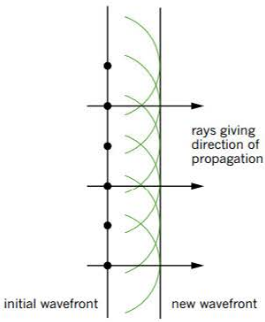


FIGURE 10.1.2 Each point on the wavefront of a plane wave can be considered a source of secondary wavelets. These wavelets combine to produce a new plane wavefront.

This process is repeated at the new wavefront, causing the wave to propagate in the direction shown.

Circular waves are propagated in a similar way, as shown in Figure 10.1.3.

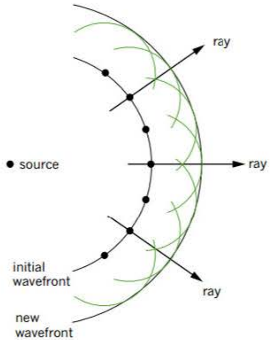
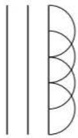


FIGURE 10.1.3 Each point on the wavefront of a circular wave can be considered a source of secondary wavelets. These wavelets combine to produce a new circular wavefront.

Worked example 10.1.1

APPLYING HUYGENS' PRINCIPLE

On the plane wave shown moving from left to right below, sketch some of the secondary wavelets on the wavefront and draw the appearance of the new wave formed after one period.

Thinking	Working
Sketch a number of secondary wavelets on the advancing wavefront.	
Sketch the new wavefront.	

Worked example: Try yourself 10.1.1

APPLYING HUYGENS' PRINCIPLE

On the circular waves shown below, sketch some of the secondary wavelets on the outer wavefront and draw the appearance of the new wave formed after one period.

Eventually scientists discovered that light has a number of properties that can only be explained using a wave model. Some of the most important of these are:

- diffraction
- interference
- polarisation.

These are described in the rest of this chapter.

DIFFRACTION

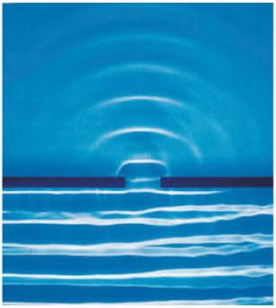


FIGURE 10.1.4 Water waves bend as they pass through a narrow opening.

FIGURE 10.1.5 The way information is printed on a CD or DVD means that it creates structures small enough to cause light to diffract, producing the colours observed in this image.

When a plane (straight) wave passes through a narrow opening, it bends. Waves will also bend as they travel around obstacles (Figure 10.1.4). This kind of 'bending' phenomenon is known as **diffraction**.

Diffraction is significant when the size of the opening or obstacle is similar to or smaller than the wavelength of the wave. Light waves range in wavelength from around 700 nm for red light to about 400 nm for violet light. 1 nm is equal to 10^{-9} m or a one millionth of a millimetre. Light waves are all less than one thousandth of a millimetre in length. This means that the diffraction of light is difficult to observe because the wavelength of light is so small, and there are not many situations in nature where light encounters structures of this size. However, diffraction does occur with artificially constructed materials like CDs (Figure 10.1.5) or commercially produced diffraction gratings.

Diffraction and slit width

In the diffraction of waves, if the wavelength, λ, is much smaller than the size of the gap or obstacle, w, the degree of diffraction is less. For example, Figure 10.1.6 shows the diffraction of water waves in a ripple tank. In Figure 10.1.6a, the gap is similar in size to the wavelength, so there is significant diffraction and the waves emerge as circular waves. In Figure 10.1.6b, the gap is much bigger than the wavelength, so diffraction only occurs at the edges.

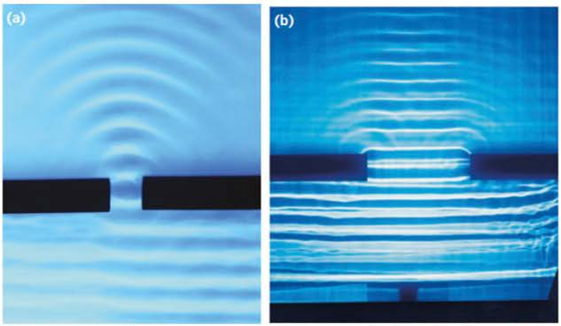


FIGURE 10.1.6 The diffraction of water waves in a ripple tank. (a) Significant diffraction occurs when the wavelength is approximately the same as the slit width, i.e. $\lambda \approx w$. (b) As the gap increases, diffraction becomes less obvious, since $\lambda \ll w$, but is still present.

Wavelengths comparable to or larger than the diameter of the obstacle or gap produce significant diffraction. This can be expressed as the ratio $\frac{\lambda}{w} \geq 1$, where λ is the wavelength of the wave and w is the width of the gap.

Diffraction gratings

As you have already seen, light diffracts as it passes through a very small gap. As the light passes through the gap, some of the wavelets making up the wavefront will diffract at the barriers that form the edges of the gap and some will pass through the centre of the gap. As a result of this, the light waves that emerge from the gap will interact, causing **interference**. In some places the interactions will be constructive (will add together) and in others the interactions will be destructive (will cancel out). When these light waves are made to shine on a screen, the areas of **constructive interference** will appear as bright bands and areas of **destructive interference** will appear as dark bands. The pattern of dark and light bands that is seen when light passes through a single small gap is called a **diffraction pattern**.

As stated earlier, the extent of diffraction of light waves is proportional to the ratio $\frac{\lambda}{w}$. This ratio also describes the spacing of dark and light bands in a diffraction pattern, and therefore the width of the overall diffraction pattern. According to this relationship, if the wavelength is held constant and the gap is made smaller, greater diffraction is seen. If different wavelengths enter the same gap, those with a smaller wavelength will undergo less diffraction than those with a longer wavelength. This is shown in Figure 10.1.7. Note that Figure 10.1.7 shows intensity. High intensity is where bright bands will appear on a screen; zero intensity corresponds to dark bands.

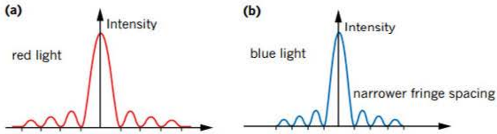


FIGURE 10.1.7 (a) Red light is diffracted more than (b) blue light. Red light's longer wavelength results in more-widely spaced fringes and a wider overall pattern.

Although some diffraction patterns can be observed using natural materials, in practice, much clearer diffraction patterns can be generated by passing light through a diffraction grating. A diffraction grating is a piece of material that contains a large number of very closely spaced parallel gaps or slits.

A diffraction grating can be thought of as a series of parallel slits all placed side by side. The diffraction pattern from one slit is superimposed on the pattern from the adjacent slit, producing a strong, clear image on the screen.

Diffraction experiments usually use only **monochromatic** light (i.e. light of only one colour). When white light, which contains a number of different colours, shines through a diffraction grating, each different colour is diffracted by a different amount and forms its own set of coloured fringes. This results in the light being dispersed into its component colours, as seen in Figure 10.1.8.

FIGURE 10.1.8 A diffraction grating disperses white light into a series of coloured spectra.

YOUNG'S DOUBLE-SLIT EXPERIMENT

Thomas Young's observation of interference patterns in light (Figure 10.1.9) was a pivotal moment in the history of science. It tipped the scales in the long-running dispute between scientists about the nature of light and paved the way for a series of discoveries and inventions that would fundamentally change scientists' understanding of energy and matter.

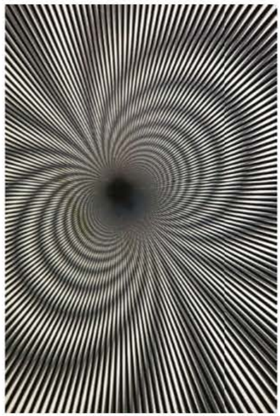


FIGURE 10.1.9 Optical interference can produce spectacular patterns.

PHYSICS IN ACTION WE CCT

Diffraction and imaging

Diffraction can be a problem for scientists using microscopes and telescopes because it can result in blurred images. For example, a significant problem is that the light from two tiny objects or two distant objects very close together can be diffracted so much that the two objects appear as one blurred object. When this happens, we say that the objects are unresolved. Essentially, the ratio $\frac{\lambda}{w}$ dictates how small an object can be clearly imaged by a particular instrument.

This means that, as a general rule, optical microscopes cannot create images of objects that are smaller than the wavelength of the light they use; otherwise, diffraction effects are too significant.

Diffraction also places a theoretical limit on the resolution of optical telescopes (Figure 10.1.10). However, for telescopes, atmospheric distortion usually has a much larger effect than diffraction. The Hubble Space Telescope, which sits above the Earth's atmosphere, is not affected by atmospheric distortion. It can resolve images right down to its diffraction limit, i.e. where the separation of the stars is approximately equal to the wavelength of the light.

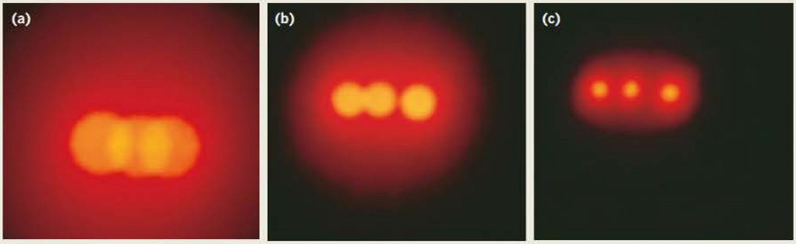


FIGURE 10.1.10 Photographs of three pinpoints of light taken with three lenses of different diameter. The lens used for (a) has the smallest diameter and therefore has the worst resolution, i.e. the points of light overlap. The lens used for (b) has a greater diameter than (a) and has better resolution. The best resolution image (c) is provided by the largest lens.

Young's experiment

In 1803, Young performed a now-famous experiment in which he shone monochromatic light (i.e. only one colour) on a screen containing two very tiny slits. On the far side of the double slits he placed another screen, on which he observed the pattern produced by the light passing through the slits (Figure 10.1.11).

According to the particle theory, light should have passed directly through the slits to produce two bright lines or bands on the screen (Figure 10.1.11a). Instead, Young observed a series of bright and dark bands or 'fringes' (Figure 10.1.11b).

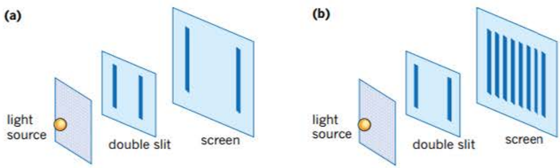


FIGURE 10.1.11 The particle theory of light predicted that Young's experiment should (a) produce two bright bands. But his actual experiment (b) produced a series of bright and dark bands or 'fringes'.

Young was able to explain this bright and dark pattern by treating light as a wave. He assumed that the monochromatic light was like plane waves and that, as they passed through the narrow slits, these plane waves were diffracted into **coherent** (in phase) circular waves as shown in Figure 10.1.12. The circular waves would interact causing interference. The interference pattern produced by these two waves would result in lines of constructive (antinodal) and destructive (nodal) interference that would match the bright and dark fringes respectively.

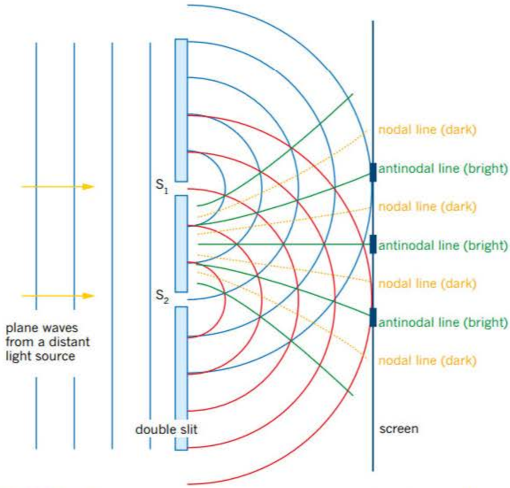


FIGURE 10.1.12 The interaction of two circular waves can produce a pattern of antinodal (constructive interference) and nodal (destructive interference) lines.

Earlier in his scientific career, Young had observed similar interference patterns in water waves, such as the ones shown in Figure 10.1.13. This gave greater credibility to the wave model for light proposed by Christiaan Huygens many years earlier.

When Young used the data to calculate the wavelength of light, it became clear why no one had ever noticed the wave properties of light before—light waves are tiny, with typical wavelengths of less than 1 micrometre (1 μm = 0.001 mm).

To understand Young's experiment more fully, you have to consider how the waves produced by the two slits interact with each other when they hit the screen. At a particular point, P, on the screen, the wave train from slit 1 (S_1) will have travelled a different distance compared with the wave train from slit 2 (S_2), i.e. the distance S_1P is different to S_2P. The difference in the distance travelled by each wave train to a point P on the screen is called the **path difference** for the waves (pd).

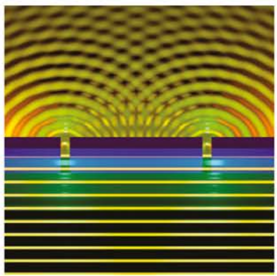


FIGURE 10.1.13 Interference patterns can be observed in water waves (lit here in yellow).

The path difference to point P from wave source S_1 and from wave source S_2 is given by:

$$pd = |S_1P - S_2P|$$

Path difference can be measured in metres, but it is far more useful to measure it in wavelengths in order to determine the light intensity on the screen.

As shown in Figure 10.1.14, at a point, M, at the centre of the screen, equidistant from each slit, each wave train will have travelled through the same distance, so there is no path difference (i.e. $S_1M = S_2M$). The light waves arrive in phase with each other. These light waves reinforce to produce an antinode. A fringe of bright light is seen, known as the 'central maximum'. This phenomenon is called constructive interference.

Constructive interference will occur whenever the path difference between the two wave trains is zero or differs by a whole number of wavelengths, i.e. pd = 0, 1λ, 2λ, 3λ... For example, in Figure 10.1.14, the path difference $S_1R - S_2R$ is equal to λ.

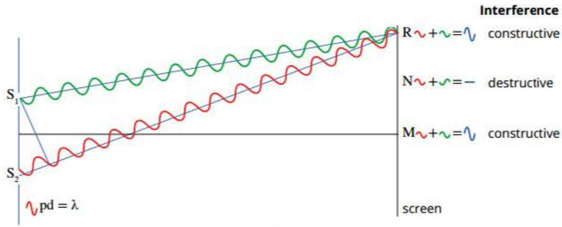


FIGURE 10.1.14 Waves meeting from each slit at R, where the path difference is λ. A bright fringe will be seen, as the wave trains arrive at this point in phase.

There will be points on the screen at which the path difference is $\frac{\lambda}{2}$; for example, point N in Figure 10.1.14. The two wave trains that meet at this point are completely out of phase and cancel each other to produce a nodal point. Destructive interference occurs at this point, and no light is seen. This creates the dark lines or fringes that appear in between the bright antinodal fringes. Destructive interference occurs when the path difference between the waves is $\frac{\lambda}{2}, \frac{3\lambda}{2}, \frac{5\lambda}{2}, \frac{7\lambda}{2}$...

In summary:

- constructive interference of coherent waves occurs when the path difference equals a whole number of wavelengths: pd = $m\lambda$, where m = 0, 1, 2, 3...
- destructive interference of coherent waves occurs when the path difference equals an odd number of half wavelengths: pd = $\left(m - \frac{1}{2}\right)\lambda$, where m = 1, 2, 3...

The sequence of constructive and destructive interference effects produces an interference pattern of regularly spaced vertical bands or fringes on the screen that can be represented graphically as shown in Figure 10.1.15.

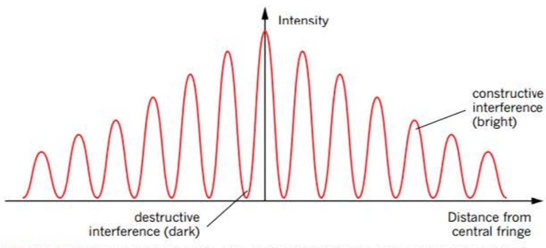


FIGURE 10.1.15 A graph of the intensity of a double-slit interference pattern. The horizontal axis represents a line drawn across the screen. The centre of the distribution pattern corresponds to the centre of the brightest central fringe, the central maximum.

Calculating fringe separation for Young's experiment

In Young's experiment, it is possible to predict where the bright fringes will occur by considering the geometry of the situation (Figure 10.1.16).

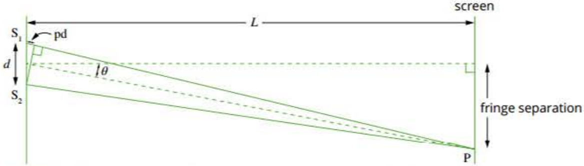


FIGURE 10.1.16 The geometry of the interference pattern from a double-slit experiment.

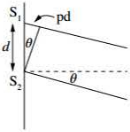


FIGURE 10.1.17 Assuming the distance to the screen is much greater than the slit separation, the rays will travel approximately parallel.

Consider two slits, S_1 and S_2, that are separated by a distance, d. If we draw a line at right angles halfway between the two slits directly across to the screen, we can then identify any point, P, on the screen by its angle, θ, from this line. When the distance to the screen (L) is much greater than the distance between the slits (d), then the exiting rays can be approximated to be travelling parallel (Figure 10.1.17). In this case, the angle θ is congruent to the internal angle of the smaller triangle such that the product $d \sin\theta$ is equal to the path difference between the two waves. Since a bright fringe will occur whenever the path difference is equal to a whole number of wavelengths (i.e. pd $= m\lambda$), then:

$$d \sin\theta = m\lambda$$

where d is the slit separation (m)

θ is the position to a bright point on the screen (as an angle from the perpendicular bisector between the slits)

m is any whole number, i.e. 0, 1, 2, 3...

λ is the wavelength of the light waves (m)

The path difference used for this formula is for constructive interference so that the angle calculated will be to a bright band. If instead you wanted to calculate the angle for a dark band, substitute the path difference for destructive interference, i.e. $d \sin\theta = (m - \frac{1}{2})\lambda$.

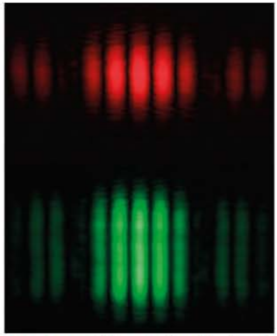


FIGURE 10.1.18 If the separation of the slits and the distance to the screen are kept the same, then the fringes produced by longer wavelength red light are further apart than those produced by shorter wavelength green light.

+ ADDITIONAL

Changing the parameters of Young's experiment

The angle θ is usually very small in Young's experiment. This means that scientists can use the approximation $\sin\theta \approx \theta \approx \tan\theta$ to simplify the Young's experiment formula to:

$$\Delta x = \frac{\lambda L}{d}$$

where Δx is the fringe separation (m)

λ is the wavelength of the light waves (m)

L is the distance from the slits to the screen (m)

d is the slit separation (m)

An advantage of this simplified formula is that it reveals some important relationships between the experimental parameters in Young's experiment:

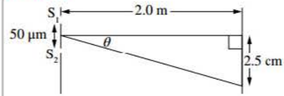
- $\Delta x \propto \lambda$, i.e. increasing the wavelength of light increases the fringe spacing (Figure 10.1.18).
- $\Delta x \propto L$, i.e. increasing the distance from the slits to the screen increases the fringe spacing.
- $\Delta x \propto \frac{1}{d}$, i.e. decreasing the distance between the slits increases the fringe spacing.

One of the challenges that scientists face when performing Young's experiment in the laboratory is that it is hard to measure the distance between the fringes because they are very close together. These relationships show how the apparatus can be adjusted to increase this distance—for example, the screen can be moved further away from the slits.

Worked example 10.1.2

CALCULATING WAVELENGTH FROM FRINGE SEPARATION

Light of an unknown wavelength emitted by a laser is directed through a pair of thin slits that are 50 μm apart. The slits are 2.0 m from a screen on which bright fringes are 2.5 cm apart. Calculate the wavelength of the laser light in nm.

Thinking	Working
Determine the angle θ.	 Since the screen is 2.0 m away and the distance between the first two bright fringes is 2.5 cm = 0.025 m: $\tan\theta = \frac{0.025}{2.0}$ $\theta = \tan^{-1}\left(\frac{0.025}{2.0}\right) = 0.72°$
Determine m.	The path difference between the rays creating the first fringe from the centre is 1λ. Therefore, $m = 1$.
Recall the equation for fringe separation.	$d\sin\theta = m\lambda$
Transpose the equation to make λ the subject.	$\lambda = \frac{d\sin\theta}{m}$
Substitute values into the equation and solve. (Note: 1 μm = 1×10^{-6} m.)	$\lambda = \frac{50\times10^{-6}\times\sin 0.72}{1} = 6.3\times10^{-7}$ m
Express your answer using the units specified—in this case nm, where 1 nm = 1×10^{-9} m.	The wavelength of the laser light is 630 nm.

Worked example: Try yourself 10.1.2

CALCULATING WAVELENGTH FROM FRINGE SEPARATION

Green laser light is directed through a pair of thin slits that are 25 μm apart. The slits are 1.5 m from a screen on which bright fringes are 3.3 cm apart. Calculate the wavelength of the green laser light in nm.

RESISTANCE TO THE WAVE MODEL

Young's wave explanation for his experiment was not immediately accepted by the scientific community. Many scientists were reluctant to abandon the corpuscular theory that had been accepted for over a century.

In 1818, the French scientist Augustin-Jean Fresnel was able to provide a mathematical explanation for Young's double-slit experiment based on Huygens' principle.

Another French scientist, Siméon Poisson, who was a passionate supporter of Newton's corpuscular theory, argued that if the same mathematics was applied to the light shining around a round disk, then there should be a bright spot in the middle of the shadow created by the disk (Figure 10.1.19). Since nobody had ever observed a bright spot in the middle of a shadow, Poisson believed this proved that the wave model was incorrect.

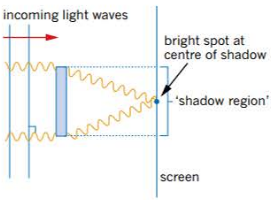


FIGURE 10.1.19 Waves of light incident on a solid disk diffract to give a point of light in the centre of the shadow zone. This is convincing evidence for the wave nature of light.

However, one of Poisson's colleagues decided to test these ideas by performing an experiment with a very small bright light source and a round disk, and observed the bright spot predicted by Poisson's calculations (Figure 10.1.20). As a consequence, for the remainder of the 19th century, the wave theory became the almost universally accepted model for light.

PA 7.3 PA 7.4 PA 7.4

This now-famous diffraction pattern has come to be known as the 'Poisson bright spot', which means it is named after the person who predicted that it would not exist!

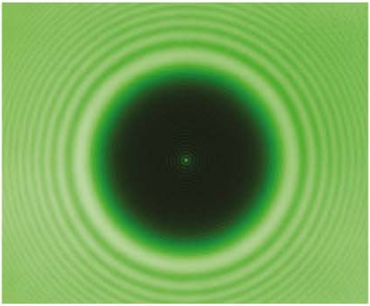


FIGURE 10.1.20 The bright spot inside the shadow region of this image is caused by the diffraction and interference of light waves. The image also shows diffraction and interference patterns surrounding the shadow.

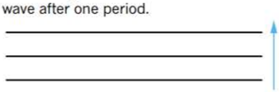
10.1 Review

SUMMARY

- A wave model explains a wide range of light-related phenomena, including diffraction and interference.
- When a plane (straight) wave passes through a narrow opening or meets a sharp object, it experiences diffraction.
- Significant diffraction occurs when the wavelength of the wave is similar to, or larger than, the size of the diffracting object.
- Young's double-slit interference experiment provided evidence to support the wave model of light.
- Data from Young's experiment can be analysed using the formula $d \sin \theta = m\lambda$.

KEY QUESTIONS

1 On the diagram below, draw the wavefront of the plane wave after one period.



2 For a diffraction experiment to produce significant diffraction of red light (wavelength of approximately 700 nm), what width opening should be used?

A 1 mm
B 0.1 mm
C 0.01 mm
D 0.001 mm

3 If Thomas Young's double-slit experiment was modelled using circular water waves in a ripple tank, which of the following events would correspond to nodal lines? (More than one correct answer is possible.)

A crests meet troughs
B troughs meet troughs
C crests meet crests
D troughs meet crests

4 A blue laser is directed through a pair of thin slits that are 40 μm apart. The slits are 3.25 m from a screen on which bright fringes are 3.7 cm apart. Calculate the wavelength of the blue laser light in nm.

10.2 Polarisation

PHYSICS INQUIRY N CCT

Interference art

What evidence supports the classical wave model of light and what predictions can be made using this model?

COLLECT THIS...

- overhead projector film
- cellophane tape
- two polarisation films
- scissors

DO THIS...

1 Place pieces of tape on the overhead film in different orientations, overlapping each other.
2 Place the overhead film between two pieces of polarising film and look through the combination at a bright light.
3 Rotate one of the polarising films, leaving the other two layers still.
4 Use the effect you have just seen to create some artwork. Try cutting different shapes and placing them in set orientations.

RECORD THIS...

Describe how rotating the polarising film changes the amount of light that comes through it.

Present your artwork.

REFLECT ON THIS...

What model of light predicts the behaviour of light travelling through polarising films?

Where have you seen this phenomenon before?

What could you do to improve your artwork?

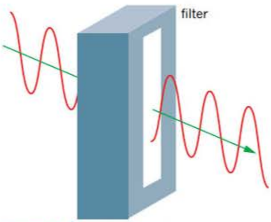


FIGURE 10.2.1 A vertically polarised wave can pass through a vertically orientated polarising filter.

One of the most convincing pieces of evidence for the wave nature of light is the phenomenon of **polarisation**. Polarisation occurs when a transverse wave is allowed to vibrate in only one direction. For example, the light wave in Figure 10.2.1 is vertically polarised—the wave oscillations occur in the vertical plane only. This also means this wave is unaffected by a polarising filter that is orientated in the vertical plane.

The wave in the Figure 10.2.2 is horizontally polarised. It is completely blocked by the vertical polarising filter.

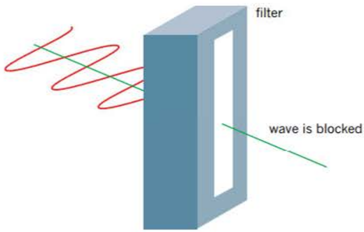


FIGURE 10.2.2 A horizontally polarised wave cannot pass through a vertically orientated polarising filter.

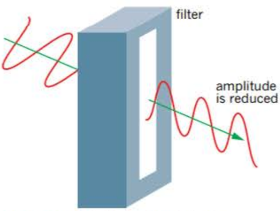


FIGURE 10.2.3 A diagonally polarised wave has its horizontal component blocked by the vertically orientated polarising filter. A vertically polarised wave of reduced amplitude passes through it.

In Figure 10.2.3, the incoming wave is polarised at 45° to the horizontal and vertical planes. The horizontal component of this wave is blocked by the vertical filter, so the ongoing wave is vertically polarised and has a smaller amplitude than the original wave.

Light produced by sources such as a light globe or the Sun is unpolarised, which means that it can be thought of as a collection of waves, each with a different plane of polarisation, as shown in Figure 10.2.4.

Certain materials can act as polarising filters for light. These only transmit the waves or components of waves that are polarised in a particular direction and absorb the rest. Polarising sunglasses work by absorbing the light polarised in a particular direction, thus reducing glare. Photographers also use polarised filters to reduce the glare in photographs or to achieve specific effects (Figure 10.2.5).

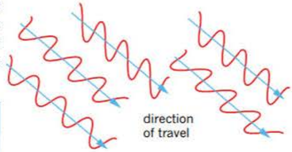


FIGURE 10.2.4 Unpolarised light consists of a collection of waves that are each polarised in a different direction.

FIGURE 10.2.5 These are photographs taken of the same tree, one without a polarising filter (left) and one with a polarising filter (right).

PHYSICSFILE N

Polarised sunglasses

Light that is reflected from the surface of water or snow is partially polarised (Figure 10.2.6). The polarising plane of polarised sunglasses is selected to absorb this reflected light. This makes polarised sunglasses particularly effective for people involved in outdoor activities such as boating, fishing or skiing.

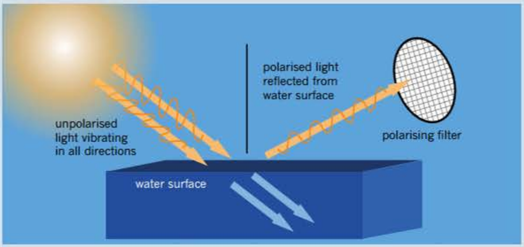


FIGURE 10.2.6 The polarising filter in a pair of sunglasses is designed to block the polarised light reflected from the surface of the water.

MALUS' LAW

The intensity of light is reduced as the light passes through a polarising filter; the amount of the reduction depends on the relative orientation of the filter and the plane of polarisation of the light. Mathematically, this relationship is described by an equation known as Malus' law:

$$I = I_{max} \cos^2 \theta$$

where

d is the slit separation (m)

I is the intensity of light passing through the filter (cd)

I_{max} is the intensity of light entering the filter (cd)

θ is the angle between the direction of polarisation of the light entering the filter, θ_0, and the axis of polarisation of the filter, θ_1 (Figure 10.2.7).

REVISION

Luminous intensity

In Module 3, you were introduced to the concept of luminous intensity, I, which is a measure of the rate at which light energy passes through an area of space. The SI unit of luminous intensity is the candela (cd), based on a candle of standard brightness. You saw in Module 3 that luminous intensity reduces with distance (r) using the relationship $I \propto \frac{1}{r^2}$. Recall that intensity is a measure of the amount of power transferred over a given area.

GO TO ➤ Year 11 Section 10.1

SKILLBUILDER N

EXPRESSING RELATIONSHIPS USING RATIOS

A ratio is the relationship between two numbers of the same kind. It could be the quantities in a recipe, the division of profits from a sale or the number of different types of the same thing.

Scientists use ratios to compare quantities, like the intensities of a ray of light before and after it passes through a polarising filter.

Ratios can be expressed in a number of different ways. For example, if a ray of light passes through a polarising filter and has its intensity reduced by 50%, we can express the relationship between the new intensity (I) and the original intensity (I_{max}) as:

$$I : I_{max} = 1 : 2$$
$$I = 0.5\,I_{max}$$
$$I = \tfrac{1}{2}I_{max}$$
$$2I = I_{max}$$
$$I = 50\% \times I_{max}$$

All of these representations are equivalent to each other. Often the context of a problem will make it clear which representation should be used.

WS 7.4

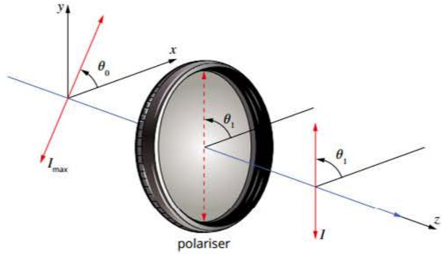


FIGURE 10.2.7 In Malus' law, θ is the angle between the direction of polarisation of the light entering the filter, θ_0, and the axis of polarisation of the filter, θ_1, i.e. $\theta = \theta_1 - \theta_0$.

If the axis of polarisation of the light is the same as that of the filter, then $\theta = 0°$. Using $I = I_{max}\cos^2\theta$, as $\cos 0° = 1$, then $I = I_{max}$, which means that the luminous intensity is not reduced at all in this case.

Conversely, if the axis of polarisation of the light is at right angles to the plane of polarisation of the filter, then $\theta = 90°$. As $\cos 90° = 0$, then $I = 0$, which means that no light can get through the filter.

Worked example 10.2.1

APPLYING MALUS' LAW TO CALCULATE RELATIVE REDUCTION IN INTENSITY

How much (as a percentage) is the intensity of a ray of light reduced if it passes through a polarising filter that is aligned at 45° to the plane of polarisation of the light?

Thinking	Working
Recall Malus' law.	$I = I_{max}\cos^2\theta$
Substitute the angle between the planes of polarisation of the filter and the light.	$I = I_{max}\cos^2 45$ $= 0.5\,I_{max}$
Express the answer as a percentage.	The intensity of the light has been reduced by 50%.

Worked example: Try yourself 10.2.1

APPLYING MALUS' LAW TO CALCULATE RELATIVE REDUCTION IN INTENSITY

How much (as a percentage) is the intensity of a ray of light reduced if it passes through a polarising filter that is aligned at 30° to the plane of polarisation of the light?

Worked example 10.2.2

APPLYING MALUS' LAW TO CALCULATE CHANGE IN INTENSITY

Vertically polarised laser light with an intensity of 50 cd passes through a polarising filter that is orientated at 25° to the vertical plane. Calculate the intensity of the light as it leaves the filter.

Thinking	Working
Recall Malus' law.	$I = I_{max}\cos^2\theta$
Substitute the values into the equation.	$I = 50\cos^2 25$ $= 41$ cd

Worked example: Try yourself 10.2.2

APPLYING MALUS' LAW TO CALCULATE CHANGE IN INTENSITY

PA 7.5 PA 7.5

Horizontally polarised laser light with an intensity of 90 cd passes through a polarising filter that is orientated at 60° to the horizontal plane. Calculate the intensity of the light as it leaves the filter.

PHYSICSFILE

Étienne Malus

Malus' law is named after Étienne Malus (Figure 10.2.8), who was a French physicist and mathematician. After serving as an engineer in Napoleon's army at the end of the 18th century, he made a number of important discoveries about the polarisation of light. He is one of only 72 people to have their names inscribed on the Eiffel tower, along with other famous French physicists such as André-Marie Ampère and Charles-Augustin de Coulomb.

FIGURE 10.2.8 Étienne Malus was a French engineer and scientist famous for his work on polarisation.

10.2 Review

SUMMARY

- A transverse wave model for light is required to explain polarisation.
- Polarisation occurs when a transverse wave is allowed to vibrate in only one direction.
- Malus' law can be used to calculate the change in the intensity of light as it passes through a polarising filter: $I = I_{max} \cos^2 \theta$.

KEY QUESTIONS

1 Explain how polarisation supports a wave model for light.

2 A ray of light passes through a polarising filter. The filter is aligned at 20° to the plane of polarisation of the light. By what percentage is the intensity of light reduced?

3 A ray of light of intensity 30 cd passes through a polarising filter which is aligned at 35° to the plane of polarisation of the light. What is the intensity of the light as it leaves the filter?

4 If the intensity of a ray of light decreases by 10% as it passes through a polarising filter, what is the angle between the planes of polarisation of the light and the filter?

Chapter review

KEY TERMS

coherent
constructive interference
destructive interference
diffraction
diffraction pattern
interference
monochromatic
path difference
polarisation

REVIEW QUESTIONS

1 Name the model of light each of the following scientists supported.
   a Young
   b Huygens
   c Newton

2 In the 18th century, why did most scientists support Newton's particle model?
   A Newton had better evidence to support his theory.
   B The speed of light in glass had been shown to be faster than in air.
   C Newton had a better reputation as a scientist than Huygens.
   D Newton was English and Huygens was from Holland.

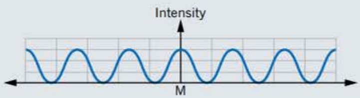
3 What phenomenon does the diagram below demonstrate?
   A diffraction
   B interference
   C reflection
   D refraction

4 Explain how the width of a double-slit interference pattern would change if all the variables were kept constant but a blue laser was replaced with a green laser.

5 Polarisation is an important phenomenon. What does it show about light?
   A It can travel instantaneously at an infinite speed.
   B It travels faster in materials like water and air than in a vacuum.
   C It is a longitudinal wave.
   D It is a transverse wave.

6 According to the particle model of light, Young's double-slit experiment should have produced two bright lines on the screen. Instead, what was observed on the screen?
   A It was completely dark.
   B It was completely light.
   C It contained three bright lines.
   D It contained a pattern of alternating bright and dark lines.

7 The following diagram shows the resulting (simplified) intensity pattern after light from two slits reaches the screen in a Young's double-slit experiment. Copy the diagram into your workbook and circle the points at which the path difference is equal to 1λ.



8 Explain why Young's double-slit experiment led to a significant change in scientists' understanding of the nature of light.

9 Light of an unknown wavelength emitted by a laser is directed through a pair of thin slits separated by 75 µm. The slits are 4.0 m from a screen on which bright fringes are 3.1 cm apart.
   a Calculate the angle between the first two bright fringes on the screen.
   b Calculate the wavelength of the light (in nm).

10 Red light with a wavelength of 650 nm is directed through a pair of thin slits separated by 80 μm onto a screen 3.5 m away. What is the angle between the central bright fringe and the third bright fringe to its right?

11 Blue light with a wavelength of 425 nm is used in a version of Young's experiment. The angle between the central bright fringe and the fifth bright fringe on the screen is 2.0°. Calculate the distance between the slits (in μm).

12 Describe Young's experiment and explain why it is considered evidence for the wave theory of light.

13 A version of Young's double-slit experiment is set up by directing the light from a red laser through a pair of thin slits. An interference pattern appears on the screen behind the slits. The following changes are made to the apparatus. Identify whether the distance between the interference fringes seen on the screen would increase, decrease or stay the same if:

a the screen is moved further away from the slits

b the slits are moved closer together.

14 Explain briefly why snowboarders and sailors are likely to wear polarising sunglasses.

15 A ray of light passes through a polarising filter that is aligned at 50° to the plane of polarisation of the light. Express the intensity of the light passing through the filter (I_2) as a percentage of the initial intensity (I_1) of the light.

16 A 15 cd beam of light passes through a polarising filter that is aligned at 75° to the plane of polarisation of the light. What is the intensity of the light as it leaves the filter?

17 What should the angle between the planes of polarisation of light and a filter be to reduce the intensity of the light by 80%?

18 Diffraction of light waves was not observed by scientists until the 19th century. Why was this phenomenon not observed earlier?

19 A monochromatic light source with a wavelength of 450 nm is shone through a pair of slits that are 0.1 mm apart, creating an interference pattern on a screen 2.0 m away. At what angle do you expect to see the third dark band?

20 After completing the activity on page 274, reflect on the inquiry question: What evidence supports the classical wave model of light and what predictions can be made using this model?

CHAPTER 11

Light: quantum model

At the end of the 19th century, most scientists had accepted that the wave model of light had replaced Newton's earlier corpuscular (particle) model. The wave model explained a number of phenomena (such as diffraction, interference and polarisation) that could not be explained using the particle model. Some scientists believed that all that was left to be done in the field of optics was to measure some constants, such as the speed of light, as accurately as possible. Few scientists expected what was to come—that the wave model's inability to explain a couple of key observations would lead to a fundamental revolution in humanity's understanding of both light and matter.

This chapter will outline the key experiments that led to the development of a 'new' particle model of light, known as the quantum model.

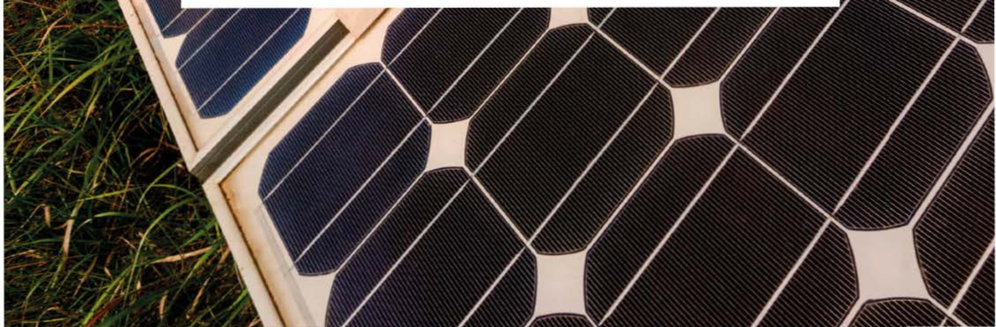
Content

INQUIRY QUESTION

What evidence supports the particle model of light and what are the implications of this evidence for the development of the quantum model of light?

By the end of this chapter you will be able to:

- analyse the experimental evidence gathered about black-body radiation, including Wien's law related to Planck's contribution to a changed model of light (ACSPH137) CCT ICT N
  - $\lambda_{max} = \frac{b}{T}$
- investigate the evidence from photoelectric effect investigations that demonstrated inconsistency with the wave model for light (ACSPH087, ACSPH123, ACSPH137) CCT ICT
- analyse the photoelectric effect ($K_{max} = hf - \phi$) as it occurs in metallic elements by applying the law of conservation of energy and the photon model of light (ACSPH119). ICT N



11.1 Black-body radiation

GO TO ➤ Section 9.1 page 240

You may recall from Chapter 9 that the term **electromagnetic radiation** includes such things as:

- visible light
- ultraviolet light
- infrared radiation.

All forms of electromagnetic radiation are essentially the same, differing only in their frequency and, therefore, their wavelength. The full range of electromagnetic radiation is called the electromagnetic spectrum. The electromagnetic spectrum is divided into different categories according to the wavelength and frequency of the radiation (Figure 11.1.1).

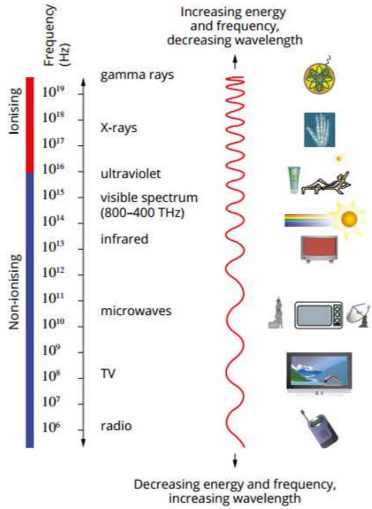


FIGURE 11.1.1 The electromagnetic spectrum showing the wavelength of the electromagnetic radiation and the corresponding frequencies (Hz). The spectrum is divided into categories based on wavelength and radiation energy levels. Visible light is a mid-frequency range of wavelengths that can be seen by the human eye.

The seven categories of electromagnetic radiation are grouped according to similar characteristics. Table 11.1.1 shows these groups from the highest energy to the lowest. There is some overlap between these groups, and they are divided largely based on application rather than distinct boundaries or values.

Electromagnetic radiation is emitted by all objects whose temperature is above absolute zero (0 K or −273°C). The actual wavelength or frequency of the emitted radiation depends almost entirely on the internal energy of the object and not on the characteristics of the material itself.

TABLE 11.1.1 Types of electromagnetic radiation.

Type of electromagnetic radiation	Description
gamma rays (γ-rays)	The highest energy, shortest wavelength energy is produced within the nucleus of an atom. Gamma rays are one of the three types of emissions that come from radioactive (unstable) atoms.
X-rays	When fast-moving electrons are fired into an atom, X-rays are produced. X-rays got their name because scientists, at first, did not know what they were, hence the letter 'X'.
ultraviolet (UV)	UV light waves have shorter wavelengths than visible violet light, but longer wavelengths than gamma rays or X-rays. UV light is known to cause skin cancer, particularly with frequent exposure.
visible light	This is the small band of wavelengths around the middle of the electromagnetic spectrum that can be detected by human eyes. Many other life forms, for example insects and birds, can perceive wavelengths well into the ultraviolet range.
infrared (IR)	Infrared or heat radiation is emitted by all objects that are not at a temperature of absolute zero. The hotter the object, the more radiation that is emitted and the shorter its wavelength within the infrared band.
microwaves	The microwaves that cook your dinner and allow remote communications and radar to work are produced by the spin of electrons or nuclei. Wavelengths range from about 1 nm to 10 cm.
radio and television waves	Electrons oscillating in a conducting wire, such as an antenna, produce the radio and television waves that bring music and pictures to your home and carry voice and data to your phone. Long-wavelength electromagnetic radio and television waves can be transmitted across very long distances.

The higher the temperature of an object, the higher the frequency and the shorter the wavelength of the emitted radiation. As temperature increases, electromagnetic radiation is emitted at increasingly higher frequencies. Consider the following examples.

- Relatively cool objects, such as the human body, emit infrared radiation. Infrared radiation is not visible to the human eye under normal circumstances.
- At higher temperatures, objects emit radiation with a higher frequency and you can see them glow red. An example is a bar heater that glows red hot.
- At even higher temperatures, for example 2000 K, objects such as the filament of an incandescent light glow yellow or white.
- Very hot objects, at temperatures of 10^6 K or more, emit most of their radiation within the gamma and X-ray regions of the electromagnetic spectrum.

PHYSICSFILE N

Scientific notation

Very small and very large numbers, such as the wavelengths and frequencies of electromagnetic radiation, can be tedious to write out in full. Scientific notation, which uses powers of 10, is a useful way of showing numbers in a more concise way. Table 11.1.2 lists some commonly used powers of 10 factors, prefixes and symbols.

TABLE 11.1.2 Common prefixes, symbols and factors for large and small numbers.

Factor	Prefix	Symbol	Factor	Prefix	Symbol
10^{12}	tera	T	10^{-3}	milli	m
10^9	giga	G	10^{-6}	micro	µ
10^6	mega	M	10^{-9}	nano	n
10^3	kilo	K	10^{-12}	pico	p

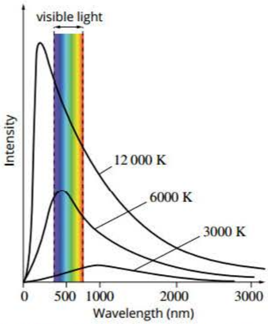


FIGURE 11.1.2 The spectrum of wavelengths emitted for an idealised object (a black body) at different temperatures. The radiation approximates the surface temperature of many real objects.

WIEN'S LAW

Wilhelm Wien, a German physicist, formulated laws that describe the properties of heat radiation. Wien discovered that the peak wavelength at which an object will emit the maximum intensity of radiation is dependent on its surface temperature. Wien's displacement law, more commonly known just as Wien's law, can be used to determine the peak wavelength for an object at a particular surface temperature:

$$\lambda_{max} = \frac{b}{T}$$

where

λ_{max} is the peak wavelength of the emitted radiation (m)

b is a constant equal to 2.898×10^{-3} m K

T is the surface temperature of the object (K)

According to this formula, no matter what the surface temperature of an object, the product of the temperature and the wavelength at which the peak intensity of the emitted radiation occurs is a constant and is equal to 2.898×10^{-3} m K.

The graph in Figure 11.1.2 shows the continuous spectra emitted by an object (e.g. a solid, liquid or even a dense gas) at particular temperatures. An object at a temperature of around 12 000 K will emit its peak wavelength in the ultraviolet range. That is, the wavelength corresponding to the highest intensity for the 12 000 K curve occurs in the ultraviolet range.

The 6000 K curve in the graph in Figure 11.1.2 corresponds to the surface temperature of our Sun. The maximum value of this intensity curve corresponds to a peak wavelength at about 500 nm that is within the visible band of the electromagnetic spectrum. The surface temperature of the Sun means that most of its electromagnetic radiation is emitted within the range between ultraviolet and infrared, including visible light.

PHYSICSFILE N

Our eyes and the Sun

It is no coincidence that the human eye is very good at detecting visible light. Human eyes have evolved to be most receptive to wavelengths of light within what is known as the visible range and which correspond to the highest intensity of light produced by our Sun. If the Sun had a lower surface temperature, of say around 3000 K, it's probable that human eyes would be adapted to the infrared range.

BLACK-BODY RADIATION

Wien's work on the wavelength of the radiation emitted by a hot, dense object was based initially on a theoretical object called a **black body**. A black body does not necessarily have to be black. An incandescent lamp or a light bulb may be, to a certain extent, regarded as a black body.

This theoretical object completely absorbs all the rays of electromagnetic radiation that fall on it regardless of the wavelength of the radiation. In other words, a black body does not reflect any radiation. The radiation emitted by many objects, such as the Sun, can be approximated as the radiation emitted by a black body at the same temperature. The spectrum emitted by a hot object is continuous but has a peak intensity at a wavelength that is inversely proportional to the surface temperature. This relationship is more simply stated by rearranging Wien's law:

$$\lambda_{max} \propto \frac{1}{T} \text{ and also } T \propto \frac{1}{\lambda_{max}}$$

Wien's law can be used to find the temperature of any hot object, including our Sun.

Worked example 11.1.1

THE TEMPERATURE AT A STAR'S SURFACE

The Sun emits a continuous electromagnetic spectrum with a peak wavelength of approximately 500 nm. Based on this wavelength, estimate the surface temperature of the Sun.

Thinking	Working
Express the peak wavelength in metres.	$\lambda_{max} = 500\,nm = 500 \times 10^{-9}\,m$
Rearrange Wien's law to solve for *T*.	$\lambda_{max} = \frac{b}{T}$ $T = \frac{2.898 \times 10^{-3}}{\lambda_{max}}$
Substitute the value for λ_{max} and solve for *T*.	$T = \frac{2.898 \times 10^{-3}}{500 \times 10^{-9}}$ $= 6000\,K$

Worked example: Try yourself 11.1.1

THE TEMPERATURE AT A STAR'S SURFACE

A newly discovered star is observed to emit radiation with a peak wavelength of approximately 90 nm. Based on this wavelength, estimate the surface temperature of this star.

PHYSICS IN ACTION ICT

Warm white or cool daylight

The simple act of replacing a light globe requires many considerations. The globe needs to be the right size and have the appropriate power rating. Further consideration needs to be given to the technology—a standard incandescent, a hot halogen, a low-energy fluorescent or an LED. Finally, there's the choice of 'warm white' or 'cool daylight'. Some of the different types of globe available are shown in Figure 11.1.4.

Warm white and cool daylight are descriptions used by manufacturers to describe the colour of the light from what is basically a white light globe. A warm-white globe provides a slightly yellow, 'warm' glow, while 'cool daylight' is a more harsh white light. This difference relates to the peak wavelength emitted by the globe. Manufacturers design globes to mimic surface temperatures corresponding to particular peak wavelengths. Warm white is usually labelled as corresponding to a surface temperature of 3000 K; cool daylight corresponds to 6500 K.

Computer manufacturers use similar colour profiles to allow computer users to correct the whiteness of a computer screen.

FIGURE 11.1.4 Standard, energy-saving and LED light bulbs come in various shades of white that relate to the surface temperature they mimic.

PHYSICSFILE L

Wilhelm Wien

Wilhelm Wien (1864–1928), shown in Figure 11.1.3, was a German physicist. Wien was awarded the 1911 Nobel Prize in Physics for making a significant contribution to the thermodynamics of radiation. In 1893 he had discovered the relationship now known as Wien's law, which paved the way for Planck's quantum theory of radiation. This is the theoretical basis of modern physics, which explains the nature and behaviour of matter and energy. So Wien's displacement law was a very significant discovery indeed!

FIGURE 11.1.3 German physicist Wilhelm Wien.

RE-RADIATED ELECTROMAGNETIC RADIATION

Radiant energy interacts with matter in three ways. It can be:

- reflected
- transmitted
- absorbed.

More often than not it will be a combination of two or more of these modes. For example, some of the radiant energy that is absorbed by the surface of the Earth causes the Earth to heat up, and the rest is re-radiated back out into space.

Each of the gases in the Earth's atmosphere absorbs a very narrow band of wavelengths of the incoming solar radiation depending upon the nature of the gas. The smaller molecules (i.e. oxygen and nitrogen) absorb very short wavelengths of solar radiation. The larger molecules (i.e. water vapour and carbon dioxide) absorb primarily longer infrared radiant energy. About 17% of the radiant energy from the Sun is absorbed by the atmosphere, leading to the heating of the upper layers of the atmosphere.

The lower layers of the atmosphere do not absorb much heat directly from the Sun. They are predominately heated by the radiation from the Earth. The Earth is much cooler than the Sun, and so emits much longer wavelength radiation. With the temperature of the Sun's surface being approximately 6000 K, the peak wavelength of solar radiation is around 500 nm. This corresponds to the visible part of the electromagnetic spectrum. The Earth has an average temperature of around 16°C or 289 K. At that significantly lower temperature, the Earth emits most of its energy in the infrared range of the electromagnetic spectrum. This can be seen in the infrared image shown in Figure 11.1.5.

Using Wien's law, the peak wavelength of re-radiated energy from the Earth can be calculated, as you can see in the Worked example below.

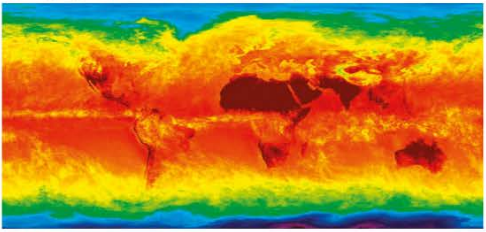


FIGURE 11.1.5 The Atmospheric Infrared Sounder (AIRS) instrument aboard NASA's Aqua satellite senses temperature using infrared wavelengths. This image shows the temperature of the Earth's surface or clouds covering it for the month of April 2003.

Worked example 11.1.2

RE-RADIATED ENERGY FROM THE EARTH

The Earth's average surface temperature is 289 K. What is the peak wavelength of the re-radiated electromagnetic radiation?

Thinking	Working
State Wien's law.	$\lambda_{max} = \frac{b}{T}$
Substitute the values for *b* and *T* and solve for λ_{max}.	$\lambda_{max} = \frac{2.898 \times 10^{-3}}{289}$ $= 1.00 \times 10^{-5}\ m = 10\ \mu m$

Worked example: Try yourself 11.1.2

RE-RADIATED ENERGY FROM THE EARTH

The Earth's average surface temperature at the equator is 300 K. What is the peak wavelength of the re-radiated electromagnetic radiation from this portion of the Earth?

PLANCK'S EQUATION

At the turn of the 20th century, a number of scientists turned their attention to light phenomena that could not be readily explained using Maxwell's electromagnetic wave model. The study of these phenomena required the development of much more sophisticated models for light, and eventually led to a revolution in the scientific understanding of the nature of energy and matter.

In 1900, the German physicist Max Planck (1958–1947), shown in Figure 11.1.6, was studying the spectrum for light emitted by hot objects. Planck and other scientists had discovered that certain features of this spectrum could not be explained using a wave model for light.

Planck proposed a controversial solution to this problem by assuming that light was emitted as discrete packets. He called the discrete packets of energy 'quanta', and developed an equation for the energy, E, of each **quantum**:

$E = hf$

where

E is the energy of a quantum of light (J)

f is the frequency of the electromagnetic radiation (Hz)

h is the constant 6.626×10^{-34} J s, now known as Planck's constant

FIGURE 11.1.6 Max Planck.

Since electromagnetic radiation is more commonly described according to its wavelength, scientists often combine Planck's equation with the wave equation for light, $c = f\lambda$, as follows:

$$E = hf \text{ and } f = \frac{c}{\lambda}$$

So

$$E = \frac{hc}{\lambda}$$

At the time, most scientists disregarded Planck's work because the particle model it suggested was so much at odds with the wave model that had become widely accepted as the correct explanation for light.

Worked example 11.1.3

USING PLANCK'S EQUATION

Calculate the energy in joules of a quantum of ultraviolet light that has a frequency of 2.00×10^{15} Hz.

Thinking	Working
Recall Planck's equation.	$E = hf$
Substitute in the appropriate values to solve.	$E = 6.626 \times 10^{-34} \times 2.00 \times 10^{15}$ $= 1.33 \times 10^{-18}$ J

PHYSICSFILE L

Max Karl Ernst Ludwig Planck

Max Planck was a German physicist. At the age of 21 he obtained a PhD in physics, and in 1889 was appointed professor at the university in Berlin. Planck was an author of numerous works about physics—about quantum theory in particular. On 14 December 1900 he presented a revised version of Wien's law and introduced a new constant that came to be known as Planck's constant. This date is now recognised as the beginning of the era of quantum mechanics. In 1918, Planck was awarded the Nobel Prize in Physics for 'the discovery of energy quanta'.

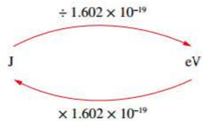
Worked example: Try yourself 11.1.3

USING PLANCK'S EQUATION

Calculate the energy in joules of a quantum of infrared radiation that has a frequency of 3.6×10^{14} Hz.

SKILLBUILDER N

The electron-volt



When studying light, the quantities of energy considered are usually very small. These are often so small that the joule is no longer a convenient unit to use. Scientists have therefore adopted a unit called an **electron-volt** (eV). An electron-volt is the amount of energy an electron gains when it moves through a potential difference of 1 V. Since the charge on an electron is -1.602×10^{-19} C, then:

$$1\,\text{eV} = q_e \times 1\,\text{V}$$
$$= 1.602 \times 10^{-19}\,\text{C} \times 1\,\text{J C}^{-1}$$
$$= 1.602 \times 10^{-19}\,\text{J}$$

Here is a simple way to convert between the units for energy:

To convert a value expressed in J into eV, divide it by $1.602 \times 10^{-19}\,\text{J eV}^{-1}$.

To convert a value expressed in eV into J, multiply it by $1.602 \times 10^{-19}\,\text{J eV}^{-1}$.

Worked example 11.1.4

CONVERTING TO ELECTRON-VOLTS

A quantum of light has 1.33×10^{-18} J of energy. Convert this energy to electron-volts.

Thinking	Working
Recall the conversion for joules to electron-volts.	$1\,\text{eV} = 1.602 \times 10^{-19}\,\text{J}$
Divide the value expressed in joules by $1.602 \times 10^{-19}\,\text{J eV}^{-1}$ to convert it to electron-volts.	$\frac{1.33 \times 10^{-18}}{1.602 \times 10^{-19}} = 8.31\,\text{eV}$

Worked example: Try yourself 11.1.4

CONVERTING TO ELECTRON-VOLTS

A quantum of light has 2.4×10^{-19} J of energy. Convert this energy to electron-volts.

As seen from worked examples 11.1.3 and 11.1.4, it is easier to compare the relative energies of quanta when they are expressed in eV.

For convenience, Planck's constant can also be given in terms of electron-volts:

$$h = 6.63 \times 10^{-34}\,\text{J s}$$
$$= \frac{6.626 \times 10^{-34}}{1.602 \times 10^{-19}}$$
$$= 4.14 \times 10^{-15}\,\text{eV s}$$

Worked example 11.1.5

WS 7.6

CALCULATING QUANTUM ENERGIES IN ELECTRON-VOLTS

Calculate the energy (in eV) of a quantum of ultraviolet light that has a frequency of 2.0×10^{15} Hz. Use $h = 4.14 \times 10^{-15}$ eV s.

Thinking	Working
Recall Planck's equation.	$E = hf$
Substitute in the appropriate values and solve for E.	$E = 4.14 \times 10^{-15} \times 2.0 \times 10^{15}$ $= 8.3$ eV

Worked example: Try yourself 11.1.5

CALCULATING QUANTUM ENERGIES IN ELECTRON-VOLTS

Calculate the energy (in eV) of a quantum of infrared radiation that has a frequency of 3.6×10^{14} Hz. Use $h = 4.14 \times 10^{-15}$ eV s.

11.1 Review

SUMMARY

- The peak wavelength, at which an object will emit the maximum intensity of radiation, is dependent on the object's surface temperature and is given by Wien's law. Wien's law states that $\lambda_{max} = \frac{b}{T}$, where $b = 2.898 \times 10^{-3}$ m K.
- Planck assumed that, on the atomic level, electromagnetic radiation is emitted or absorbed in discrete packets called quanta.
- The energy of a quantum of light is proportional to its frequency: $E = hf = \frac{hc}{\lambda}$.
- The electron-volt is an alternative (non-SI) unit of energy: $1\text{ eV} = 1.602 \times 10^{-19}$ J.

KEY QUESTIONS

1 James Maxwell developed a theory of EMR based on oscillating fields. What was Max Planck's contribution to the theory of EMR, and what in Maxwell's theory did it modify?

2 If a star emits a continuous electromagnetic spectrum with a peak wavelength of approximately 800 nm, what is the surface temperature of the star?

3 The element of an electric heater is just seen to glow a dull red. This colour corresponds to the lower end of the visible spectrum at approximately 700 nm. What temperature, in kelvin, is the element of the heater?

4 If a star has a surface temperature of 9000 K, what is the peak wavelength of the energy it is emitting?

5 Calculate the energies (in joules and electron-volts) of the quanta of the following wavelengths of light.

	Colour	Wavelength (nm)
a	red	656
b	yellow	589
c	blue	486
d	violet	397

11.2 The photoelectric effect

PHYSICS INQUIRY CCT PSC

A quantum model

What evidence supports the particle model of light and what are the implications of this evidence for the development of the quantum model of light?

COLLECT THIS…

- poster paper
- marker pens and pencils to decorate your poster

DO THIS…

1. Choose one side of the historical debate for the nature of light: 'light is a wave' or 'light is a particle'.
2. Research the evidence that supports your side. Include:
   - a description of your model
   - behaviours that this model explains and predicts
   - famous historic scientists that supported your side.
3. Create a poster advertisement to convince other students that your side of the debate is the right one.

RECORD THIS…

Describe how the scientific community deals with conflicting evidence.

Present your advertisement.

REFLECT ON THIS…

What evidence supports the particle model of light and what are the implications of this evidence for the development of the quantum model of light?

At the start of the 20th century, another phenomenon that could not be explained using the wave model for light was being observed.

Scientists noticed that when some types of electromagnetic radiation are incident on a piece of metal, the metal becomes positively charged. This positive charge is due to electrons being ejected from the surface of the metal. The electrons became known as **photoelectrons** because they were released due to light or other forms of electromagnetic radiation. The phenomenon is known as the **photoelectric effect**.

OBSERVING THE PHOTOELECTRIC EFFECT

A common apparatus used to observe the photoelectric effect is shown in Figure 11.2.1. It consists of a clean metal surface (the cathode), illuminated with light from an external source. If the light causes photoelectrons to be emitted, they are detected at the anode. The flow of electrons is called the **photocurrent** and is registered by a sensitive ammeter.

A typical circuit used to investigate the photoelectric effect includes a variable voltage supply, which can be used to make the cathode negative (and the anode positive). When this is done, the resulting electric field helps the photoelectrons to cross the gap to the anode. This happens because the photoelectrons are repelled by the negative potential at the cathode and are attracted to the positive potential at the anode. As a result, a maximum possible current will be measured. Alternatively, the voltage may be adjusted to make the cathode positive and the anode negative. This repels the photoelectrons and slows them down. As the anode voltage is increased, the photoelectrons are repelled more and more until the photocurrent drops to zero.

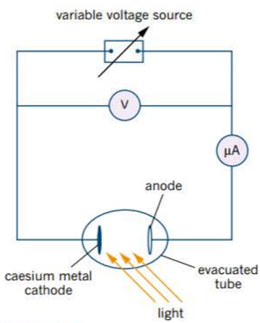


FIGURE 11.2.1 Circuit diagram of an experimental investigation of the photoelectric effect.

Using the apparatus shown in Figure 11.2.1, the German physicist Philipp Lenard made a number of surprising discoveries about the photoelectric effect. He won the Nobel Prize in Physics in 1905 for his discoveries.

Lenard used a filter to vary the frequency of the incident light. He discovered that, for a particular cathode metal, there is a certain frequency of light below which no photoelectrons are observed. This is called the **threshold frequency**, f_0. For frequencies of light greater than the threshold frequency (i.e. $f > f_0$), photoelectrons will be collected at the anode and registered as a photocurrent. For frequencies below the threshold frequency (i.e. $f < f_0$), no photoelectrons will be detected.

Lenard also discovered that, for light that has a frequency greater than the threshold frequency, i.e. $f > f_0$, the rate at which the photoelectrons are produced varies in proportion with the intensity (brightness) of the incident light as shown in Figure 11.2.2.

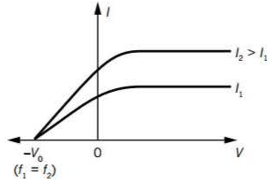


FIGURE 11.2.2 Photocurrent (I) plotted as a function of the voltage (V) applied between the cathode and the anode for different light intensities. For brighter light ($I_2 > I_1$) of the same frequency ($f_1 = f_2$), there is a higher photocurrent, but the same stopping voltage, V_0.

This graph shows a number of important properties of the photoelectric effect.

- When the light intensity increases, the photocurrent increases.
- When the applied voltage is positive, photoelectrons are attracted to the collector electrode (anode). A small positive voltage is enough to ensure that every available photoelectron is collected. The current therefore reaches a maximum value and remains there even if the voltage is increased.
- When the applied voltage is negative, photoelectrons are attracted back towards the illuminated cathode and are repelled by the collector electrode (anode), and the photocurrent is reduced. The photocurrent is reduced because fewer and fewer photoelectrons have the energy to overcome the opposing electric potential. There is a voltage, V_0, for which no photoelectrons reach the collector. This is known as the **stopping voltage**. For a particular frequency of light on a particular metal, this stopping voltage is a constant.

GO TO ➤ Section 5.1 page 149

Recall from earlier studies of electricity (Chapter 5) that the work done on a charge (by an applied voltage) is given by $W = qV$. In this case, the voltage used is designated the stopping voltage, V_0, and the charge value is equal to the magnitude of the charge on an electron, $q_e = 1.60 \times 10^{-19}$ C. Hence the work done on the electron is given by $W = q_e V_0$. Since the stopping voltage is large enough to stop even the fastest-moving electrons from reaching the anode, this expression gives the value of the maximum possible kinetic energy of the emitted photoelectrons. For example, should the stopping voltage be 2.5 V, then the maximum kinetic energy of any photoelectron is 2.5 eV.

When the light sources have the same intensity but different *frequencies*, they produce the same maximum current. However, the higher frequency light has a higher stopping voltage (Figure 11.2.3).

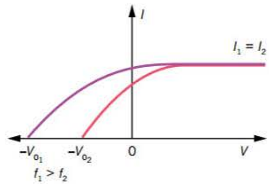


FIGURE 11.2.3 Photocurrent (I) plotted as a function of the voltage (V) applied between the cathode and the anode for different frequencies ($f_1 > f_2$) of incident light with the same intensity ($I_1 = I_2$). Both frequencies produce the same maximum photocurrent; however, light with the higher frequency requires a larger stopping voltage.

Finally, as long as the incident light has a frequency above the threshold frequency of the cathode material, photoelectrons are found to be emitted without any appreciable time delay. This fact holds true regardless of the intensity of the light.

When illuminated with light above the threshold frequency, some photoelectrons are emitted from the first layer of atoms at the surface of the metal and have the maximum kinetic energy possible. Other photoelectrons come from deeper inside the metal and lose some of their kinetic energy due to collisions on their way to the surface. Hence, the emitted photoelectrons have a range of kinetic energies from the maximum value K_{max} downwards.

EXPLAINING THE PHOTOELECTRIC EFFECT

The characteristics of the photoelectric effect could not be explained using a wave model of light. According to the wave model, the frequency of light should be irrelevant to whether or not photoelectrons are ejected. Since a wave is a form of continuous energy transfer, it would be expected that energy from the wave would build up in the metal over time. This means that even low-frequency light should transfer enough energy to emit photoelectrons if left incident on the metal for long enough. Similarly, the wave model predicts that there should be a time delay between the light striking the metal and photoelectrons being emitted.

FIGURE 11.2.4 Albert Einstein helped revolutionise our understanding of the nature of light.

The dual nature of light

In 1905, Albert Einstein (Figure 11.2.4) proposed a solution to this problem. Einstein drew on Planck's earlier work by assuming that light exists as particles, or **photons** (like Planck's 'quanta'), each with an energy of $E = hf$. This assumption made the properties of the photoelectric effect relatively easy to explain.

Einstein's work was actually a significant extension of Planck's ideas. Although Planck had assumed that light was being emitted in quantised packets, he never questioned the assumption that light was fundamentally a wave phenomenon.

Einstein's work went further, challenging scientists' understanding of the nature of light itself.

Einstein and the photoelectric effect

Einstein identified that, for a particular metal, the amount of energy required to eject a photoelectron is a constant value that depends on the strength of the bonding within the metal. This energy was called the **work function**, ϕ, of the metal. For example, the work function of lead is 4.14 eV, which means that 4.14 eV of energy is needed to just release one electron from the surface of a piece of lead.

According to Einstein's model, shining light on the surface of a piece of metal is equivalent to bombarding it with photons. When a photon strikes the metal, it can transfer its energy to an electron. That is, a single photon can interact with a single electron, transferring all of its energy at once to the electron. What happens next depends on whether or not the photon contains enough energy to overcome the work function.

If the energy of the photon is less than the work function, then photoelectrons will not be released as the electrons will not gain enough energy to let them break free of the lead atoms. For example, the photons of violet light ($f = 7.50 \times 10^{14}$ Hz) each contain 3.11 eV of energy.

$$\begin{aligned} E &= hf \\ &= 4.14 \times 10^{-15} \times 7.50 \times 10^{14} \\ &= 3.11\,\text{eV} \end{aligned}$$

This means that violet light shining on lead would not release photoelectrons since the energy of each photon, 3.11 eV, is less than the work function of lead, 4.14 eV.

However, ultraviolet photons of frequency 1.20×10^{15} Hz each contain 4.97 eV of energy.

$$\begin{aligned} E &= hf \\ &= 4.14 \times 10^{-15} \times 1.20 \times 10^{15} \\ &= 4.97\,\text{eV} \end{aligned}$$

Therefore ultraviolet light of this frequency would release photoelectrons from the lead since the energy of each photon, 4.97 eV, is greater than the work function of lead, 4.14 eV.

Each metal has a threshold frequency—this is the frequency at which the photons have an energy equal to the work function of the metal:

$$\phi = hf_0$$

where ϕ is the work function (J or eV)
h is Planck's constant (6.63×10^{-34} J s or 4.14×10^{-15} eV s)
f_0 is the threshold frequency for that metal (Hz)

Worked example 11.2.1

CALCULATING THE WORK FUNCTION OF A METAL

Calculate the work function (in J and eV) for aluminium, which has a threshold frequency of 9.8×10^{14} Hz.

Thinking	Working
Recall the formula for the work function of a metal.	$\phi = hf_0$
Substitute the threshold frequency of the metal into this equation.	$\phi = 6.626 \times 10^{-34} \times 9.8 \times 10^{14}$ $= 6.5 \times 10^{-19}$ J
Convert this energy from J to eV. See the SkillBuilder on page 288.	$\phi = \frac{6.5 \times 10^{-19}}{1.602 \times 10^{-19}}$ $= 4.1$ eV

Worked example: Try yourself 11.2.1

CALCULATING THE WORK FUNCTION OF A METAL

Calculate the work function (in J and eV) for gold, which has a threshold frequency of 1.2×10^{15} Hz.

THE KINETIC ENERGY OF PHOTOELECTRONS

If the energy of the photon is greater than the work function of the metal, then a photoelectron is released. The remainder of the energy in excess of the work function is transformed into the kinetic energy of the photoelectron.

Einstein described this relationship with his photoelectric equation:

$K_{max} = hf - \phi$

where

K_{max} is the maximum kinetic energy of an emitted photoelectron (J or eV)

ϕ is the work function of the metal (J or eV)

h is Planck's constant (6.63×10^{-34} J s or 4.14×10^{-15} eV s)

f is the frequency of the incident photon (Hz)

Graphing Einstein's equation results in a linear (straight line) graph like the one shown in Figure 11.2.5. A graph like this is useful because it clearly shows key information such as the work function and threshold frequency for a particular metal.

Einstein's equation, $K_{max} = hf - \phi$, can be compared with the equation of a straight line, $y = mx + c$. In making this comparison, it can be seen that extrapolating (extending) the graph back to the vertical axis will give the magnitude of the work function, ϕ (Figure 11.2.5). The gradient of the graph is Planck's constant, h. From the graph it is also apparent how, as soon as the threshold frequency is exceeded, an electron can be ejected and escape with some kinetic energy. The greater the frequency of the light, the greater the kinetic energy of the photoelectron. At the threshold frequency, electrons are no longer bound to the metal, but they have no kinetic energy.

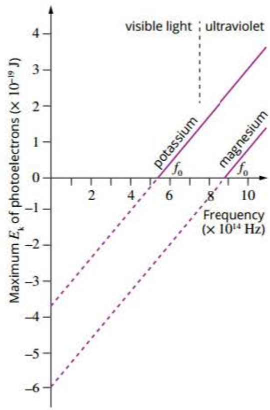


FIGURE 11.2.5 Magnesium has a high threshold frequency, which is in the ultraviolet region. The threshold frequency for potassium is in the visible region. The gradient of the graph for each metal is Planck's constant, h. The x-intercept gives the threshold frequency, f_0. The magnitude of the y-intercept gives the work function, ϕ.

WS 7.7

Worked example 11.2.2

CALCULATING THE KINETIC ENERGY OF PHOTOELECTRONS

Calculate the kinetic energy (in eV) of the photoelectrons emitted from lead by ultraviolet light with a frequency of 1.2×10^{15} Hz. The work function of lead is 4.14 eV. Use $h = 4.14 \times 10^{-15}$ eV s.

Thinking	Working
Recall Einstein's photoelectric equation.	$K_{max} = hf - \phi$
Substitute values into this equation.	$K_{max} = 4.14 \times 10^{-15} \times 1.2 \times 10^{15} - 4.14$ $= 4.97 - 4.14$ $= 0.83$ eV

Worked example: Try yourself 11.2.2

CALCULATING THE KINETIC ENERGY OF PHOTOELECTRONS

Calculate the kinetic energy (in eV) of the photoelectrons emitted from lead by ultraviolet light which has a frequency of 1.5×10^{15} Hz. The work function of lead is 4.14 eV. Use $h = 4.14 \times 10^{-15}$ eV s.

PHYSICSFILE

Albert Einstein

Although Albert Einstein is most famous for his work on relativity (and its related equation $E = mc^2$), he gained his Nobel Prize 'for his services to Theoretical Physics, and especially for his discovery of the law of the photoelectric effect'. His work on relativity was never formally recognised with a Nobel Prize.

PHYSICS IN ACTION ICT CC S

Photovoltaic cells

The photovoltaic cells that are used in many solar panels (Figure 11.2.6) work on the principle of the photoelectric effect. Sunlight falling on the solar panel provides energy that causes photoelectrons to be emitted as a current that can be used to drive electrical appliances.

However, whereas many photoelectric-effect experiments use high-energy photons of ultraviolet light, photovoltaic cells use materials that will produce photoelectrons when exposed to visible light. Most commonly, these are semiconducting materials based on silicon 'doped' with small amounts of other elements.

Although solar cells are designed to produce the highest current possible from sunlight, most commercially available solar cells have an energy efficiency of less than 20%. Scientists hope to improve this in order to make solar cells a better alternative to fossil fuels for large-scale energy generation.

FIGURE 11.2.6 Solar panels are used to convert sunlight into electrical energy using the photoelectric effect.

Resistance to the quantum model of light

This new particle or 'quantum' model of light was not initially well received by the scientific community. It had already been well established that a discrete particle model for light could not explain many of light's properties such as polarisation and the interference patterns produced in Young's experiment, both discussed in Chapter 10.

Most scientists believed instead that wave explanations for the photoelectric effect would eventually be found. However, eventually the quantum model of light was accepted and the Nobel Prize in Physics was awarded to both Planck (1918) and Einstein (1921) for their groundbreaking work in this field.

In order to explain the photoelectric effect, Einstein used the photon concept that Planck had developed. However, like many great discoveries in science, the development of the quantum model of light raised almost as many questions as it answered. It had already been well established that a wave model was needed to explain phenomena such as diffraction and interference. How could these two contradictory models be reconciled to form a comprehensive theory of light?

Answering this question was one of the great scientific achievements of the 20th century and led to the extension of the quantum model to matter as well as energy. It led to a fundamental shift in the way the universe is viewed. Some of the great scientists of that time are shown in the historic photograph in Figure 11.2.7.

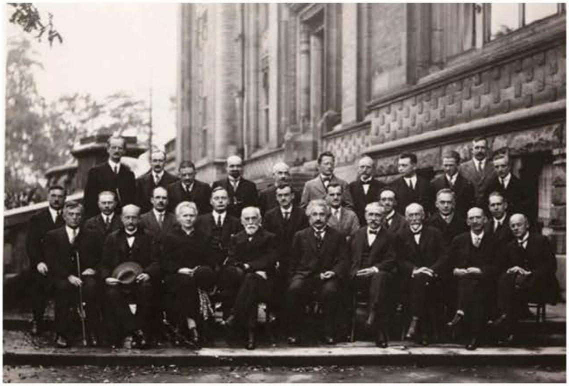


FIGURE 11.2.7 This photo shows the 5th Solvay Conference in Brussels in 1927, which was attended by great scientists including Albert Einstein, Max Planck, Niels Bohr, Marie Curie, Paul Dirac, Erwin Schrödinger and Louis de Broglie. All of these scientists contributed to the current knowledge of the universe, the atom and quantum mechanics.

Experimental evidence for the dual nature of light

In the early years of quantum theory, some scientists believed that the wave properties of light observed in Young's double-slit experiment (discussed in Chapter 10) might have been due to some sort of interaction between photons as they passed through the slits together.

To test this, experiments were done with light sources that were so dim that scientists were confident that only one photon was passing through the apparatus at a time. In this way, any interactions between photons could be eliminated. Over time, these experiments produced identical interference patterns to those done with bright sources (Figure 11.2.8), thus demonstrating the dual nature of light. This feature of light is known as **wave–particle duality** and it will be discussed in greater detail in Chapter 15.

Interestingly, when a detector is used to measure which slit the photon passes through, the wave pattern disappears and the photon acts just like a particle.

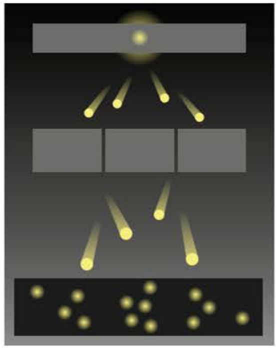


FIGURE 11.2.8 An interference pattern can be built up over time by a series of single photons passing through an apparatus like that used in Young's experiment, demonstrating the wave–particle duality of light.

GO TO ➤ Section 15.2 page 404

11.2 Review

SUMMARY

- The photoelectric effect is the emission of photoelectrons from a clean metal surface due to incident light with a frequency greater than a threshold frequency, f_0.
  - If $f < f_0$, no electrons are released.
  - If $f > f_0$, the *rate* of electron release (the photocurrent) is proportional to the intensity of the light and occurs without any time delay.
- The wave model for light could not explain various features of the photoelectric effect:
  - the existence of a threshold frequency
  - the absence of a time delay when using very weak light sources
  - increased intensity of light resulting in a greater rate of electron release rather than increased electron energy.
- Einstein used Planck's concept of a photon to explain the photoelectric effect, stating that each electron release was due to an interaction with only one photon.
- The work function, ϕ, for the metal is given by $\phi = hf_0$, and is different for each metal. If the frequency of the incident light is greater than the threshold frequency, then a photoelectron will be ejected with some kinetic energy up to a maximum value.
- A graph of a photoelectron's kinetic energy, K_{max}, versus frequency, f, will have a gradient equal to Planck's constant, h, and a y intercept equal to the work function, ϕ.
- $K_{max} = q_eV_0$, where q_e is the charge on an electron and V_0 is the stopping voltage.
- The maximum kinetic energy of the photoelectrons emitted from a metal is the energy of the photons minus the work function, ϕ, of the metal: $K_{max} = hf - \phi$.
- Light exhibits wave properties in some situations and particle properties in other situations. The concept of wave–particle duality is used to describe the dual nature of light.

KEY QUESTIONS

1 When light shines on a metal surface, why might the metal become positively charged?

2 Which of the following statements about the photoelectric effect are true and which are false? For those that are false, rewrite them to make them correct.

a When the intensity of light shining on the surface of the metal increases, the photocurrent increases.

b When light sources of the same intensity but different frequencies are used, the higher frequency light has a higher stopping voltage and produces a higher maximum current than the lower frequency.

c When the applied voltage is positive, photoelectrons are attracted to the collector electrode.

3 Calculate the work functions (in electron-volts) of the following metals (using $h = 4.14 \times 10^{-15}$ eV s).

	Metal	Threshold frequency ($\times 10^{15}$ Hz)
a	lead	1.0
b	iron	1.1
c	platinum	1.5

4 In an experiment on the photoelectric effect, different frequencies of light were shone on a piece of magnesium with a work function of 3.66 eV. Identify which of the following frequencies listed would be expected to produce photoelectrons.

A 3.0×10^{14} Hz

B 5.0×10^{14} Hz

C 7.0×10^{14} Hz

D 9.0×10^{14} Hz

5 Light with a frequency of 9.0×10^{14} Hz is shone onto a piece of magnesium with a work function of 3.66 eV. Calculate the maximum kinetic energy, in electron-volts, of the emitted photoelectrons.

Chapter review

KEY TERMS

black body
electromagnetic radiation
electron-volt
photocurrent
photoelectric effect
photoelectron
photon
quantum
stopping voltage
threshold frequency
wave–particle duality
work function

REVIEW QUESTIONS

1 Identify which of the following phenomena can be explained using a particle (i.e. quantum or photon) model of light and which can be explained using a wave model:

- black-body radiation
- interference patterns
- photoelectric effect
- polarisation

2 The peak wavelength of a star is 502 nm. Use Wien's law to calculate the surface temperature of the star.

3 If a Bunsen burner flame is blue, with a typical wavelength of 455 nm, what is the temperature (in Kelvin) of the flame?

4 The constellation Orion contains two bright stars, Betelgeuse and Rigel. Betelgeuse is red and Rigel is blue. Which star has the higher surface temperature?

5 The star Rigel has a surface temperature of 11 000 K. What is the peak wavelength of the energy emitted by the star?

6 Calculate the energy (in J) of a quantum of light that has a frequency of 8.0×10^{14} Hz.

7 Calculate the energy (in J and eV) of a quantum of light that has a wavelength of 500 nm.

8 What is the energy, in electron-volts, of light with a frequency of 6.0×10^{14} Hz?

9 What is the approximate value of the energy in J of a quantum of light that has an energy of 5.0 eV?

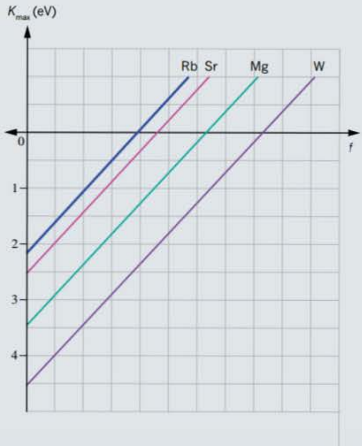
10 What name is given to the electrons released from a metal surface due to the photoelectric effect?

11 If the work function for nickel is 5.0 eV, what is the threshold frequency for nickel?

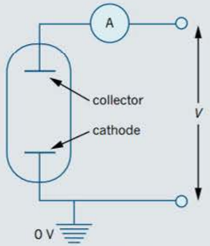
12 Platinum has a threshold frequency of 1.5×10^{15} Hz. Calculate the maximum kinetic energy, in electron-volts, of the emitted photoelectrons when ultraviolet light with a frequency of 2.2×10^{15} Hz shines on it.

13 The stopping voltage obtained using a particular photocell is 1.95 V. Determine the maximum kinetic energy of the photoelectrons in electron-volts.

14 From the graph, determine the value of the work function for each of the metals.



15 The cathode of a particular photocell, shown below, is coated with rubidium. Incident light of varying frequencies is directed onto the cathode of the cell and the maximum kinetic energy of the photoelectrons is logged. The results are summarised in the following table.



Frequency (Hz) × 10^{14}	K_{max} (eV)
5.20	0.080
5.40	0.163
5.60	0.246
5.80	0.328
6.00	0.411
6.20	0.494

a Plot the points from the table on a graph.
b Calculate the gradient of the graph.
c Based on your graph, what is the threshold frequency for rubidium?
d Will red light of wavelength 680 nm cause photoelectrons to be emitted from the rubidium surface? Justify your answer.

16 The following statements describe the value of the stopping voltage obtained when light is incident on a metal cathode. Which statements are true, and which are false? For those that are false, rewrite them to make them true.
a The stopping voltage indicates how much work must be done to stop the most energetic photoelectrons.
b The stopping voltage is reached when the photocurrent is reduced almost to zero.
c If only the intensity of the incident light is increased, the stopping voltage will not alter.
d For a given metal, the value of the stopping voltage is affected only by the frequency of the incident light.

17 The metal sodium has a work function of 1.81 eV. Which of the following types of electromagnetic radiation would cause photoelectrons to be emitted?
A infrared radiation, $\lambda = 800$ nm
B red light, $\lambda = 700$ nm
C violet light, $\lambda = 400$ nm
D ultraviolet radiation, $\lambda = 300$ nm

18 Blue light with a wavelength of 475 nm is shone on a piece of sodium with a work function of 2.36 eV. Calculate the maximum kinetic energy, in electron-volts, of the emitted photoelectrons.

19 When yellow-green light with a wavelength of 500 nm is shone on a metal, the photoelectrons require a stopping voltage of 0.80 V. Calculate the work function of the metal in electron-volts.

20 After completing the activity on page 290, reflect on the inquiry question: What evidence supports the particle model of light and what are the implications of this evidence for the development of the quantum model of light?

CHAPTER 12

Light and special relativity

Galileo and Newton laid the foundations of the 'clockwork universe', a mechanical picture of the world which has underpinned most modern world views. Einstein, along with others such as Bohr and Heisenberg, presented a much richer and more mysterious universe, one that challenges people to think beyond the mechanical picture they so often take for granted.

In this chapter, you will explore the concepts of classical physics, as described by Galileo and Newton, and the evidence that pointed towards the need for some different thinking. Einstein's special relativity is presented as a solution to the problem of classical physics at speeds approaching the speed of light.

Content

INQUIRY QUESTION

How does the behaviour of light affect concepts of time, space and matter?

By the end of this chapter you will be able to:

- analyse and evaluate the evidence confirming or denying Einstein's two postulates:
  - the speed of light in a vacuum is an absolute constant
  - all inertial frames of reference are equivalent (ACSPH131)
- investigate the evidence, from Einstein's thought experiments and subsequent experimental validation, for time dilation $\left(t = \frac{t_0}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}\right)$ and length contraction $\left(l = l_0\sqrt{\left(1-\frac{v^2}{c^2}\right)}\right)$ and analyse quantitatively situations in which these are observed; for example:
  - observations of cosmic-origin muons at the Earth's surface ICT N
  - atomic clocks (Hafele–Keating experiment) CCT ICT N
  - evidence from particle accelerators CCT ICT N
  - evidence from cosmological studies ICT
- describe the consequences and applications of relativistic momentum with reference to:
  - $p_v = \frac{m_0 v}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$ ICT N
  - the limitation on the maximum velocity of a particle imposed by special relativity (ACSPH133) CCT
- use Einstein's mass–energy equivalence relationship ($E = mc^2$) to calculate the energy released by processes in which mass is converted to energy; for example: (ACSPH134) ICT N
  - production of energy by the Sun
  - particle–antiparticle interactions, e.g. positron–electron annihilation
  - combustion of conventional fuel.

12.1 Einstein's postulates

PHYSICS INQUIRY N CCT

Spacetime diagram

How does the behaviour of light affect concepts of time, space and matter?

COLLECT THIS...

- sheet of paper approximately 20 cm × 20 cm
- ruler
- scissors
- thumbtacks or pins
- pinboard or thick piece of cardboard/foam
- marker pens and pencils

DO THIS...

1 Copy the following grid onto the piece of paper. Each dot along the red line represents a stationary observer ($v = 0\ \text{m s}^{-1}$), the R along the purple line represents a rocket ($v = 0.5c$) and the P along the blue line represents a photon ($v = c$). Each row along the grid represents a time of 1 s and each vertical line represents a distance of 1.5×10^8 m.

2 Cut the grid into equal rows for each second of time. Each row will have all three objects in them.

3 Line the strips of paper up on the pinboard, and pin in place with the stationary observers all still in a straight line.

4 Slide the rows of paper along to line up the rockets. Observe how this affects the path of the photon in spacetime.

5 Use rotations and translations to adjust your grid so the path of the photon is continuous when the rockets are all aligned. Record the transformations needed.

RECORD THIS...

Describe how the speed of light being absolute (the same in all inertial reference frames) requires an adjustment in space and time.

Present your model of spacetime from the stationary observer, and adjusted for the rocket observer, highlighting that the speed of light is $3 \times 10^8\ \text{m s}^{-1}$ in both.

REFLECT ON THIS...

How does the behaviour of light affect concepts of time, space and matter?

Why was a dilation in the time axis needed?

How would the speed of the rocket change the adjustments required?

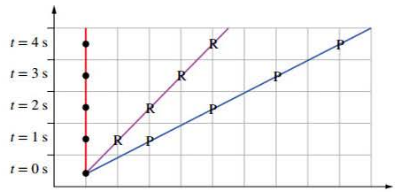


FIGURE 12.1.1 Albert Einstein statue in Washington, D.C.

Galileo and Newton developed theories of motion. These theories allowed the relative motion of low-speed objects to be modelled mathematically. This section presents the observations that challenged Galilean relativity and Newtonian physics, and explains the key principles that led to the new physics described by Albert Einstein (Figure 12.1.1), known as the theory of special relativity.

FRAMES OF REFERENCE

Perhaps it was lucky that in his early twenties Einstein was not part of the physics 'establishment'. He was working as a patent clerk in the Swiss Patent Office. It was an interesting enough job, but it left him time to think about electromagnetic waves (light) and their relationship to the Galilean principle of relativity.

Galileo was particularly interested in relative motion. One of his famous experiments involved the dropping of a cannon ball from the top of the mast of a moving ship. Galileo found that the motion of the cannon ball was not affected by the motion of the ship; the cannon ball landed next to the base of the mast. His principle of relativity was that you cannot tell if you are moving or not without looking outside of your own **frame of reference**. A frame of reference describes where an observation is being made from. Sometimes this may be from a stationary point, i.e. standing on a platform watching a train pull into a station. But at other times the frame of reference may also be in motion.

Based on the work of Galileo, Isaac Newton established detailed models for the motion of objects such as planets, moons and comets, even falling oranges. According to his equations, the velocity of objects can be calculated relative to any frame of reference as long as the velocity of the frame of reference is known. The Newtonian principle that the velocities of objects and frames of reference can be added together to determine the velocity of the object in another frame of reference is common throughout his equations and laws.

Consider an object moving in a frame of reference, A. This frame of reference is moving in another frame of reference, B. The velocity of the object in frame B is given by:

$$v_{\text{object in B}} = v_{\text{object in A}} + v_{\text{A in B}}$$

A practical example of this could be when a person runs forwards while on a train. Here, the train is frame of reference A and the track along which the train moves is frame B. Imagine that the person runs at $5\,\text{m s}^{-1}$ forwards, while the train travels at a velocity of $20\,\text{m s}^{-1}$ forwards. The velocity of the person relative to B, the track, is:

$$\begin{aligned} v_{\text{person along track}} &= v_{\text{person in train}} + v_{\text{train along track}} \\ &= 5 + 20 \\ &= +25\,\text{m s}^{-1} \end{aligned}$$

That is, the person is moving with a velocity of $25\,\text{m s}^{-1}$ forwards when measured against the track.

Einstein was a typical theoretician; the only significant experiments he ever did were thought experiments. Many of his experiments involved thinking of situations that involved two frames of reference moving with a steady relative velocity, in which the principles of Galilean relativity applied. Newton had referred to these as **inertial frames of reference**, as the law of inertia applied within them.

EINSTEIN AND GALILEAN RELATIVITY

Einstein decided that the elegance of the principle of Galilean relativity was such that it simply had to be true. Nature did not appear to have a special frame of reference, and Einstein could see no reason to believe that there was one waiting to be discovered. In other words, there is no such thing as an absolute velocity. It is not possible to have a velocity relative to space itself, only to other objects within space. So the velocity of any object can always be stated as relative to some other object. In the case of the person running on the train, their velocity can be stated as either $5\,\text{m s}^{-1}$ relative to the train or $25\,\text{m s}^{-1}$ relative to the track.

Einstein expanded the Galilean principle to state that all inertial frames of reference must be equally valid, and that the laws of physics must apply equally in any frame of reference that is moving at a constant velocity. So there is no physics experiment you can do that is entirely within a frame of reference that will tell you that you are moving. In other words, as you speed along in your train with the blinds down, you cannot measure your speed. You can tell if you are accelerating easily enough: just hang a pendulum from the ceiling. However, the pendulum will hang straight down whether you are travelling steadily at $100\,\text{km h}^{-1}$ or are stopped at the station. Consider Figure 12.1.2a and b on page 302. There is no way of telling which of the trains is stationary relative to the ground, or which is moving at a constant velocity.

Einstein decided that the relativity principle could not be abandoned. Recall that Einstein was, at the time, thinking about the relationship between light and relativity. Whatever the explanation for the strange behaviour of light, it could not be based on a flaw in the principle of Galilean relativity.

Einstein's fascination with the nature of light had led him to a deep understanding of Maxwell's work on the electromagnetic nature of light waves. He was convinced of the elegance of Maxwell's equations and their prediction of a constant speed of light. Most physicists believed that the constant speed predicted by Maxwell's equations referred to the speed of light relative to a **medium** (a substance it travelled through). It was thought that the speed predicted would be the speed in the medium in which light travelled, and the measured speed would have to be adjusted for one's own speed through that medium.

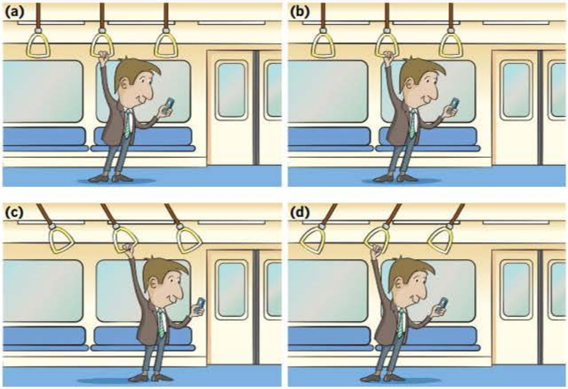


FIGURE 12.1.2 There is no observation or experiment that shows the difference between two inertial frames of reference (a) and (b). In one of the situations illustrated, the train is stationary, and in the other it is moving smoothly at $100\,km\,h^{-1}$. There is no observation that will tell which one is which. In (c) and (d), the motion of the handles hanging from the ceiling of the train indicate that these trains are not moving at a constant speed.

As light travelled through the vacuum of space between the Sun and Earth, clearly the medium was no ordinary material. Physicists gave it the name **aether**, as it was an 'ethereal' substance. It was thought, following Maxwell's work, that the aether must be some sort of massless, rigid medium that 'carried' electric and magnetic fields.

This was a real problem for Einstein. A speed of light that is fixed in the aether and which depended on the velocity of an inertial frame in the aether would be in direct conflict with the principle of Galilean relativity, which Einstein was reluctant to abandon.

Resolving the problem of the aether

As in any conflict, the resolution is usually found by people who are prepared to look at it in new ways. This was the essence of Einstein's genius. Instead of looking for faults in what appeared to be two perfectly good principles of physics, he decided to see what happened if they were both accepted, despite the apparent contradiction.

So Einstein swept away the problem of the aether, saying that it was simply unnecessary. It had been invented only to be a medium for light waves, and no one had found any evidence for its existence. Electromagnetic waves, he said, could apparently travel through empty space without a medium. Doing away with the aether, however, did not solve the basic conflict between the absolute speed of light and the principle of relativity.

EINSTEIN'S THEORY OF SPECIAL RELATIVITY

Though Einstein accepted both Galileo's and Maxwell's theories despite the apparent contradiction, this still left the question: How could two observers travelling at different speeds measure the same light beam travelling at the same speed? The answer, Einstein said, was in the very nature of space and time.

In 1905 he sent a paper to the respected physics journal *Annalen der Physik* entitled 'On the electrodynamics of moving bodies'. In this paper he put forward two simple **postulates** (statements assumed to be true) and followed them to their logical conclusion. It was this conclusion that was so astounding.

PHYSICSFILE ICT

The Michelson–Morley experiment

The existence of an aether appeared to be a serious blow for the principle of relativity. It seemed that, after all, there may be a frame of reference attached to space itself. If this was the case, there was the possibility of an absolute zero velocity.

Scientists needed to test the idea of electromagnetic waves moving through the aether. Since the Earth is in orbit around the Sun, an aether wind should be blowing past the Earth. This suggested to American physicist Albert Michelson that it should be possible to measure the speed at which the Earth was moving through the aether by measuring the small changes in the speed of light as the Earth changed its direction of travel. For example, if the light was travelling in the same direction as the Earth, through the aether, the apparent speed should be slower than usual at $c - v$ (Figure 12.1.3). It would be as if the light was travelling against an aether 'wind' created by the motion of the Earth through it. If the light was travelling against the Earth's motion, the apparent speed should be faster as it would be travelling with the 'wind' at $c + v$ (Figure 12.1.3). The differences would be tiny, less than 0.01%, but Michelson was confident that he could measure them.

In the 1880s Michelson and his collaborator Edward Morley set up a device known as an interferometer. The device cannot measure the speed of light but it can detect changes in the speed of light that might have been due to the aether wind. In fact, it was used to attempt to measure the very small differences in the time taken for light to travel in two mutually perpendicular directions. They were able to rotate the whole apparatus and hoped to detect the small difference that should result from the fact that one of the directions was to be the same as that in which the Earth was travelling and the other at right angles. However, they found no difference. Perhaps, then, the Earth at that time was stationary with respect to the aether? Six months later, however, when the Earth would have to be travelling in the opposite direction relative to the aether, there was still no difference in the measured speeds! Other people performed similar experiments, virtually always with the same null result. Whatever direction the Earth was moving it seemed to be at rest in the aether. Or perhaps there was no aether at all.

While Michelson and Morley's results were consistent with Maxwell's prediction that the speed of light would always appear to be the same for any observer, the apparent absurdity of such a situation led most physicists to believe that some flaw in the theory behind the experiment, or in its implementation, would soon be discovered. Einstein, however, wondered about the consequences of actually accepting their prediction about the speed of light but at the same time holding on to the relativity principle.

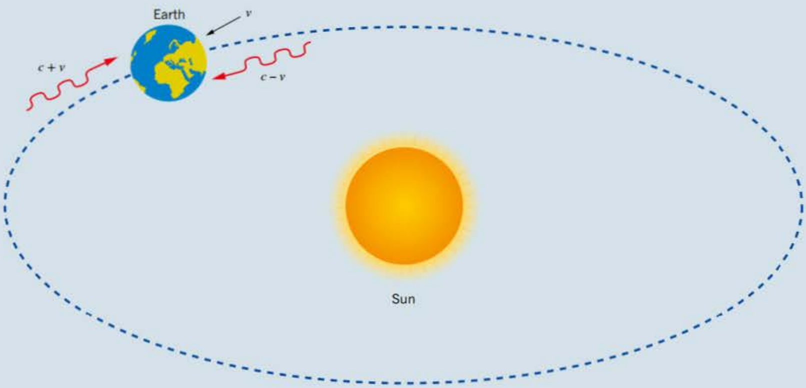


FIGURE 12.1.3 The basic principle of the Michelson–Morley experiment. If the aether is fixed relative to the Sun, and the light is travelling at c relative to the aether and in the same direction as the Earth, the apparent speed should be less than c, i.e. $c - v$. If the light was travelling in the opposite direction to Earth it should appear faster than c, i.e. $c + v$.

Einstein's two postulates:

- The laws of physics are the same in all inertial (non-accelerated) frames of reference.
- The speed of light has a constant value for all observers regardless of their motion or the motion of the source.

(The first postulate means that there is no preferred frame of reference and so is sometimes stated as: no law of physics can identify a state of absolute rest.)

EINSTEIN'S POSTULATES

The first postulate is basically that of Newton, but Einstein extended it to include the laws of electromagnetism, so elegantly expressed by Maxwell. The second postulate simply takes Maxwell's prediction about the speed of electromagnetic waves in a vacuum at face value.

These two postulates sound simple enough; the only problem was that, according to early Newtonian physics, they were contradictory.

Consider the example illustrated in Figure 12.1.4. Binh is in his spaceship travelling away from Clare at a speed v, and Clare turns on a laser beam to signal Binh. The first postulate seems to imply that the speed of the laser light, as measured by Binh, should be $c - v$, where c is the speed of light in Binh's frame of reference. This is what you would expect if, for example, you were to measure the speed of sound as you travel away from its source; as your velocity gets closer to the speed of sound, the slower the soundwaves appear to be travelling.

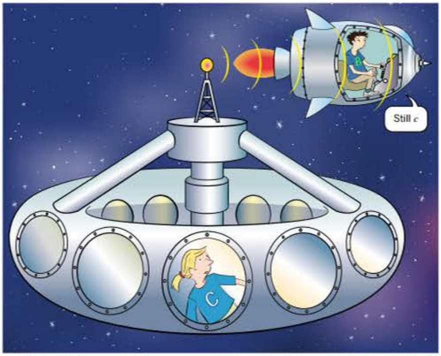


FIGURE 12.1.4 Einstein's two postulates are seemingly contradictory. His first postulate indicates that the speed of the laser light, as measured by Binh, should be $c - v$, whereas his second postulate indicates it should be c. Einstein revisited Newton's assumptions to resolve this problem.

The second postulate, however, tells you that when Binh measures the speed of Clare's laser light, he will find it to be c; that is, $3.00 \times 10^8\,\mathrm{m\,s^{-1}}$. So at first glance, these two postulates appear to be mutually exclusive. To resolve this problem, Einstein went back to the assumptions on which Newton based his theories.

Newton's assumptions

In 1687, Isaac Newton published his famous *Principia*. At the start of this incredible work, which was the basis for all physics in the next two centuries and beyond, he notes the two fundamental assumptions.

The following two statements are assumed to be evident and true:

- Absolute, true, and mathematical time, of itself, and from its own nature, flows equably without relation to anything external.
- Absolute space, in its own nature, without relation to anything external, remains always similar and immovable.

Newton based all of his laws on these two assumptions: that space and time are constant, uniform and straight. So according to Newton, space is like a big set of *xyz* axes that always have the same scale, and in which distances can be calculated exactly according to Pythagoras' theorem. You expect a metre ruler to be the same length whether it is held vertically or horizontally, north–south or east–west, in your classroom or in the International Space Station.

In this space, time flows on at a constant rate, which is the same everywhere. One second in Perth is the same as one second in Sydney, and one second on the ground is the same as one second up in the air.

Einstein realised that the assumptions that Newton made may not be valid, at least not on scales involving huge distances and speeds approaching the speed of light. The only way in which Einstein's two postulates can both be true is if both space and time are not fixed and unchangeable.

Einstein's train

To illustrate the consequences of accepting the two postulates he put forward, Einstein discussed a simple thought experiment. It involves a train, moving at a constant velocity.

Amaya and Binh have boarded Einstein's train and Clare is outside on the platform (Figure 12.1.5). This train has a flashing light bulb set right in the centre of the carriage. Amaya and Binh observe the flashes of light as they reach the front and back walls of the carriage. They find that the flashes reach the front and back walls at the same time, which is not surprising. Outside, Clare measures the same flashes of light. Einstein was interested in when Clare saw the flashes reach the end walls.

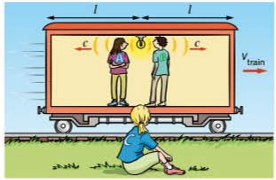


FIGURE 12.1.5 Amaya and Binh observe that the light takes the same time, $\frac{l}{c}$ seconds, to reach the front and back walls.

To appreciate Einstein's ideas, you need to contrast them with what you would normally expect. Consider a situation in which Amaya and Binh are rolling balls towards opposite ends of a train carriage. It is important to appreciate that, while Clare, the outside observer, measures the ball's velocity differently from Amaya and Binh, the times at which various events (balls hitting the ends of the carriage) occur must be the same.

If you had discussed a pulse of soundwaves travelling from the centre of the train, you would find exactly the same result: Clare always agrees with Amaya and Binh that the time taken for balls, or soundwaves, to reach the end walls is the same. But what about light?

Einstein's second postulate tells you that all observers measure light travelling at the same speed. Amaya, Binh and Clare will all measure the light travelling at $3.00 \times 10^{8}\,\mathrm{m\,s^{-1}}$; they do not add or subtract the speed of the train.

If Clare observes the light travelling at the same speed in the forward and backward directions, she will measure the light reaching the back wall first (Figure 12.1.6). This is because that wall is moving towards the light, whereas the front wall is moving away from the light, and so the light will take longer to catch up to it. This is against the principles of Newtonian physics. Amaya and Binh observe that the light flashes reach the ends of the carriage at the same time; Clare's measurements saw them reach the walls at different times.

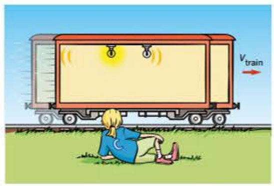


FIGURE 12.1.6 Clare measures the light reaching the back wall first, and then the front wall.

The idea that two events that are **simultaneous** (occur at the same time) for one set of observers but are not simultaneous for another seems outrageous.

PHYSICSFILE L ICT

Measurement in a thought experiment

The people in Einstein's train would need extremely good measuring devices, such as an atomic clock (see Section 12.2), and amazingly quick reflexes in order to take their measurements.

Under normal circumstances, there is no chance of detecting the lack of simultaneity of light beams hitting the front and back walls of a train. This is because the differences in time are about a millionth of a microsecond, well beyond the capacity of even the best stopwatches. The reflexes required to see the light reach the back wall, then see the light encounter the front wall, would also be beyond human ability.

Simultaneity and spacetime

The big difference between the situation for light and that for balls or sound is the strange notion that both sets of observers measure the speed of light as exactly the same. The velocity of a thrown ball or the velocity of sound in Amaya and Binh's frame of reference will always be different from that in Clare's frame of reference by exactly the velocity of the train. For light, however, there is no difference. As a result, events that are simultaneous for one set of observers are *not simultaneous* for the others. This is a very strange situation that is referred to as a lack of simultaneity.

While Einstein's experiments are purely hypothetical, other experiments based on these ideas are well within the capacity of modern experimental physics. In all cases they confirm Einstein's ideas to a high degree of accuracy.

Einstein said that the only reasonable explanation for how two events that were simultaneous to one set of observers were not simultaneous to another is that time itself is behaving strangely. The amount of time that has elapsed in one frame of reference is not the same as that which has elapsed in another (Figure 12.1.7).

FIGURE 12.1.7 The famous clock tower in Bern, Switzerland, near Einstein's apartment. Its hands move at one minute per minute, but only in the same frame of reference as the clock.

In the example shown in Figure 12.1.5, Amaya and Binh saw the light flashes that went forwards and backwards take the same time to reach the walls. In Clare's frame of reference the times were different. Time, which has one dimension, seems to depend on the frame of reference in which it is measured, and a frame of reference is just a way of defining three-dimensional space. Clearly time and space are somehow interrelated. This four-dimensional relationship, which includes the three dimensions of space and the one dimension of time, is called **spacetime**. Special relativity is all about spacetime. Spacetime coordinates then describe events with a spatial reference at a specific time.

This was a profound shock to the physicists of Einstein's time. Many of them refused to believe that time was not the constant and unchanging quantity that it was assumed always to have been. And to think that it might 'flow' at a different rate in a moving frame of reference was too mind-boggling for words. That could mean that if you went for a train trip, your clock would go slower, and you would come back having aged slightly less than those who stayed behind.

Einstein's idea was that time and distance are relative. They can have different values when measured by different observers. Simultaneous events in one frame of reference are not necessarily simultaneous when observed from another frame of reference. This is difficult to comprehend at first and will take some time to fully appreciate. Our basic understanding of time and distance (and perhaps mass too) needs adjustment when objects travel close to the speed of light. A certain observer might measure light travelling through a distance d in a time t at a speed c. A different observer might measure light travelling through a different distance, d', in a different time, t', but still at the same speed, c.

Probably because of the tiny differences involved and the highly abstract nature of the work, many physicists simply disregarded the concepts and got on with their work. They thought it could never have any practical results.

12.1 Review

SUMMARY

- Einstein decided that Galileo's principle of relativity was so elegant it simply had to be true, and he was also convinced that Maxwell's electromagnetic equations, and their predictions, were sound.
- Einstein's two postulates of special relativity can be abbreviated to:
  - I The laws of physics are the same in all inertial frames of reference.
  - II The speed of light is the same to all observers.
- Einstein realised that accepting both of these postulates implied that space and time were not absolute and independent, but were related in some way.
- Two events that are simultaneous in one frame of reference are not necessarily simultaneous in another.
- This implies that time measured in different frames of reference might not be the same. Time and space are related in a four-dimensional universe of spacetime.

KEY QUESTIONS

1 Why did the physicists of the late 19th century feel the need to invent the idea of the aether?

A It was required to satisfy the principle of relativity.
B It was required to satisfy Maxwell's equations.
C They thought that it would be impossible that totally empty space could occur in nature.
D They thought that there should be a medium that carries light waves just as air carries soundwaves.

2 Which of the following are reasonably good inertial frames of reference? More than one correct answer is possible.

A an aircraft in steady flight
B an aircraft taking off
C a car turning a corner
D a car driving up a hill of constant slope at a steady velocity

3 Two spaceships are travelling for a while with a constant relative velocity. Then one begins to accelerate. A passenger with a laser-based velocity measurer finds that the relative velocity increases. Give an example of how this passenger could tell whether it was his own or the other ship that began to accelerate.

4 Tom, who is in the centre of a train carriage moving at constant velocity, rolls a ball towards the front of the train, while at the same time he blows a whistle and shines a laser towards the front of the train. What will Jana, who is on the ground outside the train, observe compared with Tom about the speed of the ball, the sound and the light?

5 If the speed of sound in air is $340\,\text{m}\,\text{s}^{-1}$, at what speed would the sound from a fire truck siren appear to be travelling in the following situations?

a You are driving towards the stationary fire truck at $30\,\text{m}\,\text{s}^{-1}$.
b You are driving away from the stationary truck at $40\,\text{m}\,\text{s}^{-1}$.
c You are stationary and the fire truck is heading towards you at $20\,\text{m}\,\text{s}^{-1}$.
d You are level with the fire truck and are about to overtake it, while it is travelling at $20\,\text{m}\,\text{s}^{-1}$ in the same direction.

6 In order to resolve the apparent conflict resulting from his two postulates, Einstein rejected some of Newton's assumptions. Which of the following statements is a consequence of this?

A Time is not constant in all frames of reference.
B Absolute, true, and mathematical time, of itself, and from its own nature, flows equably without relation to anything external.
C One second in any inertial frame of reference is the same as one second in any other inertial frame of reference.
D Space and time are independent of each other.

12.1 Review *continued*

7 Anna is at the front end of a train carriage moving at $10\,m\,s^{-1}$. She throws a ball back to Ben, who is 5 m away at the other end of the carriage. Ben catches it 0.2 s after it was thrown. Chloe is watching all this from the side of the track.

a At what velocity does Chloe measure the thrown ball travelling?

b How far, in Chloe's frame of reference, did the ball move while in flight?

c How long was it in flight in Chloe's frame of reference?

8 Imagine that the speed of light has suddenly slowed down to only $50\,m\,s^{-1}$ and this time Anna (still at the front of the 5 m train moving at $10\,m\,s^{-1}$ in Question 7) sends a flash of light towards Ben.

a From Anna's point of view, how long does it take the light flash to reach Ben?

b How fast was the light travelling in Ben's frame of reference?

c In Chloe's frame of reference, how far did the train travel in 0.1 s?

d How fast was the light travelling in Chloe's frame of reference?

e Approximately when did Chloe measure the light reaching Ben?

12.2 Evidence for special relativity

Measuring time is an exercise in precision, replicating an interval of one second over and over again, until 86 400 of them equals the time for one rotation of the Earth, or one day. There have been many mechanical solutions to this problem in the past using cogs and levers, weights and dials. The accuracies of these devices varied, with some of them gaining or losing seconds or minutes per day.

Before 1967, the standard of one second was based on a fraction of the time it took for the Earth to orbit the Sun, a far-from-ideal standard. From 1967 onwards, the basis for the unit of time was changed to be a certain number of transitions of the outermost electron of a caesium-133 isotope. In fact, one second is now defined as 9 192 631 770 oscillations of the 6s electron of the Cs-133 isotope. The remarkable precision of this oscillation resulted in atomic clocks with an accuracy of 1 second in 1.4 million years, and the ability to measure time to an incredible number of decimal places. It is at these levels of measurement that the predictions of Newton's laws of motion vary from the measured values.

Extremely precise atomic clocks (Figure 12.2.1) enabled very short-lived events to be measured to a large number of decimal places. At this level of precision, some unusual observations were made regarding the life spans of some high-speed subatomic particles when compared to the life spans of those same particles at rest. This section explores the concept of **time dilation** and length contraction as an explanation for these observations.

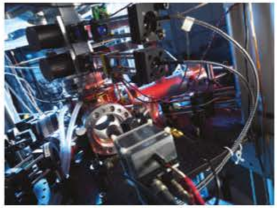


FIGURE 12.2.1 The duration of one second can be measured very precisely using a caesium atomic clock like this one.

TIME IN DIFFERENT FRAMES OF REFERENCE

The consequences of Einstein's two postulates have been discussed, in general terms, when they are applied to a simple thought experiment situation, such as a moving train. Observers inside the train measure two simultaneous events, while those outside measure the same two events occurring at different times. Certainly the differences are extremely small and would not be noticeable by an observer in any actual train, unless they had a clock that could measure very, very small time intervals. For aircraft flying at supersonic speeds, the differences, while very small, become measurable by the most precise clocks. For subatomic particles, such as pions in accelerators like the Australian Synchrotron, the differences in time become more significant, and so in situations like this, where speeds approach the speed of light, it is important to use calculations that take Einstein's theory into account.

The light clock

Consider Amaya and Binh riding in a spaceship that can travel at speeds close to the speed of light. Clare is going to watch from a space station, which according to Clare is a stationary frame of reference. Amaya and Binh have taken along a clock, which (it is assumed) Clare can read, even from a large distance away.

Like any clock, this clock is governed by a regular oscillation that defines a period of time.

Amaya's clock has a light pulse that bounces back and forth between two mirrors. One mirror is on the floor and the other on the ceiling, as shown in Figure 12.2.2. When a light pulse oscillates from one mirror to the other and back, you can consider that period of time to be 'one unit'. Clare has an identical clock in her own space station, which she can compare to Amaya's clock.

The advantage of this clock is that it can be used to predict how motion will affect it by using Pythagoras's theorem and some algebra. The clock has been set up in the spaceship so that the light pulses oscillate up and down a distance d that is at right angles to the direction of travel. The distance d is shown by a black arrow in the centre position of the moving spacecraft in Figure 12.2.3 on page 310. As the spaceship speeds along, the light will trace out a zigzag path, as shown by the red dotted line in Figure 12.2.3.

Only one of the oscillations of the light pulse needs to be considered, as all the other oscillations will have the same geometry.

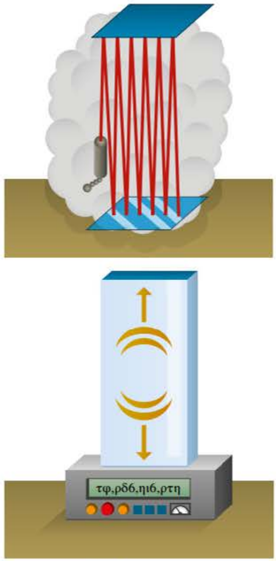


FIGURE 12.2.2 The light clock 'ticks' each time the light pulse reflects off the bottom mirror.

One 'unit of time' will be the time taken for the light pulse to oscillate once. In the frame of reference of the spaceship, Amaya and Binh measure a unit of time equal to t_a. Clare, from her frame of reference, will measure a different time, t_c. The relationship between these two times will now be determined.

Amaya and Binh see the light pulse travel at the speed of light, c, along the distance $2d$, from the bottom mirror to the top and back again, in time t_a. So the distance that the light pulse travels is given by:

$$2d = c \times t_a$$

On the other hand, Clare measures the light travelling a longer path that is shown as the red dotted line in Figure 12.2.3.

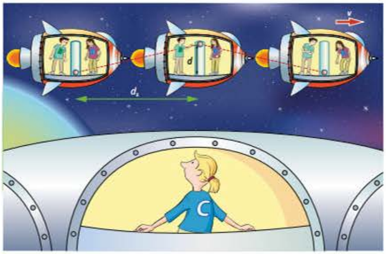


FIGURE 12.2.3 Clare can measure that in one unit of time the light clock 'ticks' each time the light pulse reflects off the bottom mirror. She also finds that the light pulses travel a zigzag path between the mirrors.

The ship moves with a speed v, and so in one unit of time as measured by Clare, t_c, the spaceship will travel a distance $2 \times d_s$, equal to the velocity multiplied by the time taken for her to see one oscillation:

$$2d_s = v \times t_c$$

Consider only half of the light oscillation for now. The light pulse not only travels the vertical distance d in the clock, but also travels forwards as the spaceship moves through the distance d_s, making the combined distance d_c. Therefore, according to Pythagoras's theorem:

$$d_c^2 = d^2 + d_s^2$$

$$d_c^2 = d^2 + \left(\frac{vt_c}{2}\right)^2$$

$$d_c = \sqrt{\left(d^2 + \left(\frac{vt_c}{2}\right)^2\right)}$$

Clare measures this light pulse travelling twice this combined distance at the speed of light, c, in a period of time t_c measured on her clock. So:

$$2d_c = c \times t_c$$

Equating and rearranging the two expressions for d_c gives:

$$\frac{c \times t_c}{2} = \sqrt{\left(d^2 + \left(\frac{vt_c}{2}\right)^2\right)}$$

$$c \times t_c = 2 \times \sqrt{\left(d^2 + \left(\frac{vt_c}{2}\right)^2\right)}$$

$$c \times t_c = \sqrt{\left(4d^2 + 4 \times \left(\frac{vt_c}{2}\right)^2\right)}$$

$$t_c = \frac{\sqrt{4d^2 + (vt_c)^2}}{c}$$

From Amaya and Binh's frame of reference, where they measure the light pulse travelling a distance $2d$ at speed c in a time t_a, the previously given equation can be rewritten in terms of d as:

$$d = \frac{c \times t_a}{2}$$

Note that you have used the same value for c in both of these equations, something you would never do in **classical physics**, but something Einstein insists you must.

Substituting this expression for d into the previous equation gives:

$$t_c = \frac{\sqrt{4 \times \left(\frac{ct_a}{2}\right)^2 + (vt_c)^2}}{c}$$

Now square both sides and simplify to make t_c^2 the subject:

$$t_c^2 = \frac{\frac{4(ct_a)^2}{4} + (vt_c)^2}{c^2}$$

$$= \frac{c^2 t_a^2 + v^2 t_c^2}{c^2}$$

$$= \frac{c^2 t_a^2}{c^2} + \frac{v^2 t_c^2}{c^2}$$

$$= t_a^2 + \frac{v^2 t_c^2}{c^2}$$

Group the terms with t_c^2 together and factorise:

$$t_c^2 - \frac{v^2 t_c^2}{c^2} = t_a^2$$

$$t_c^2\left(1 - \frac{v^2}{c^2}\right) = t_a^2$$

Take the square root of both sides and make t_c the subject:

$$t_c\sqrt{\left(1 - \frac{v^2}{c^2}\right)} = t_a$$

$$t_c = \frac{t_a}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}}$$

As v can never be larger than c, the denominator in the equation above must be less than one. Any number divided by a number less than one must result in a larger number, so $t_c > t_a$.

This final equation shows that the time that Clare measures, t_c, is greater than the time that Amaya and Binh measure, t_a, for the same event.

TIME DILATION

In Einstein's equation for time dilation, the symbol t is used to represent the time that the observer measures for an event occurring in a separate, moving frame of reference. The symbol t_0 is then the time that passes on the moving clock, which is also known as the **proper time**.

The factor that the proper time is multiplied by is given the Greek symbol gamma, γ, so that:

$t = t_0\gamma$

where

$$\gamma = \frac{1}{\sqrt{\left(1 - \frac{v^2}{c^2}\right)}}$$

v is the speed of the moving frame of reference

c is the speed of light in a vacuum ($3.00 \times 10^8\,\text{m s}^{-1}$)

t is the time observed in the stationary frame

t_0 is the time observed in the moving frame (proper time)

The physicist H. A. Lorentz first introduced the factor γ in an attempt to explain the results of the Michelson–Morley experiment, so it is often known as the **Lorentz factor**.

PHYSICSFILE N

The zigzag path of light

Mathematically, you can see that time dilation results from the strange behaviour of light. As light travels on the diagonal zigzag path, it does so at speed c, not at a faster speed resulting from the additional component of the spaceship's motion as, for example, would be true for a boat zigzagging across a river as it is carried along by the current.

Table 12.2.1 and Figure 12.2.4 show the effect of varying the value of v on the value for γ.

TABLE 12.2.1 The value of the Lorentz factor at various speeds.

v ($m\,s^{-1}$)	$\frac{v}{c}$	γ
3.00×10^2	0.000001	1.000000000
3.00×10^5	0.00100	1.0000005
3.00×10^7	0.100	1.005
1.50×10^8	0.500	1.155
2.60×10^8	0.866	2.00
2.70×10^8	0.900	2.29
2.97×10^8	0.990	7.09
2.997×10^8	0.999	22.4

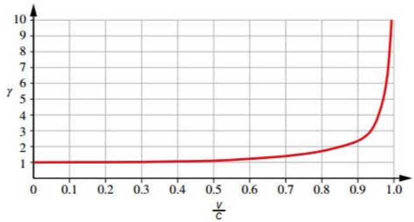


FIGURE 12.2.4 The graph of the Lorentz factor versus $\frac{v}{c}$.

From the data in Table 12.2.1, a velocity of $300\,m\,s^{-1}$ results in a Lorentz factor of essentially 1. So for relatively low-speed spaceships, a stationary observer measures the oscillation of light in the light clock on the spaceship to be the same as in their own stationary light clock. This implies that time is passing at essentially the same rate in both frames of reference.

When the spaceship is travelling at $0.990c$, a stationary observer like Clare will measure that a single oscillation of light in the spaceship's light clock will take seven oscillations of her own stationary light clock. According to Clare, time for the objects and people in the moving frame of reference has slowed down to one-seventh of 'normal' time.

As the speed approaches the speed of light, time in the moving frame, as viewed from the stationary frame, slows down more and more. So, if you were able to see the clock travelling on a light wave, the clock would not be 'ticking' at all. In other words, time would be seen to stand still.

It is important to realise that Amaya and Binh do not perceive their time slowing down at all. To them, their clock keeps ticking away at the usual rate and events in their frame of reference take the same time as they normally would. It is the series of events that Clare measures in Amaya and Binh's frame that go slowly. Binh and Amaya are moving in slow motion because, according to Clare's observations, time for them has slowed down (Figure 12.2.5).

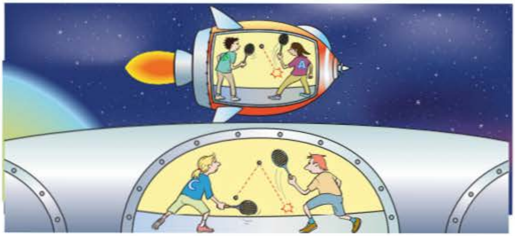


FIGURE 12.2.5 As Clare watches Amaya and Binh play space squash, the ball seems to be moving much more slowly than in her own game.

Worked example 12.2.1

TIME DILATION

A stationary observer on Earth measures a very fast car passing by, travelling at $2.50 \times 10^8\,\text{m s}^{-1}$. In the car's frame of reference, 3.00 s are observed for this event. Calculate how many seconds pass by on the stationary observer's clock during this observation. Use $c = 3.00 \times 10^8\,\text{m s}^{-1}$.

Thinking	Working
Identify the variables: the time for the stationary observer is t, the proper time for the moving clock is t_0, and the velocities are v and the constant c.	$t = ?$ $t_0 = 3.00\,\text{s}$ $v = 2.50 \times 10^8\,\text{m s}^{-1}$ $c = 3.00 \times 10^8\,\text{m s}^{-1}$
Use Einstein's time dilation formula and the Lorentz factor.	$t = t_0\gamma$ $= \frac{t_0}{\sqrt{1-\frac{v^2}{c^2}}}$
Substitute the values for t_0, v and c into the equation and calculate the answer, t.	$t = \frac{3.00}{\sqrt{1-\frac{(2.50\times10^8)^2}{(3.00\times10^8)^2}}}$ $= \frac{3.00}{0.55277}$ $= 5.43\,\text{s}$

Worked example: Try yourself 12.2.1

TIME DILATION

A stationary observer on Earth measures a very fast scooter passing by, travelling at $2.98 \times 10^8\,\text{m s}^{-1}$. On the wrist of the rider is a watch on which 60.0 s pass. Calculate how many seconds pass by on the stationary observer's clock during this observation. Use $c = 3.00 \times 10^8\,\text{m s}^{-1}$.

Looking back to the stationary observer

So far you have been looking at the situation from Clare's point of view, not Amaya's and Binh's. Galileo had said that all inertial frames of reference are equivalent. It follows then that, according to Amaya and Binh, as they look out their window at Clare in her space station receding from them, they can consider that it is they who are at rest and it is Clare and her space station that are moving away at a velocity near the speed of light. This is what Galileo's principle of relativity and Einstein's first postulate are all about.

If Amaya and Binh watch the light clock in Clare's space station, they see that time has slowed down for Clare, as they would observe Clare's moving light-clock oscillation taking longer than their stationary light-clock oscillation. This raises the question: Whose time actually runs slowly?

The answer is that they are both right. The whole point of relativity is that you can only measure quantities relative to some particular frame of reference, not in any absolute sense. Certainly Amaya and Binh see Clare as though in slow motion and Clare sees them in slow motion. Remember that there is no absolute frame of reference and so there is no absolute clock ticking away the absolute 'right' time. All that you can be sure of is that time in your own inertial frame of reference is ticking away at a rate of one second per second.

PHYSICS IN ACTION ICT

The Hafele–Keating experiment

Earlier in this section, it was stated that observers in a frame of reference moving relative to a second frame of reference would measure time to be 'different'. As discussed, these differences would be very small and could only be measured if the first frame of reference was moving at supersonic speeds and if very precise clocks were utilised to measure the times.

This is exactly what Joseph C. Hafele and Robert E. Keating set out to do in 1971. Hafele, a physicist, and Keating, an astronomer, performed what is now known as the Hafele–Keating experiment to test Einstein's theory of relativity, and more specifically time dilation.

Hafele and Keating used four caesium-beam atomic clocks, placing two in aeroplanes to fly around the world. They compared the times from these clocks to two clocks that remained in the United States Naval Observatory. The aeroplanes first flew eastwards (in the same direction as the rotation of the Earth), and then they flew westwards. Theory suggested a greater time dilation would occur when the aeroplanes flew westwards, as there is a greater relative motion between the Earth and the aeroplane in this situation. It was predicted that the clocks that flew eastwards would lose 40 ± 23 ns and the clocks that flew westwards would gain 275 ± 21 ns.

The actual results were a loss of 59 ± 10 ns for the eastward trip and a gain of 273 ± 7 ns for the westward trip. Hafele and Keating's results matched the prediction within 10% and experimentally confirmed time dilation.

The twin paradox

If Clare measures time for Amaya and Binh running slowly, then Amaya and Binh will age slowly. But if Amaya and Binh measure that time for Clare has slowed down, then Clare will age more slowly. So what happens when Amaya and Binh decide to turn their spaceship around and come home? Who will have aged more?

To solve this **paradox**, or contradiction, Einstein described a thought experiment in which one of a set of twins heads off on a long space journey, while the other twin stays on Earth.

The travelling twin finds that when she returns, her remaining twin has become quite elderly (Figure 12.2.6). While each twin is in constant motion relative to the other, they both measure the other twin ageing more slowly. So why did the twin on the spaceship age more slowly than the twin on Earth?

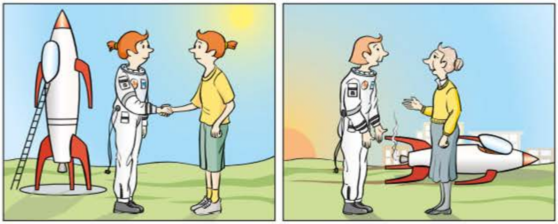


FIGURE 12.2.6 The twin paradox describes the phenomenon where one twin ages less quickly than the other after travelling in a non-inertial frame.

While one twin has spent the entire time in an inertial (non-accelerating) frame of reference, the other twin spent some time in non-inertial frames of reference. The twin that got on the spaceship accelerated away from the Earth, decelerated as she slowed down, then accelerated back towards the Earth, and finally decelerated as she slowed down to land back on the Earth.

If you apply the twin paradox situation to Amaya, Binh and Clare, as Clare watched from her inertial frame of reference, relativity tells you that her view of Amaya and Binh in the non-inertial frame shows them ageing slowly. During this time, Amaya and Binh measure Clare's time passing quickly. As a result, they will see Clare age more rapidly while they are accelerating, and more slowly when they are travelling at constant velocity. Clare sees Amaya and Binh ageing slower and slower as they gain speed, then ageing constantly but slowly as they travel at a constant speed. Amaya and Binh never age rapidly.

But how do you know that it is Amaya and Binh that have accelerated and not Clare, because that is what it would look like for Amaya and Binh looking out of their window at Clare? For the answer to this you need to ask Amaya and Binh if they noticed anything unusual in their frame of reference; for example, did the surface of the water in their bottles tilt at an angle to the horizontal, or did the handles hanging down from the ceiling lean forwards or backwards. If you asked Clare these questions she would say no, while Amaya and Binh would say yes. So it was Amaya and Binh that accelerated and not Clare.

Although it is often called a paradox, there is actually nothing impossible or illogical about this story. Einstein himself pointed out that, due to the Earth's rotation, and therefore centripetal acceleration, a clock on the Earth's equator would run a little more slowly than one at the poles. This has now actually been found to be the case. In fact, in 1971 accurate atomic clocks were flown around the world on commercial flights. When compared with those left behind, the difference of about a quarter of a microsecond was just what Einstein's theory predicted. Now there are many satellites in orbit around the Earth, so the theory has been well and truly tested many times. Indeed, global positioning systems (GPS) must take the relativistic corrections into account to ensure their accuracy.

Explaining high-altitude muons

When certain unstable particles (like pions, which have a precisely known decay rate) are accelerated to almost the speed of light, their life spans are measured to be longer than when the particles are stationary. For example, the mean lifetime of the positive pion, π^+, is 0.000 000 026 033 s (26.033 ns) when it is stationary relative to the atomic clock that is measuring it. However, when it is moving at 99% of the speed of light, its mean lifetime as measured by the stationary atomic clock is 184.54 ns. This means that the moving pion exists seven times longer than a stationary pion.

Even if you ignore the accelerated part of the spaceship twin's motion, you still find that the traveller ages less than the twin that remains on Earth. The reason is because the traveller's journey actually takes place in two frames: the frame of the outbound journey and that of the inbound one. The reduced journey time experienced by the traveller is consistent with the fact that they experience a length contraction of their distance covered relative to that measured by the observer on the Earth.

PHYSICSFILE

Is light slowing down?

Recently there has been publicity given to research that has suggested that the speed of light is slowing down. Some have even suggested that Einstein's theory of relativity itself is under threat. The research, based on analysis of light from very distant quasars, actually suggests that there have been very small changes in what is called the fine structure constant, which is made up of three more basic constants: the speed of light, the charge on an electron and Planck's constant.

Prominent theoretical physicist Professor Paul Davies and others have suggested that if the evidence is correct, then it is probably the speed of light that is changing. If proved correct, no doubt this new data will modify some aspects of relativity, but to suggest that it will overturn relativity is a wild exaggeration.

In the Earth's atmosphere, high-energy cosmic rays interact with the nuclei of oxygen atoms 15 km above the surface of the Earth to create a cascade of high-velocity subatomic particles. One of these particles is a muon, which is unstable. The mean lifetime of a stationary muon, as measured by an atomic clock, is 0.000 002 196 s (2.196 μs). The muons created by cosmic radiation typically travel at 99.97% of the speed of light, so at this speed Newtonian physics would predict that a muon would travel about 659 m:

$$\begin{aligned} s &= v\Delta t \\ &= 0.9997 \times 3.00 \times 10^{8} \times 2.196 \times 10^{-6} \\ &= 658.6\,\text{m (or roughly 659 m)} \end{aligned}$$

After 10 lifetimes, you can expect there to be essentially no muons remaining. So after beginning at a height of 15 km and travelling through a distance of 6.58 km, to a height of about 8.42 km above the surface of the Earth, you would expect that no muons would be detected.

However, muons created by cosmic radiation are actually detected at the surface of the Earth. This means that the fast-moving muons have existed for a much longer period of time than they should have. A muon that strikes the surface would have existed at least 22.8 times its predicted life span as a stationary muon, based on Newtonian physics. Once again, Newtonian physics and Galilean relativity cannot explain this observation.

This surprising observation could only be explained if the mean lifetime of the short-lived particles were extended far beyond their normal mean lifetime. Time dilation provides the explanation for this unusual observation.

The 'normal' mean lifetime of a muon is about 2.2 μs. However, this is the mean lifetime when measured in a stationary frame of reference. Muons travel very fast; in fact a speed as great as $0.999c$ is possible. At this speed, an observer on Earth will measure the lifetime of a muon as far greater than 2.2 μs:

$$\begin{aligned} t &= t_0\gamma \\ &= \frac{t_0}{\sqrt{1-\frac{v^2}{c^2}}} \\ &= \frac{2.2\times10^{-6}}{\sqrt{1-\frac{(0.999c)^2}{c^2}}} \\ &= \frac{2.2\times10^{-6}}{\sqrt{1-0.999^2}} \\ &= 49.21\,\mu\text{s (which is more than 22 times as long as in the stationary frame!)} \end{aligned}$$

An observer on Earth would see the muon's time run much slower. The slower time means that many muons live long enough to reach the Earth's surface.

LENGTH CONTRACTION

Because of the constancy of the speed of light, this effectively means that time appears to have slowed in a moving frame relative to the frame of an observer. Einstein describes how space and time are interrelated, so it follows then that space, and therefore length, is not absolute (Figure 12.2.7).

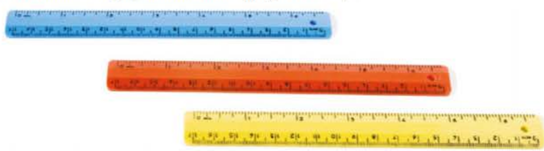


FIGURE 12.2.7 Length is relative to the frame of reference and the direction of motion.

The light clock analysis is appropriate to compare the proper time on the clock in the moving frame of reference and the time measured on a clock in the stationary frame (*with* Clare). The light clock was used as it only depends on light, not some complicated mechanical arrangement that may well include other factors that are altered by relative motion. There was, however, one other condition in this clock analysis—that both Amaya and Clare would agree on the distance, d, between the mirrors. This enabled the two expressions for d to be equated in order to find the relationship between proper time, t_0, and time, t.

The clock was deliberately set up in the spaceship so that this light path, of distance d, was perpendicular (at right angles) to the velocity. Distances in this perpendicular direction are unaffected by motion. Indeed, Einstein showed that while perpendicular distances are unaffected, relative motion affects length only in the direction of travel (Figure 12.2.8).

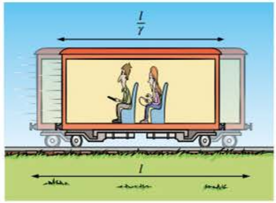


FIGURE 12.2.8 Einstein showed that the length of a moving object is foreshortened by the Lorentz factor, γ. The height and width of the carriage though remain unchanged.

Length contraction

Consider the situation in which Clare is standing on a train platform while Amaya and Binh pass by at a speed v. Both Clare and Binh want to measure the length of the train platform on which Clare is standing. Using a measuring tape, Clare measures the length of the platform (which is at rest according to her) as l_0, and says that Binh and Amaya cover this distance in a time equal to:

$$t = \frac{l_0}{v}$$

Binh observes the platform passing in a time t_0, as he and Amaya move past the station. The relationship between the time in Binh's frame of reference and the time that Clare measures is:

$$t_0 = \frac{t}{\gamma}$$
$$= t\sqrt{1 - \frac{v^2}{c^2}}$$

Substituting the first equation into the equation above gives us:

$$t_0 = \frac{l_0}{v}\sqrt{1 - \frac{v^2}{c^2}}$$

Binh sees the platform moving at a velocity of v relative to him, so he can say that the distance from the start to the end of the platform is:

$$l = vt_0$$

Substituting the previous equation for t_0 into the equation above gives us:

$$l = v \times \frac{l_0}{v}\sqrt{1 - \frac{v^2}{c^2}}$$

This simplifies to:

$$l = l_0\sqrt{1 - \frac{v^2}{c^2}}$$

This is Einstein's **length contraction** equation that incorporates the Lorentz factor. This equation shows that an object with a **proper length** of l_0, when measured in its own frame of reference, will have a shorter length l, parallel to the motion of its moving frame of reference when measured by an observer. The proper length is contracted by a factor of $\frac{1}{\gamma}$. Length contraction can be represented as:

> $$l = l_0\sqrt{1 - \frac{v^2}{c^2}} = \frac{l_0}{\gamma}$$
>
> where
>
> $$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$
>
> l_0 is the proper length, i.e. the length measured at rest, in the stationary frame of reference
>
> l is the length in the moving frame, measured by an observer

> The effects of time dilation and length contraction would not be 'seen' in a real-world sense. If you were to watch the relativistic train used in this example as it approaches, the light reflecting off each end of the train would reach you at different times; as the speed of light is finite, the light travelling from the front of the train would reach you before that reflecting off the back of the train. This causes an optical aberration, distorting the image. As the train is moving towards you it will appear elongated or stretched; while it moves away it will seem to be squished.

Worked example 12.2.2

LENGTH CONTRACTION

A stationary observer on Earth measures a very fast car travelling by at $2.50 \times 10^8\,\text{m s}^{-1}$. When stationary, the car is 3.00 m long. Calculate the length of the car as seen by the stationary observer. Use $c = 3.00 \times 10^8\,\text{m s}^{-1}$.

Thinking	Working
Identify the variables: the length measured by the stationary observer is l, the proper length of the car is l_0, and the velocities are v and the constant c.	$l = ?$ $l_0 = 3.00\,\text{m}$ $v = 2.50 \times 10^8\,\text{m s}^{-1}$ $c = 3.00 \times 10^8\,\text{m s}^{-1}$
Use Einstein's length contraction formula and the Lorentz factor.	$l = \frac{l_0}{\gamma}$ $= l_0\sqrt{1 - \frac{v^2}{c^2}}$
Substitute the values for l_0, v and c into the equation and calculate the answer, l.	$l = 3.00 \times \sqrt{1 - \frac{(2.50 \times 10^8)^2}{(3.00 \times 10^8)^2}}$ $= 3.00 \times 0.553$ $= 1.66\,\text{m}$

Worked example: Try yourself 12.2.2

LENGTH CONTRACTION

A stationary observer on Earth measures a very fast scooter travelling by at $2.98 \times 10^8\,\text{m s}^{-1}$. The stationary observer measures the scooter's length as 45.0 cm. Calculate the proper length of the scooter, measured when the scooter is at rest. Use $c = 3.00 \times 10^8\,\text{m s}^{-1}$.

Length contraction of distance travelled

So far you have been looking at situations in which objects that are in a moving frame of reference are seen as being shorter in the direction of the motion according to an observer that is in a stationary frame of reference. You can also apply length contraction to the distance that a moving object covers as it travels at very high speed.

Recall that no inertial frame of reference is special. Consider Amaya and Binh in their spacecraft. According to them, they are stationary and it is space itself that rushes by at high speed. As space zooms by Amaya and Binh, they are travelling a proper distance of 384 400 km from the Earth to the Moon. This is the proper length as it is measured by a device that is in the same frame of reference as the Earth and the Moon. As Binh looks out of the window, he sees a much shorter distance to travel.

Worked example 12.2.3

LENGTH CONTRACTION OF DISTANCE TRAVELLED

A pilot of a spaceship travelling at 0.997c is travelling from the Earth to the Moon. The proper distance from the Earth to the Moon is 384 400 km. When the pilot looks out of the window, the distance between the Earth and the Moon looks much less than that. Calculate the distance that the pilot measures.

Thinking	Working
Identify the variables: the length seen by the pilot is l, the proper length of the distance is l_0 and the velocity is v.	$l = ?$ $l_0 = 384\,400\text{ km}$ $v = 0.997c\text{ m s}^{-1}$
Use Einstein's length contraction formula and the Lorentz factor.	$l = \frac{l_0}{\gamma}$ $= l_0\sqrt{1 - \frac{v^2}{c^2}}$
Substitute the values for l_0 and v into the equation. Cancel c and calculate the answer, l.	$l = 384\,400 \times \sqrt{1 - \frac{(0.997c)^2}{c^2}}$ $= 384\,400 \times \sqrt{1-(0.997)^2}$ $= 384\,400 \times 0.0774$ $= 29\,800\text{ km}$

Worked example: Try yourself 12.2.3

LENGTH CONTRACTION OF DISTANCE TRAVELLED

A stationary observer on the Earth sees a very fast train approaching a tunnel at a speed of 0.986c. The stationary observer measures the tunnel's length as 123 m long. Calculate the length of the tunnel as measured by the train's driver.

The result from Worked example: Try yourself 12.2.3 leads to an interesting phenomenon. If the proper length of the train is 100 m, then the driver could park the train in the 123 m tunnel with 11.5 m of tunnel extending beyond each end of the train. But when the train is moving at 0.986c, then according to the train driver the train will not fit in the tunnel. There will be approximately 39.8 m of train extending past each end of the tunnel. This phenomenon is illustrated in Figure 12.2.9.

Similarly, a train that is longer than the tunnel will fit completely inside the tunnel if its length was measured by a stationary observer as it was moving past very quickly. In this scenario, the length of the train would be contracted according to the stationary observer (Figure 12.2.10).

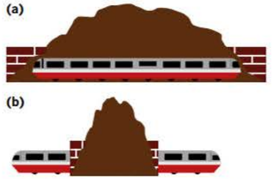


FIGURE 12.2.9 The train both fits in the tunnel and doesn't fit in the tunnel, depending on your frame of reference. In (a), both train and tunnel are stationary; in (b), the tunnel is moving towards the observer in the train.

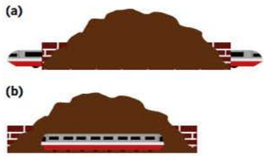


FIGURE 12.2.10 The train both doesn't fit in the tunnel and does fit in the tunnel, also depending on your frame of reference. In (a), the stationary train does not fit in the tunnel; in (b), the fast-moving train is contracted according to a stationary observer.

PROPER TIME AND PROPER LENGTH

The time t_0 and the length l_0 are referred to as the proper time and proper length. They are the quantities measured by the observer, who is in the same frame of reference as the event or the object being measured.

Proper time

The proper time is the time between two events that occur at the same point in space. For example, when a light bulb in the train flashes and Amaya measures the time for the flash to reflect off a mirror and return to her, then she has measured the proper time. This is because the stopwatch remained at the point in space inside the frame of reference where the light originated and where it ended up. Proper time is illustrated in Figure 12.2.11.

It is important that a clock isn't moved from one place to another if you want to measure proper time. This is because, as soon as the clock is in motion, the time for that clock slows slightly.

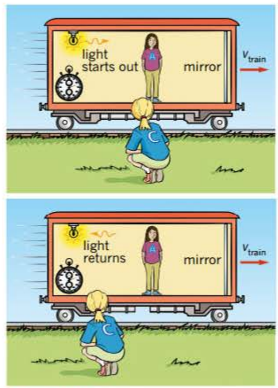


FIGURE 12.2.11 A clock measuring proper time. The clock is positioned at the place where the event started (the light starting out) and is at the same place when the event ends (the light returning).

WS 7.8

Proper length

The proper length is the distance between two points whose positions are measured by an observer at rest with respect to the two points.

Recall the example of Amaya on a train and Clare on the platform observing the passing train. As Amaya reads her measuring tape at either end of the carriage and is at rest with respect to the train, her measurement of the carriage is the proper length. Clare's measurement of the carriage will be of the contracted length.

Clare, on the other hand, measures the length of the platform as the proper length, while Amaya and Binh measure the platform as contracted in length. Remember that length contraction occurs only in the direction of travel, not in any perpendicular direction. To Clare, the carriage will appear shortened, but its width and height (the dimensions of the train perpendicular to the direction of travel) will remain unaltered.

An example of length contraction is shown with a tennis ball in Figure 12.2.12. The length in the direction of the motion is contracted, but the height is not.

FIGURE 12.2.12 As the tennis ball moves faster to the right, its length in this dimension is contracted, but its height and depth remain the same.

FINAL THOUGHTS

Length contraction and time dilation are easy to confuse. When viewed from a frame of reference where objects are seen to be moving, they appear shorter and their clocks tick slower. All lengths and all clocks seem normal when viewed from within their own frame of reference.

12.2 Review

SUMMARY

- The pulses in a light clock in a moving frame of reference have to travel further when observed from a stationary frame.
- Because of the constancy of the speed of light, this effectively means that time appears to have slowed in a moving frame.
- Time in a moving frame seems to flow more slowly according to the equation: $t = t_0\gamma$ where t_0 is the time in the moving frame (proper time), t is the time observed from the stationary frame and γ is the Lorentz factor:
$$\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$$
- Two observers in relative motion both measure time slowing in the other frame of reference; that is, each sees the other ageing more slowly.
- If one observer accelerates in order to return to meet the other, then that accelerated observer will have aged less than the other.
- Time dilation explains how muons can reach the Earth's surface after originating 15 km up in the upper atmosphere, when they should all decay within 7 km of their journey according to classical physics.
- Observations of the lifetimes of subatomic particles that are accelerated to high speeds indicate that they exist for longer than when they are stationary.
- The theory of special relativity states that time and space are related. Motion affects space in the direction of travel.
- A moving object will appear shorter, or appear to travel less distance, by the inverse of the Lorentz factor, γ. Einstein's length contraction equation is given by $l = \frac{l_0}{\gamma}$, where l_0 is the proper length in the stationary frame, l is the contracted length as measured in the moving frame and γ is the Lorentz factor.
- The proper time, t_0, is the time measured by an observer at the same point in an inertial frame of reference.
- The proper length, l_0, is the length measured by an observer at rest with respect to the object being measured.

KEY QUESTIONS

For the following questions, let $c = 3.00 \times 10^8\,\text{m s}^{-1}$ unless stated otherwise.

1 Complete the following sentences by selecting the correct term from those in bold.
In a device called a **light/mechanical/digital** clock, the **speed/oscillation/wavelength** of light is used as a means of measuring **time/mass**, as the speed of light is **unknown/variable/constant** no matter from which inertial frame of reference it is viewed.

2 To what does the term 'proper time', t_0, refer?

3 An observer is standing on a train platform as a very fast train passes by at a speed of $1.75 \times 10^8\,\text{m s}^{-1}$. The observer notices the time on a passenger's phone as the passenger drops the phone to the floor. According to the clock on the phone, it takes 1.05 s to hit the floor. Calculate how much time has passed on the platform's clock during this time.

4 An observer standing on a comet is watching as a satellite approaches at a speed of $2.30 \times 10^8\,\text{m s}^{-1}$. The observer times on her watch that the solar panels on the satellite unfold in 75.0 s. Calculate how much time the observer measures as having passed on the satellite's clock.

5 If Anna measured Ben flying by at 0.5c, how long, in her frame, would it take Ben's clock to tick 1 s?

6 Briefly explain why Einstein said that a clock at Earth's equator should run slightly slower than one at the Earth's poles. Why do we not find this to be a problem?

7 To what does the term 'proper length', l_0, refer?

8 An observer is standing on a train platform as a very fast train passes by at a speed of $1.75 \times 10^8\,\text{m s}^{-1}$. The observer notices that a passenger is holding a metre ruler in line with the direction that the train is moving. Calculate the length of the metre ruler that the stationary observer would measure.

12.3 Momentum and energy

FIGURE 12.3.1 Mass is relative to the frame of reference in which it is measured.

In classical physics and chemistry, the conservation of mass is assumed: the particles that start out in a chemical reaction are still there at the end, and applying a force to an object classically does not change its mass (Figure 12.3.1). This section refers to both energy and momentum being conserved. You will learn about the implications of Einstein's relativistic principles, to show the development of Einstein's most famous equation relating energy, mass and the speed of light, $E = mc^2$. In order to do this, this section will first look at what happens to the momentum of an object as its speed approaches the speed of light.

APPROACHING THE SPEED OF LIGHT

Recall the Lorentz factor that was introduced in Section 12.1:

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}$$

DS 7.1

At low speeds, γ is so close to 1 that the effects of special relativity can be ignored, but γ rapidly increases as the speed, v, comes closer to the speed of light, c. At 99.9% of the speed of light, γ has a value of approximately 22, and so anything moving at that speed, relative to a stationary observer, will appear to have shrunk to $\frac{1}{22}$ of its normal length. As you watch the action inside a spaceship travelling at that speed, events would appear to be going 22 times more slowly than they would if they occurred in a stationary observer's frame of reference.

The closer that the speed of the spaceship gets to the speed of light, the more the Lorentz factor increases towards infinity. It is reasonable to wonder what happens at the speed of light. According to Einstein's equations, the length of the spaceship shrinks to zero and time inside it appears to stop altogether. Einstein took this to mean that it is not possible to reach the speed of light in any real spaceship. However, the difficulties with time and length for the spaceship were not the only reasons Einstein came to this conclusion.

RELATIVISTIC MOMENTUM

If a rocket ship like the one in Figure 12.3.2 is travelling at $0.99c$, why can't it simply turn on its rocket motor and accelerate up to c, or more? A full answer to this question was not given in Einstein's original 1905 paper on relativity. Some years later he showed that as the speed of a spaceship approaches c, its momentum increases, but this is not reflected in a corresponding increase in speed.

Although his analysis is beyond the scope of this course, you can get a feel for his approach if you take some short cuts.

The acceleration, a, of any object is inversely proportional to its mass m, the mass that appears in Newton's second law:

$$F = ma$$

Newton originally stated this law as: a force, F, is equal to the rate of change in momentum p. That is:

$$F = \frac{\Delta p}{\Delta t}$$

A change in momentum is classically defined as the change in the product of the mass, m, and the velocity, v. If you rearrange the above equation and substitute the relationship $\Delta p = m\Delta v$, you get:

$$F\Delta t = m\Delta v$$

Now you can see that time is involved, but at relativistic speeds you know that time is not the constant entity it was once believed to be.

FIGURE 12.3.2 This rocket ship is moving at 0.99c and accelerating, and yet it can never reach a speed of c.

PHYSICSFILE L

Travel at the speed of light

Einstein said that at the speed of light distances shrink to zero and time stops. No ordinary matter can reach c, but light always travels at c. Strange though it may seem, for light there is no time. It appears in one place and disappears in another, having got there in no time (in its own frame of reference, not ours!). When you stay still, you travel through spacetime in the time dimension only. Light does the opposite: all its spacetime travel is through space and none through time.

Imagine that you have a rocket ship accelerating from rest to a high speed as viewed by an observer in a stationary frame of reference. You can say that the change in momentum of the ship will be given by:

$$F\Delta t_0 = m_0\Delta v$$

where t_0 is the time in the ship's frame of reference, and $m_0\Delta v$ is just the classical Newtonian change in momentum.

In the stationary observer's frame, the time is dilated:

$$\Delta t = \gamma\Delta t_0$$

$$\Delta t_0 = \frac{\Delta t}{\gamma}$$

Substituting Δt_0 into the change of momentum equation above:

$$F\frac{\Delta t}{\gamma} = m_0\Delta v$$

$$F\Delta t = \gamma m_0\Delta v$$

That is, the impulse as measured by the stationary observer is equal to the product of the Lorentz factor, γ, and the change in Newtonian momentum. This means that as the spaceship approaches the speed of light, the impulse is multiplied by a factor that grows very rapidly. You can interpret this as meaning that the change in momentum in the stationary observer's frame of reference is equal to:

$$\Delta p = \gamma m_0\Delta v$$
$$= \gamma\Delta p_0$$

If we assume an object starts at zero velocity, the final relativistic momentum becomes:

$$p_v = \gamma m_0 v = \frac{m_0 v}{\sqrt{\left(1-\frac{v^2}{c^2}\right)}}$$

$$p = \gamma p_0$$

where

p_0 is the momentum m_0v, as you would define it in classical mechanics, i.e. where m_0 is the mass measured at rest, in the stationary frame of reference

p_v is the relativistic momentum in the moving frame measured by an observer

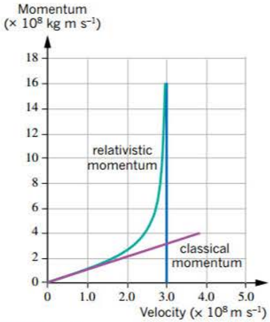


FIGURE 12.3.3 The relationship between classical momentum and velocity, and the relationship between relativistic momentum and velocity, for a 1 kg mass.

PHYSICSFILE N

Getting beyond *c*

A rocket ship travelling at 0.99*c* (speed *U*) fires a small ship at 0.02*c* relative to it (speed *v*). Isn't the small ship moving at 1.01*c*? No! First you need to be careful to specify in which frame of reference you are measuring the speeds. The rocket ship has speed *U* in your frame, while the small ship has speed *v* in the frame of the rocket ship. (Capital letters for your frame, small for the rocket frame.) Because of length contraction you measure the small ship fired at much less than 0.02*c*. Einstein showed that in these cases the speed (*V*) of the small ship in your frame is given by:

$$V = \frac{U+v}{1+\frac{Uv}{c^2}}$$

You can use this expression to show that you measure the small ship travelling at 0.9904*c*.

If velocity, v, is needed when the mass and relativistic increase in momentum is known, the formula $p_v = \gamma m_0 v$ can be rearranged to give the following:

$$v = \frac{p_v}{m_0\sqrt{\left(1+\frac{p_v^2}{m_0^2c^2}\right)}}$$

The momentum increases very rapidly as a spaceship approaches the speed of light. You might argue that this is expected—after all, momentum is a function of velocity. If you graph the relativistic momentum, p_v, against the velocity, v, and on the same graph show the classical momentum, you can see that the relativistic momentum increases at a rate far greater than it would if it were due to the increase in velocity alone (Figure 12.3.3).

This result can be interpreted by thinking of the mass as a quantity that also increases at high speeds. Thus there is a relationship between the rest mass, m, which is the mass measured while the object is at rest in the frame of reference, and the relativistic mass, γm_0, which is the mass measured as the object is moving relative to the observer.

As the Lorentz factor increases with the increase in the velocity, then the relativistic mass also increases. Einstein was never happy with the term 'relativistic mass', and preferred that people only spoke of the relativistic momentum of an object.

Notice too how the classical treatment allows the object to have a speed greater than the speed of light, but the relativistic treatment causes the mass to become very large so that the speed of light is never actually reached.

Now return to the example of the rocket ship that is attempting to increase its velocity to the speed of light. With the increase in the relativistic mass of the rocket ship, it becomes harder for the force of the engines to cause a change in velocity. The closer the rocket ship approaches c, the greater the amount of impulse that is required to accelerate the ship to the speed of light. In fact, as the velocity approaches c, the relativistic mass, γm_0, approaches infinity. You can now see why your rocket ship cannot reach the speed of light.

Worked example 12.3.1 illustrates this point. Notice that the result in part (b) shows that if you double the impulse required to get the rocket ship to 0.99c, then you will only add 0.007c to your top speed. When you've completed Worked example: Try yourself 12.3.1, consider the change in velocity achieved by tripling the impulse.

Worked example 12.3.1

RELATIVISTIC MOMENTUM

a Calculate the momentum, as measured by a stationary observer, provided to a rocket ship with a rest mass of 1000 kg, as it goes from rest up to a speed of 0.990c. Use $c = 3.00 \times 10^8\,\text{m s}^{-1}$.

Thinking	Working
Identify the variables: the rest mass is m_0, and the velocity of the rocket ship is v.	$p_v = ?$ $m_0 = 1000\,\text{kg}$ $v = 0.990 \times 3.00 \times 10^8$
Use the relativistic momentum formula.	$p = \gamma m_0 v$
Substitute the values for m_0 and v into the equation and calculate the answer, p_v.	$p_v = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} m_0 v$ $= \frac{1}{\sqrt{1-\frac{0.990^2c^2}{c^2}}} \times 1000 \times 0.990 \times 3.00 \times 10^8$ $= 2.11 \times 10^{12}\,\text{kg m s}^{-1}$

b If twice the relativistic momentum from part a is applied to the stationary rocket ship, calculate the new final speed of the rocket ship in terms of *c*.	
Thinking	**Working**
Identify the variables: the rest mass is m_0, and the relativistic momentum of the rocket ship is *p*.	$p_v = 2 \times 2.11 \times 10^{12}$ $= 4.21 \times 10^{12}\,\text{kg m s}^{-1}$ $m_0 = 1000\,\text{kg}$ $v = ?$
Use the relativistic momentum formula, rearranged.	$p = \gamma m_0 v$ $= \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} m_0 v$ $v = \frac{p}{m_0\sqrt{\left(1+\frac{p^2}{m_0^2c^2}\right)}}$
Substitute the values for m_0 and *p* into the rearranged equation and calculate the answer, *v*.	$v = \frac{p}{m_0\sqrt{\left(1+\frac{p^2}{m_0^2c^2}\right)}}$ $= \frac{4.21\times10^{12}}{(1000)\sqrt{1+\frac{(4.21\times10^{12})^2}{1000^2(3.00\times10^8)^2}}}$ $= \frac{4.21\times10^{12}}{1000\times14.07}$ $= 2.99 \times 10^8\,\text{m s}^{-1}$ $= 0.997c$

Worked example: Try yourself 12.3.1

RELATIVISTIC MOMENTUM

a Calculate the momentum, as measured by a stationary observer, provided to an electron with a rest mass of 9.11×10^{-31} kg, as it goes from rest to a speed of 0.985*c*. Use $c = 3.00 \times 10^8\,\text{m s}^{-1}$.

b If three times the relativistic momentum from part (a) is applied to the electron, calculate the new final speed of the electron in terms of *c*.

EINSTEIN'S FAMOUS MASS–ENERGY EQUATION

As the momentum of an object increases, so does its kinetic energy. The classical relationship between the two can be written as:

$$\begin{aligned} K &= \tfrac{1}{2}mv^2 \\ &= \tfrac{1}{2}(mv) \times v \\ &= \tfrac{1}{2}pv \end{aligned}$$

This form of the equation shows that the kinetic energy of an object is related to an object's momentum as well as its velocity.

Einstein showed, however, that the classical expression for kinetic energy was not correct at high speeds. The mathematics involved is beyond the scope of this course, but Einstein, working from the expression for relativistic momentum and the usual assumptions about work, forces and energy, was able to show that the kinetic energy of an object was given by the expression:

$$K = (\gamma - 1)mc^2$$

Although it is not very obvious from this expression, if the velocity (which is hidden in the γ term) is small, this expression actually reduces to the classical equation for K of $\frac{1}{2}mv^2$. A small velocity in this context means small in comparison to c. But even for speeds up to $0.10c$, the classical expression is accurate to better than ±1%.

Einstein's expression can be expanded to:

$$K = \gamma mc^2 - mc^2$$

This kinetic energy equation, in turn, can be rearranged as:

$$\gamma mc^2 = K + mc^2$$

Einstein interpreted the left-hand side of this expression as being an expression for the total energy of the object:

$$E_{total} = \gamma mc^2$$

The right-hand side appeared to imply that there were two parts to the total energy: the kinetic energy, K, and another term that only depended on the rest mass, m. The second term, mc^2, he referred to as the rest energy of the object, as it does not depend on the speed of the object. This appeared to imply that somehow there was energy associated with mass (Figure 12.3.4). This would be an astounding proposition to a classical physicist, but as you have seen, in relativity, mass increases as you add kinetic energy to an object. The conservation of energy relationship is therefore:

$$E_{total} = K + E_{rest}$$

$$E_{total} = \gamma mc^2$$

where

γm is the relativistic mass (kg)
c is the speed of light ($m\,s^{-1}$)
E_{total} is the total energy (J)

You will have seen part of this equation before:

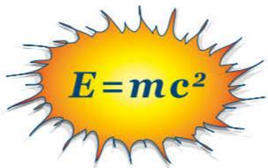


FIGURE 12.3.4 Einstein's famous equation.

This equation tells you that mass and energy are totally interrelated. In a sense, you can say that mass has energy, and energy has mass.

Converting mass to energy or energy to mass

Nuclear reactions involve vastly more energy per atom than chemical ones (Figure 12.3.5). When a uranium atom splits into two fission fragments, about 200 million electron volts of energy are released. By comparison, most chemical reactions involve just a few electron volts.

FIGURE 12.3.5 In a nuclear bomb, a few grams of mass are converted into energy. As the uranium undergoes fission, it releases the equivalent of hundreds of gigajoules (10^{12} J) of energy. Millions of tonnes of a chemical explosive (TNT) would be required to produce this much explosive energy.

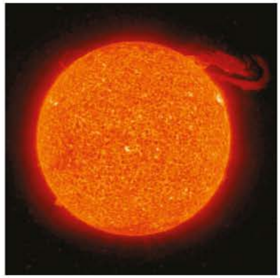


FIGURE 12.3.6 Nuclear fusion in the Sun results in about 4 million tonnes of mass being converted into energy every second, which is radiated from the Sun.

In the fission of uranium, it is possible to find the original mass of the uranium nucleus and the fission fragments accurately enough to determine the mass defect (change in mass). This difference in mass agrees exactly with the prediction of Einstein's famous equation.

Likewise, nuclear fusion reactions deep inside the Sun release the huge amounts of energy that stream from the Sun, resulting in a conversion of about 4 million tonnes of mass into energy every second (Figure 12.3.6).

Nuclear fusion

Nuclear fusion occurs when two light nuclei are combined to form a larger nucleus (Figure 12.3.7).

As in the cases of radioactive decay and nuclear fission, the mass of the reactants is slightly greater than the mass of the products when the nuclei combine during fusion.

The energy created by this missing mass—known as the **mass defect**—can again be determined from:

$$\Delta E = \Delta mc^2$$

where

ΔE is the change in energy (J)

Δm is the mass defect (kg)

c is the speed of light ($3.0 \times 10^8\,\mathrm{m\,s^{-1}}$).

Nuclear fusion is a very difficult process to recreate in a laboratory. The main problem is that nuclei are positively charged, and thus repel one another.

Slow-moving nuclei with relatively small amounts of kinetic energy will not be able to get close enough for the strong nuclear force to come into effect, and so fusion will not happen. Only if nuclei have enough kinetic energy to overcome the repulsive force can they come close enough for the strong nuclear force to start acting. If this happens, fusion will occur (Figure 12.3.8 on page 328). The process of fusion will be discussed in greater detail in Chapter 16.

Typically, temperatures of the order of hundreds of millions of degrees are required. These are exactly the conditions that are present inside the Sun.

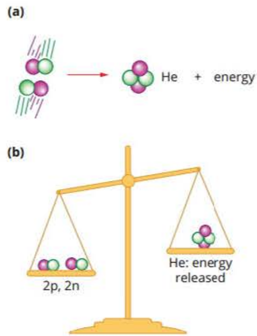


FIGURE 12.3.7 (a) When two isotopes of hydrogen fuse to form a helium nucleus, energy is released. (b) The binding energy of the nucleus appears as a loss in mass, Δm, which can be calculated using $\Delta E = \Delta mc^2$.

GO TO ➤ Section 16.3 page 437

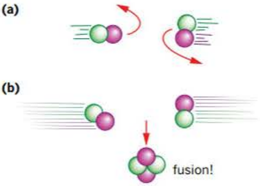


FIGURE 12.3.8 (a) Slow-moving nuclei do not have enough energy to fuse together. The electrostatic forces cause them to be repelled from each other. (b) If the nuclei have sufficient kinetic energy, then they will overcome the repulsive forces and move close enough together for the strong nuclear force to come into effect. At this point, fusion will occur and energy will be released.

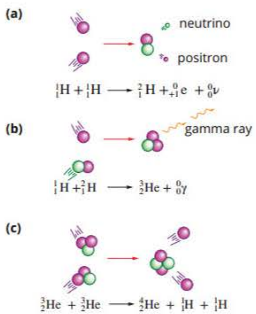


FIGURE 12.3.9 The three main fusion reactions taking place inside the Sun.

Fusion in the Sun and similar stars

In the Sun, many different fusion reactions are taking place. The main reaction is the fusion of hydrogen nuclei to form helium. Each second, about 657 million tonnes of hydrogen and hydrogen isotopes fuse to form about 653 million tonnes of helium. Each second, a mass defect of 4 million tonnes results from these fusion reactions. The amount of energy released is enormous, and can be found by using the equation $\Delta E = \Delta mc^2$. A tiny proportion of this energy reaches Earth and sustains life as we know it.

The sequence of fusion reactions shown in Figure 12.3.9 has been occurring inside the Sun for the past 5 billion years and is expected to last for another 5 billion years or so. Hydrogen nuclei are fused together and, after several steps, a helium nucleus is formed. This process releases about 25 MeV of energy.

The Sun is a second- or third-generation star. It was formed from the remnants of other stars that exploded much earlier in the history of the galaxy. As this giant gas cloud contracted under the effect of its own gravity, the pressure and temperature at the core reached extreme values, sufficient to sustain these fusion reactions.

Worked example 12.3.2

FUSION

Consider the fusion reaction shown below. A proton (a hydrogen nucleus) fuses with a deuterium nucleus (a hydrogen nucleus with one neutron) in the Sun. A helium nuclide is formed and a γ-ray is released. 20 MeV of energy is released during this process.

$${}^1_1p + {}^2_1H \rightarrow {}^3_2He + \gamma$$

a How much energy is released in joules?

Thinking	Working
$1\,eV = 1.602 \times 10^{-19}\,J$	$20\,MeV = 20 \times 10^6 \times 1.602 \times 10^{-19}$ $= 3.2 \times 10^{-12}\,J$

b Calculate the mass defect for this reaction.

Thinking	Working
Use $\Delta E = \Delta mc^2$.	$\Delta m = \frac{\Delta E}{c^2}$ $= \frac{3.2 \times 10^{-12}}{(3.0 \times 10^8)^2}$ $= 3.6 \times 10^{-29}\,kg$

Worked example: Try yourself 12.3.2

FUSION

A further fusion reaction in the Sun fuses two helium nuclides. A helium nucleus and two protons are formed, and 30 MeV of energy is released.

$${}^3_2He + {}^3_2He \rightarrow {}^4_2He + 2{}^1_1H$$

a How much energy is released in joules?

b Calculate the mass defect for this reaction.

Electron–positron annihilation

The positron is the antiparticle of the electron. It has the same mass as an electron but has a positive charge that is equal in magnitude to the charge on an electron. Positrons are produced when proton-rich radioactive nuclei decay (through the weak interaction) as the result of a proton decaying to a neutron.

If a positron collides with an electron at low energies, annihilation occurs and gamma rays are produced. This is shown in the following equation:

$$^{0}_{-1}\mathrm{e} + {}^{0}_{+1}\mathrm{e} \rightarrow \gamma + \gamma$$

where $^{0}_{-1}\mathrm{e}$ is an electron and $^{0}_{+1}\mathrm{e}$ is a positron.

In this annihilation event, charge is conserved as the total charge prior to the event is zero (+1 for the positron and −1 for the electron) and the total charge after the event is also zero as gamma rays do not have charge. Momentum is also conserved, as the total momentum of the positron and electron prior to the event will equal the total momentum of the gamma rays after the event.

Initially, mass appears not to be conserved in this process, as the positron and electron both have mass, but the gamma rays do not.

Einstein's equation shows us that the mass of an object or objects is a measure of its energy content. Thus, the mass of the electron and positron are converted into energy, or gamma rays.

Taking the rest mass of an electron (and positron) to be 9.1×10^{-31} kg, the energy produced during an electron–positron annihilation event can be determined.

$$\begin{aligned} E &= mc^2 \\ &= 2 \times (9.1 \times 10^{-31}) \times (3.0 \times 10^{8})^2 \\ &= 1.638 \times 10^{-12}\,\mathrm{J} \\ &= 1.02\,\mathrm{MeV} \end{aligned}$$

PHYSICSFILE N

Feynman diagrams

Richard Feynman (1918–88) was an American theoretical physicist. He is known for his work on quantum electrodynamics, as well as for his contribution to the Manhattan Project (the development of the atomic bomb) and for being a member of the Rogers Commission that investigated the *Challenger* Space Shuttle disaster.

Feynman developed what has now become known as Feynman diagrams. These diagrams can be used to show the behaviour of subatomic particles and their interactions, such as the positron and electron annihilation discussed in this section and shown in Figure 12.3.10, through a simple pictorial representation.

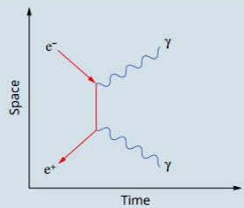


FIGURE 12.3.10 An electron (e⁻) and a positron (e⁺) collide. Antiparticles like the positron are always shown as travelling backwards in time.

Fuel combustion

Einstein's equation, $E = mc^2$, does not only apply to nuclear reactions, but also to other mass-energy scenarios. In many situations, the mass difference is so small that the change in mass goes unnoticed.

For example, the heat of combustion of coal is approximately $34\,\mathrm{MJ\,kg^{-1}}$. In other words, 3.4×10^{7} J is produced for every kg of coal that is burnt. If 1 kg of coal is burnt, what is the difference in mass before and after the combustion process?

$$\begin{aligned} E &= 3.4 \times 10^{7}\,\mathrm{J} \text{ for } 1\,\mathrm{kg} \text{ of coal} \\ \Delta E &= \Delta mc^2 \\ 3.4 \times 10^{7} &= \Delta m \times (3.0 \times 10^{8})^2 \\ \Delta m &= 3.8 \times 10^{-10}\,\mathrm{kg} \end{aligned}$$

Considered as a percentage of the original mass of coal, this is only $\frac{3.8\times10^{-10}}{1.0} \times 100 = 3.8 \times 10^{-8}\,\%$. This is a very small percentage and therefore the difference in the mass of coal before combustion and mass of the products after combustion is not significant.

12.3 Review

SUMMARY

- Relativistic momentum includes the Lorentz factor, γ, and hence, as more impulse is added, the mass seems to increase towards infinity as the speed gets closer, but never equal, to c. The relativistic momentum equation is:
$$p_v = \gamma m_0 v = \gamma p_0$$
- A term called relativistic mass, γm, may be used to indicate the mass of an object which is moving.
- Einstein found that the total energy of an object was given by:
$$E_{total} = K + E_{rest} = \gamma mc^2$$
- The kinetic energy of an object is given by:
$$K = (\gamma - 1)mc^2$$
- The rest energy, which is the energy associated with the rest mass of an object, is given by:
$$E_{rest} = mc^2$$
- Mass and energy are seen as different forms of the same thing. This means that mass, m, can be converted into energy, and energy can be converted into mass.
- Nuclear fission and fusion reactions result in a mass defect (change). It is this difference in mass that is converted to the energy released in nuclear reactions. This mass defect is related to the energy produced according to $\Delta E = \Delta mc^2$.
- Nuclear fusion is the combining of light nuclei to form heavier nuclei. Extremely high temperatures are required for fusion to occur. This is the process occurring in stars.
- Hydrogen nuclei fuse to form deuterium. Further fusions result in the formation of isotopes of helium.
- Particle–antiparticle interactions cause mass to convert into energy. For example, the mass of a positron–electron interaction is converted into gamma rays.
- The combustion of fuel such as the burning of coal is another example of converting mass into energy.

KEY QUESTIONS

1 Calculate the relativistic momentum of the *Rosetta* spacecraft as observed by the scientists at the European Space Agency. *Rosetta*'s rest mass was 1230 kg and its speed was $775\,\text{m s}^{-1}$.

2 Calculate the relativistic momentum of a carbon-12 nucleus in a linear accelerator if its rest mass is $1.99264824 \times 10^{-26}$ kg and it is travelling at 0.850c.

3 Calculate the relativistic momentum of another carbon-12 nucleus in the solar wind if its rest mass is $1.99264824 \times 10^{-26}$ kg and it is travelling at a speed of $800\,\text{m s}^{-1}$.

The following information relates to questions 4–6.

A very fast arrow has a rest mass of 12.3 g and a speed of 0.750c.

4 Calculate the relativistic kinetic energy of the arrow.

5 Calculate the kinetic energy of the arrow according to the classical equation.

6 What accounts for the difference between the kinetic energy of the arrow in the relativistic calculation and the kinetic energy in the classical calculation?

A the difference in the arrow's velocity in the two calculations

B the difference in the arrow's momentum in the two calculations

C the difference in the arrow's rest mass in the two calculations

D the presence of the Lorentz factor in the relativistic calculation

7 Calculate the total energy of a very fast vintage Vespa scooter if its rest mass is 210 kg and it is travelling at a speed of $2.55 \times 10^8\,\text{m s}^{-1}$.

8 Calculate the energy produced by the Sun in one day if 4.00 million tonnes of matter is converted into energy every second.

Chapter review

12

KEY TERMS

aether
classical physics
frame of reference
inertial frame of reference
length contraction
Lorentz factor
mass defect
medium
paradox
postulate
proper length
proper time
simultaneous
spacetime
time dilation

REVIEW QUESTIONS

1 Prove that for an object travelling at any possible velocity, the value of the term below must be less than 1.

$$\sqrt{1-\frac{v^2}{c^2}}$$

2 One of the fastest objects ever made on Earth was the Galileo Probe which, as a result of Jupiter's huge gravity, entered its atmosphere in 1995 at a speed of nearly 50 000 m s^{-1}. Give an estimate of the Lorentz factor for the probe to nine decimal places. (You may use the expression $\gamma \approx 1+\frac{v^2}{2c^2}$.)

3 In 1905 Einstein put forward two postulates. Which two of the following best summarise them?

A All observers will find the speed of light to be the same.
B In the absence of a force, motion continues with constant velocity.
C There is no way to detect an absolute zero of velocity.
D Absolute velocity can only be measured relative to the aether.

4 Whereabouts on the Earth's surface are we closest to an inertial frame of reference?

5 Which of the following is closest to Einstein's first postulate?

A Light always travels at $3 \times 10^8 \text{m s}^{-1}$.
B There is no way to tell how fast you are going unless you can see what's around you.
C Velocities can only be measured relative to something else.
D Absolute velocity is that measured with respect to the Sun.

6 Very briefly explain why Einstein said that we must use four-dimensional spacetime to describe events that occur in situations where high speeds and large distances are involved.

7 Imagine that Amaya is at the front end of a train carriage moving forwards at 10.0 m s^{-1}. She shines a laser towards Binh, who is at the other end of the carriage. Clare is watching this from the side of the track. At what velocity does Clare measure the light travelling?

8 Which one or more of the following conditions is sufficient to ensure that we will measure the proper time between two events? We must:

A be in the same frame of reference
B be in a frame of reference which is travelling at the same velocity
C be stationary
D not be accelerating with respect to the frame of the two events

9 Spaceships A and B leave the Earth and travel towards Vega, both at a speed of 0.9*c*. Observer C back on Earth measures the crews of A and B moving in 'slow motion'. Describe how the crew of A measures the change in time for both the crew of B and the stationary observers back on Earth (C).

A B will appear normal, C will be sped up.
B B will appear normal, C will be slowed down.
C B will appear slowed down, C will be normal.
D B will appear sped up, C will be slowed down.
E None of these.

10 If you were riding in a very smooth, quiet train with the blinds drawn, how could you tell the difference between the train (i) being stopped in the station, (ii) accelerating away from the station, (iii) travelling at a constant speed?

11 You are in a spaceship travelling at very high speed past a new colony on Mars. Do you notice time going slowly for you; for example, do you find your heart rate is slower than normal? Do the clocks on Mars appear to be moving normally? Explain your answers.

12 A passenger is sitting in a very fast jet plane looking out of the window at a clock placed on top of a mountain. The passenger, using the mountain's clock, notes that it takes a goat 20.0 s to run along a rocky slope. If the plane is flying at a speed of $2.00 \times 10^8 \text{m s}^{-1}$, calculate how much time has passed on the passenger's clock.

13 A spectator is standing next to the pool clock and watching as a swimmer races at a speed of $2.25 \times 10^8\,m\,s^{-1}$. The spectator times on the pool clock that the swimmer completes one stroke in 1.50 s.

a Calculate how much time the spectator measures on the swimmer's wristwatch.

b Calculate how much time the swimmer measures on the pool clock, during which time her own wristwatch shows that 1.50 s have passed.

14 **a** At what speed would a rocket ship be travelling if it is observed to be half its normal length? Give your answer to four significant figures.

b The same rocket ship is then observed to accelerate to a certain speed so that its length halved again. Did that mean that it doubled its speed? To what speed did it accelerate?

15 Binh and Amaya are playing table tennis in their spaceship. They rush past Clare in her space station at a relative speed of $240\,000\,km\,s^{-1}$. Binh says that after he hits the ball it returns to his bat after 1.00 s. Their table is 3.00 m long in the direction of their spaceship's motion and is 1.00 m high.

a Calculate the time between hits, as measured by Clare.

b Calculate the length and height of the table, as measured by Clare.

16 Star Xquar is at a distance of 5 light-years from the Earth. Space adventurer Raqu heads from the Earth towards Xquar at a speed of 0.9*c*.

a For those watching from the Earth, how long will it take for Raqu to reach Xquar?

b From Raqu's point of view, how long will it take her to reach Xquar?

c Although Xquar was 5 light-years from the Earth and Raqu travelled at 0.9*c*, her journey took much less time than might be expected from these figures. Explain why this is.

17 A space shuttle travels at close to $8000\,m\,s^{-1}$. Imagine that as it travels east–west it is to take a photograph of Australia, which is close to 4000 km wide. Because of its speed, the space camera will see everything on Earth as slightly contracted.

a About how much less than 4000 km wide will Australia appear to be in this photograph?

b Will the north–south dimension of Australia be smaller as well?

18 Imagine that as we watch a traveller from Earth to the star Vega travelling at 99.5% of the speed of light, we find that their clocks slow down by a factor of about 10 times.

a Explain how this factor of 10 was arrived at.

b Does this mean that the traveller experiences this slowing down of time?

c Vega is about 25 light-years from Earth, so in our frame of reference it takes light from Vega 25 years to reach us. How long will it take our space traveller to reach Vega?

d How long will the traveller find that it takes to travel to Vega?

e Does your answer to part d imply that they were able to get to Vega in less time than light? Explain your answer.

19 Muons are high-speed particles that are created some 15 km above the Earth's surface. Classical physics dictates that due to their short life spans, muons should not ever reach the Earth's surface even though they travel at incredible speeds (approx. 0.992*c*). However, they do. Explain how this is possible, referring to each of the frames of reference of an observer on Earth and the muon itself.

20 If a spaceship is travelling at 99% of the speed of light, which of the following best explains why it can't simply turn on its engine and accelerate through and beyond the speed of light, *c*, as the increase in momentum should be equal to the impulse applied?

A The law that impulse equals the change in momentum does not apply at speeds close to *c*.

B While the momentum increases with the impulse, it is the mass rather than the speed that is getting greater.

C The spaceship does actually exceed *c*, but it doesn't appear to from another frame of reference because of the length contraction of the distance it covers.

D Given enough impulse the spaceship could exceed *c*, but no real spaceship could carry enough fuel.

21 Find the speed of a proton if it has kinetic energy equal to its rest-mass energy.

22 Find the relativistic mass of the proton described in the question above, if $m_p = 1.67 \times 10^{-27}\,kg$.

23 Calculate the relativistic kinetic energy of a bus with a rest mass of 5.30 tonnes and travelling at a speed of 0.960*c*.

24 A student standing by the side of a road sees a very fast MG sports car driving past. The driver times on his car's clock that it takes 5.50 s for the student to pick up her bag. If the MG is moving at a speed of $2.75 \times 10^8\,m\,s^{-1}$, calculate how much time the driver sees has passed on the student's watch as she picks up the bag.

25 Anna's light clock has a height of 1 m between the mirrors, and relative to Chloe her spaceship is travelling at 90% of the speed of light ($c = 3.0 \times 10^8\,\text{m s}^{-1}$). One tick is the time for light to go from one mirror to the other.

a How far does the light flash travel in Anna's frame of reference in one tick, t_A?

b What is the tick time, t_A, for the clock in Anna's frame?

As the light takes a zigzag path in her frame, Chloe sees the clock ticking at a slower rate, t_c.

c In terms of c and t_c what is the length of the zigzag path that the flash travels in one tick in Chloe's frame?

d What is the tick time of the clock in Chloe's frame?

e What is the ratio of Chloe's tick to Anna's tick?

26 A muon created at an altitude of 15.0 km above the Earth is moving at a speed of 0.992 times the speed of light. The mean lifetime of a muon at rest is 2.20×10^{-6} s.

a Calculate the lifetime of the moving muon as timed by a stationary observer.

b Using classical physics equations and the results from part a, calculate the non-relativistic distance and the relativistic distance travelled by the moving muon during one lifetime.

27 A stationary observer on a platform measures the time with a stopwatch for a train carriage, moving at 0.99c, to pass her by. What time has she measured, t or t_0? Explain.

28 A jet plane zooms past an observer standing on the ground at a speed of $660\,\text{m s}^{-1}$. If the length of the jet is 23.5 m when it is parked on the tarmac, how long does the jet appear to be to the observer?

29 Natural gas in Australia consists mainly of methane. If the heat of combustion of methane is approximately $50\,\text{MJ kg}^{-1}$, and 5 kg of natural gas is burnt, what is the difference in mass before and after the combustion process?

30 After completing the activity on page 300, reflect on the inquiry question: How does the behaviour of light affect concepts of time, space and matter?

MODULE 7 • REVIEW

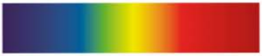
REVIEW QUESTIONS

WS 7.8 WS 7.9 MR 7

The nature of light

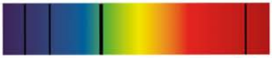
Multiple choice

1 Which of the following images show emission spectra?

A


B


C


D


E


2 A gamma ray is composed of an electric and magnetic field which are:

A perpendicular to each other and to the direction of motion
B perpendicular to each other and parallel to the direction of motion
C parallel to each other and perpendicular to the direction of motion
D parallel to each other and to the direction of motion

3 The light of which of the options A–D is:

a generated by the electrons in individual atoms as they drop one or more energy levels?
b generated by random thermal motion of atoms in the material?
c generated as electrons fall from the conduction band to the valence band in a semiconductor?
d coherent

A an incandescent light bulb
B a sodium vapour lamp
C a laser
D a light-emitting diode

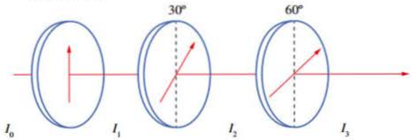
4 Which of the following has the longest wavelength?

A electron, $m = 9.109 \times 10^{-31}$ kg, $v = 7.5 \times 10^{6}\,\text{m s}^{-1}$
B blue light, $\lambda = 470$ nm
C X-ray, $f = 5 \times 10^{17}$ Hz
D proton, momentum $= 1.7 \times 10^{-21}\,\text{kg m s}^{-1}$

5 Radiation of wavelength 3.0×10^{-5} m is detected. What type of light is this?

A X-ray
B ultraviolet
C visible
D infrared

6 Unpolarised light, with intensity I_0, is passed through three polarising filters shown below. The angles given are in relation to the first axis of polarisation. What is the intensity of light, I_3, after it has passed through the third filter?



A $\frac{1}{4} I_0$
B $\frac{1}{8} I_0$
C $\frac{9}{16} I_0$
D $\frac{9}{32} I_0$

7 Laser light with a wavelength of 555 nm is directed through a pair of slits separated by 5.00 µm. The slits are 1.0 m from the screen. What is the angle to the third bright fringe from the centre of the interference pattern?

A 19.0°
B 19.5°
C 20.0°
D 20.5°

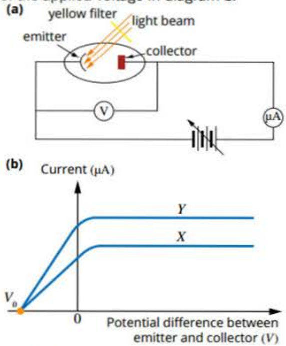
8 The following data was taken of the intensity of light after a polariser was added to the set-up. Each polariser was rotated by the same angle from the previous polarisation axis. What is this angle?

Polariser	$\frac{I_n}{I_0}$
I_0	1.0
I_1	0.97
I_2	0.94
I_3	0.91
I_4	0.88
I_5	0.86

A 5°
B 10°
C 15°
D 20°

The following information relates to questions 9–11

Light passing through a yellow filter is incident on the cathode in a photoelectric-effect experiment as shown in diagram a. The reverse current in the circuit can be altered using a variable voltage. At the stopping voltage, V_0, the photocurrent is zero. The current in the circuit is plotted as a function of the applied voltage in diagram b.



9 Which of the following changes would result in an increase in the size of V_0?

A replacing the yellow filter with a red filter
B replacing the yellow filter with a blue filter
C increasing the intensity of the yellow light

10 Which one of the following options best describes why there is zero current in the circuit when the applied voltage equals the stopping voltage?

A The threshold frequency of the emitter increases to a value higher than the frequency of yellow light.
B The work function of the emitter is increased to a value higher than the energy of a photon of yellow light.
C The emitted photoelectrons do not have enough kinetic energy to reach the collector.

11 Which of the following descriptions of the graphs X and Y in diagram b are correct?

A Both graphs are produced by yellow light of different intensities.
B Graph X is produced by yellow light, while graph Y is produced by blue light.
C Each graph is produced by light of a different colour and different intensity.

12 Which one or more of the following phenomena can be modelled by a pure wave model of light?

A the photoelectric effect
B refraction
C the double-slit interference of light
D reflection
E diffraction

13 A star is losing mass at the rate of $4 \times 10^9 \text{ kg s}^{-1}$. How much energy in total radiation is being produced per second?

A 3.6×10^{26} J
B 3.7×10^{26} J
C 3.8×10^{26} J
D 3.9×10^{26} J

14 One of the fastest human-made objects was the Galileo Probe which, as a result of Jupiter's huge gravity, entered its atmosphere in 1995 at a speed of nearly $50\,000 \text{ m s}^{-1}$. Which of the following is the best estimate of the Lorentz factor for the probe?

A Less than 1
B 1.00000000
C 1.00000001
D 1.1

15 In 1905 Einstein put forward two postulates. Which two of the following best summarise them?

A All observers will find the speed of light to be the same.
B In the absence of a force, motion continues with constant velocity.
C There is no way to detect an absolute zero of velocity.
D Absolute velocity can only be measured relative to the aether.

16 You are in interstellar space and know that your velocity relative to Earth is $4 \times 10^6 \text{ m s}^{-1}$ away from it. You then notice another spacecraft with a velocity, towards you, of $4 \times 10^5 \text{ m s}^{-1}$. Which one or more of the following best describes the velocity of the other craft relative to Earth?

A Away from Earth at $3.6 \times 10^6 \text{ m s}^{-1}$
B Towards Earth at $3.6 \times 10^6 \text{ m s}^{-1}$
C Away from Earth at $4.4 \times 10^6 \text{ m s}^{-1}$
D Towards Earth at $4.4 \times 10^6 \text{ m s}^{-1}$

17 Which one of the following best represents the basis of Einstein's considerations, which eventually led to the theory of special relativity?

A The results of numerous experiments to determine the speed of light.
B The work of Isaac Newton and Michael Faraday.
C His consideration of the consequences of accepting the implications of Maxwell's equations.
D His own experiments in electromagnetism.

18 According to a speed versus distance travelled graph, which of the following is true?

A At the maximum speed, the distance travelled is the largest.
B Velocity and distance travelled are directly proportional variables.
C At values close to the speed of light, the distance travelled is near to zero.
D None of the above.

19 If you are standing on Earth and observe a speeding rocket ship, what do you notice about its dimensions? Select from the following:

A its length (in the direction of travel) is shorter than normal

B its length (in the direction of travel) is longer than normal

C its height (at right angles to the direction of travel) is shorter than normal

D its width (at right angles to the direction of travel) is shorter than normal

20 What is the result of a large increase in the impulse provided to an object moving at a speed near that of light?

A a large change in the velocity of the object

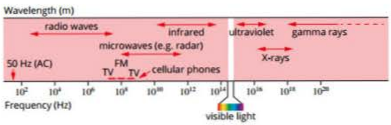
B a proportional increase in the velocity of the object

C a very small increase in the velocity of the object

D no change in the velocity of the object at all

Short answer

21



a At what speed do the waves shown in the diagram travel in air?

b What is the wavelength of waves with a frequency of 10^{16} Hz?

c What specific type of waves are those described in part (b)?

d Explain one helpful use of the waves described in part (b).

22 Explain briefly why a microwave oven is tuned to produce electromagnetic waves of a particular frequency.

23 A Young's double-slit experiment is performed using red light and the resulting interference pattern is observed. What is the effect on the observed interference pattern of:

a halving the separation of the slits

b doubling the distance between the slits and the screen

c halving the frequency of the light used

d covering one of the slits?

24 Two polarising sheets are placed together, with their polarising axes parallel, and are held up to a light source. One of the polarising sheets is rotated through 360° while the other is held still. The transmitted light alternates between maximum and zero intensities for different alignments of the polarising sheets.

a What alignment of the two polarising sheets produces the maximum light intensity?

b Why are there two maximums and two zeros of intensity during a 360° rotation?

25 At the time when Thomas Young carried out his famous double-slit experiment, there were two competing models claiming to explain the nature of light.

a What were the names of the two competing models?

b Explain how Young's experiment supported one of these models and not the other.

26 Physicists replicating Young's famous double-slit experiment determine that particular adjacent dark bands on the interference pattern (e.g. the third dark fringe and the fourth dark fringe) are different in their distance from one of the slits by only 500 nm. Determine the wavelength of the monochromatic light being used.

27 When investigating the photoelectric effect, the relationship between the maximum kinetic energy of emitted photoelectrons and the frequency of the light incident on the metal plate is:

$$K_{max} = hf - \phi$$

a Explain the meaning of the terms K_{max}, f and ϕ in this equation.

b If the intensity of the light striking the metal is increased, but the frequency is unaltered, what effect does this have on the value of K_{max}?

c If the intensity of the light striking the metal is increased, but the frequency is unaltered, what effect does this have on the value of the current flowing in the apparatus?

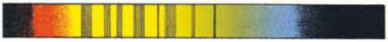
28 You are in a spaceship travelling at very high speed past a new colony on Mars. Do you notice time going slowly for you; for example, do you find your heart rate is slower than normal? Do the people on Mars appear to be moving normally? Explain your answers.

29 A high-speed subatomic particle is accelerated by a linear accelerator to a speed of $2.83 \times 10^8\,m\,s^{-1}$. A researcher measures that it only leaves, on average, a track that is 2.50 cm long in the bubble chamber. Calculate the mean lifetime of the same particle if it were at rest relative to the researcher and her timer.

30 Complete the following sentences by selecting the correct term in bold.

Muons have **very short/prolonged** lives. On average, muons live for approximately 2.2 **s/μs**. Their speeds are measured as they travel through the atmosphere. A muon's speed is **about a tenth of/very similar to** the speed of light. According to Newtonian laws, muons **should/should not** reach the Earth's surface. However, many **do/do not**.

31 An astronomer is analysing the following spectrum.



- **a** Explain what is meant by the term redshift and how you could analyse the spectrum of the star to find its translational velocity.
- **b** List three other qualities of a star that can be found by analysing its spectrum and describe how the spectrum of light can provide this information.
- **c** By comparing it to the OBAFGKM classification system, what colour do you expect this star to be, and in what temperature range?

32 A student carries out an experiment similar to Young's double-slit experiment. They shine a laser light of wavelength 600 nm through a double slit with a separation of 45 μm towards a screen which is 60 cm away from the slits.

- **a** Create a diagram of the experimental set-up the student used.
- **b** Create a ray diagram to describe how the fringes of bright and dark bands are created on the screen.
- **c** Calculate the distance between the bright bands.
- **d** If the laser light is changed to one with a wavelength of 500 nm, what is the percentage increase or decrease in the band separation?

33 When conducting a photoelectric effect experiment, a student correctly observes that the energy of emitted electrons depends only on the frequency of the incident light and is independent of the intensity.

- **a** Explain how the particle model accounts for this observation.
- **b** Explain why the wave model cannot account for this observation.
- **c** Make three statements about how the particle (photon) model of light is supported by features of the photoelectric effect and discuss the implications for the wave model of light in each case.
- **d** The student obtains the following data.

Wavelength (nm)	Maximum kinetic energy (eV)
350	1.4
400	0.91
450	0.60
500	0.25

Copy this table into your workbook and add a column containing the frequency for each of these values.

- **e** Create a graph of the energy as a function of frequency.
- **f** Analyse the graph to find a value of Planck's constant and the work function of the metal.
- **g** If there is an uncertainty of ±0.1 in the analysis of the work function, use the table below to find which metal was used in this experiment.

Metal	Work function
aluminium	4.08
carbon	4.8
copper	4.7
potassium	2.3
caesium	2.1

34 The star Xquar is at a distance of 5 light-years from Earth. Space adventurer Raqu heads from Earth towards Xquar at a speed of 0.9c.

- **a** For those watching from Earth, how long will it take for Raqu to reach Xquar?
- **b** From Raqu's point of view, how long will it take her to reach Xquar?
- **c** Explain why it is that, although Raqu knew that Xquar was 5 light-years from Earth, and that she was to travel at 0.9c, it took much less time than might be expected from these figures.
- **d** Why was the development of atomic clocks important to the advancement of Einstein's special theory of relativity?

35 The fusion reaction that powers the Sun effectively combines four protons (rest mass 1.673×10^{-27} kg) to form a helium nucleus of two protons and two neutrons (total rest mass 6.645×10^{-27} kg). The total power output of the Sun is a huge 3.9×10^{26} W.

- **a** How much energy is released by each fusion of a helium nucleus?
- **b** How many helium nuclei are being formed every second in the Sun?
- **c** How much mass is the Sun losing every day?
- **d** The Sun shoots out mass in the solar wind. Calculate the relativistic momentum of a carbon-12 nucleus in the solar wind if it is travelling at a speed of $750\,m\,s^{-1}$ and it has a rest mass of 1.99×10^{-26} kg.

MODULE

8 From the universe to the atom

Humans have always been fascinated with the finite or infinite state of the universe and whether there ever was a beginning to time. Where does all the matter that makes up the universe come from? Ideas and theories about the beginnings of the universe, based on sound scientific evidence, have come and gone. Current theories such as the big bang theory and claims of an expanding universe are based on scientific evidence available today through investigations that use modern technologies. Evidence gathered on the nucleosynthesis reactions in stars allows scientists to understand how elements are made in the nuclear furnace of stars. On scales as large as the universe to those as small as an atom, humans look to the sky for answers through astronomical observations of stars and galaxies.

Beginning in the late 19th and early 20th centuries, experimental discoveries revolutionised the accepted understanding of the nature of matter on an atomic scale. Observations of the properties of matter and light inspired the development of better models of matter, which in turn have been modified or abandoned following further experimental investigations.

By studying the development of the atomic models through the work of Thomson and Rutherford, who established the nuclear model of the atom—a positive nucleus surrounded by electrons—you will come to a better understanding of the limitations of theories and models. The work of Bohr, de Broglie and, later, Schrödinger demonstrated that the quantum mechanical nature of matter was a better way to understand the structure of the atom. Experimental investigations of the nucleus have led to an understanding of radioactive decay, the ability to extract energy from nuclear fission and fusion, and a deeper understanding of the atomic model.

Particle accelerators have revealed that protons themselves are not fundamental, and have continued to provide evidence in support of the Standard Model of matter. In studying this module, you will appreciate that the fundamental particle model is forever being updated and that our understanding of the nature of matter remains incomplete.

Outcomes

By the end of this module you will be able to:

- analyse and evaluate primary and secondary data and information PH12-5
- solve scientific problems using primary and secondary data, critical thinking skills and scientific processes PH12-6
- communicate scientific understanding using suitable language and terminology for a specific audience or purpose PH12-7
- explain and analyse the evidence supporting the relationship between astronomical events and the nucleosynthesis of atoms and relate these to the development of the current model of the atom PH12-15

CHAPTER 13

Origins of the elements

Astrophysics attempts to answer some of the biggest questions often asked about the universe, such as how old it is, how it began, whether it will last forever, and whether this is the only universe. Astrophysics is a fascinating mix of basic physics, the very latest theories and new technology. The development of supercomputers, along with space travel, has given access to methods of seeing the universe in ways never dreamed of before.

This chapter identifies the theories that provide the most current and best answers to some of the big questions about the universe: from the big bang to the formation of stars.

Content

INQUIRY QUESTION

What evidence is there for the origins of the elements?

By the end of this chapter you will be able to:

- investigate the processes that led to the transformation of radiation into matter that followed the 'big bang' CCT ICT
- investigate the evidence that led to the discovery of the expansion of the universe by Hubble (ACSPH138) ICT N
- analyse and apply Einstein's description of the equivalence of energy and mass and relate this to the nuclear reactions that occur in stars (ACSPH031) CCT
- account for the production of emission and absorption spectra and compare these with a continuous black-body spectrum (ACSPH137) CCT ICT
- investigate the key features of stellar spectra and describe how these are used to classify stars N
- investigate the Hertzsprung–Russell diagram and how it can be used to determine the following about a star: CCT ICT N
  - characteristics and evolutionary stage
  - surface temperature
  - colour
  - luminosity
- investigate the types of nucleosynthesis reactions involved in main-sequence and post-main-sequence stars, including but not limited to: CCT ICT
  - proton–proton chain
  - CNO (carbon–nitrogen–oxygen) cycle.

13.1 The big bang

PHYSICS INQUIRY N CCT

The expanding universe

What evidence is there for the origins of the elements?

COLLECT THIS...

- balloon
- sticky dots, or small pieces of paper with tape on them
- permanent marker
- string
- ruler

DO THIS...

1 Inflate the balloon a small amount and twist it to keep the air in. Do not tie yet.
2 Draw a grid on the side of the balloon. This will be your reference frame.
3 Stick a cluster of four dots on the grid side of the balloon, separated by a couple of centimetres each, so that some are closer to each other than others. Measure the separation between each of the dots by using the string on the curved surface, and then measuring the string with the ruler. Each dot represents a galaxy.
4 Draw a sinusoidal wave on the other side of the balloon. Measure its wavelength.
5 Untwist the balloon and inflate it further. Measure the distance between the galaxies and the wavelength again. Repeat this step one or two times depending on the stretchiness of the balloon.

RECORD THIS...

Describe how the change in wavelength can be evidence for an expanding universe.

Present a table of your results, including the galaxies that are moving away from each other faster than others.

REFLECT ON THIS

What evidence is there for the origins of the elements?

What is the difference between a Doppler redshift and the cosmological redshift?

What limitations does this inflating balloon analogy have in explaining the expansion of the universe?

What other evidence exists for the origin of the elements?

WS 8.1

Throughout recorded history, scientists have always sought to find an answer to how the universe began. They have also wondered if, when and how it will end. Astronomers have constructed hypotheses called cosmological models to try to find the answer. In the 1940s, the steady state theory and the big bang theory were the two contenders.

The big bang theorists won out in the end, and it's now the main theory explaining the origins of the universe we see today. This section looks at the big bang theory (Figure 13.1.1) and how the evidence supports this theory.

FIGURE 13.1.1 The big bang is often mistakenly thought of as an explosion, but it's much more complicated than that.

PREDICTION OF A DYNAMIC UNIVERSE

Scientists in the early 1900s worked to develop theories to explain the origin of the features of the universe they could observe. Using Einstein's theory of general relativity, the Russian scientist Alexander Friedmann developed a mathematical proof that predicted the universe was either expanding or contracting. At the time even Einstein did not think this was correct.

For scientists to decide if the universe was expanding, contracting or static, they needed some evidence. In a flat universe the expansion would slow and could eventually stop (Figure 13.1.2). In an open universe the expansion would speed up, and in a closed universe the expansion would eventually reverse and contract. Observational or experimental evidence that supported one of these models would allow scientists to reach a consensus on which prediction was correct.

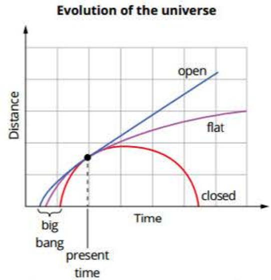


FIGURE 13.1.2 Predictions like Friedmann's proposed different fates for the universe. Are there other possibilities on the graph that are more catastrophic?

Modelling the expanding universe

The idea that space itself is expanding is complicated. It is easy enough to think of things expanding *in* space—a balloon being blown up, for example. Space seems so permanent and fixed, but this is not the case. It can be helpful to think of a two-dimensional analogy of what is really a three-dimensional problem. Imagine a little ant on the surface of the balloon as it is being blown up, as shown in Figure 13.1.3.

The balloon is so big the ant thinks of the surface as flat (as we do of the Earth's surface). The balloon has little paper stars stuck onto it. As the balloon expands the ant would see all the stars moving away from it. Wherever it wandered on the balloon's surface the same pattern would apply—and the further apart the stars the faster they would be receding.

The stars on the balloon were stuck on and not drawn on for a good reason. As the balloon expands, drawn stars would expand also, but the stick-on stars stay the same size. This is how astrophysicists picture the expansion of space. It is space that is expanding, not the stars and galaxies in it.

FIGURE 13.1.3 As the balloon is blown up the stars all move apart. The more distant stars will move away faster than the closer ones.

EVIDENCE FOR EXPANSION

In the 1920s astronomers were discovering that the universe was a much bigger place than they had previously thought. Astronomers had not yet determined that objects they called 'spiral nebulae' were actually other galaxies like our Milky Way but vast distances away (Figure 13.1.4). Edwin Hubble pioneered a technique to find the distance to these other galaxies and showed they were outside our own.

In Chapter 9, absorption and emission spectra were introduced. Imagine passing a whole visible spectrum of light through a gas. When light passes through the gas, it doesn't pass through unchanged. Different gases will absorb different wavelengths of the light. If the light that has passed through a gas is examined, any wavelength that was absorbed by the gas will be missing from the spectrum. This will appear as a black line corresponding to the absorbed wavelength. A quantitative analysis of how these spectra are produced will be discussed in Chapter 15.

GO TO ➤ Section 9.2 page 249

GO TO ➤ Section 15.1 page 392

FIGURE 13.1.4 The Andromeda galaxy is one of hundreds of billions of galaxies in the universe. It is the closest galaxy of similar size to our own Milky Way galaxy.

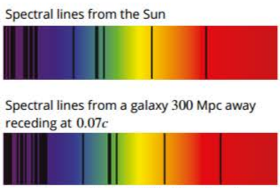


FIGURE 13.1.5 Notice that the black lines in both spectra form the same arrangement but the bottom set of lines are 'shifted' slightly to the right when compared to the lines in the spectrum from the Sun. This is known as 'redshift'.

When Hubble carefully examined the light from stars in the galaxies he found that it showed the familiar spectra of the elements, such as hydrogen and helium, that we know on Earth, but all the lines were redshifted, or in other words moved towards the red end of the spectrum (Figure 13.1.5, page 343). Astronomers interpreted this redshift as meaning that the galaxies were receding from us at huge speeds. This allowed them to calculate a velocity of recession from the redshift.

They are not moving through the universe relative to us. They appear to be moving away from us because the universe between us and them is getting bigger. Astronomers know today that the redshift of light from distant galaxies is due to the expansion of the universe. This effect is called cosmological redshift and is illustrated in Figure 13.1.6. As light travels, the wavelength of light is stretched by the expansion of the universe, and the further it travels the more it is stretched.

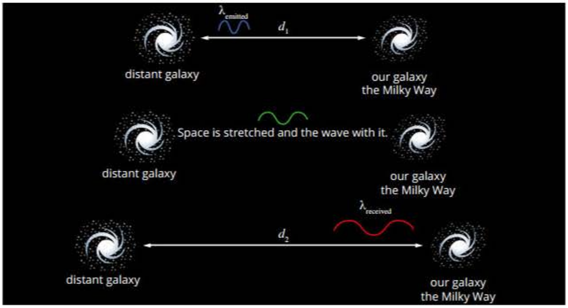


FIGURE 13.1.6 As the universe expands, so too does the wavelength of light travelling through the vast distances between galaxies.

PHYSICSFILE N

Parallax and parsecs

Hold your arm out straight and point towards the corner of the room, where the ceiling and two walls meet. Close one eye and line up your finger so it points exactly into the corner. Now open that eye and close the other without moving your finger. You will find that your finger is no longer pointing to the corner. If you keep your arm still, but keep swapping eyes, it will seem as though your finger is moving compared with the corner. This is an example of **parallax movement**.

With the development of better telescopes in the nineteenth century, it was noticed that some stars showed a very small movement against the background stars over the course of a year. It was realised that this was actually due to the Earth's rotation around the Sun causing parallax.

As shown in Figure 13.1.7, the parallax angle, p, is clearly a measure of the distance to the star. The larger the angle, the closer the star. In fact, the key measure of distance in space is defined as the reciprocal of the parallax angle measured in arc seconds (that is $\frac{1}{3600}$ of a degree). Stated simply, a star with a parallax angle of 1 arcsec is said to be at a distance of one parsec (1 pc). Astronomers define the parallax angle as the angle subtended by the radius of the Earth's orbit (1 astronomical unit, or 1 AU). A comparison of different astronomical distances is given in Table 13.1.1.

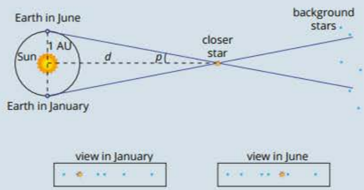


FIGURE 13.1.7 The parallax movement of a close star relative to the background 'fixed' stars. The parallax angle, p, is half the total apparent shift in angle.

TABLE 13.1.1 Relationships between astronomical distance units.

	Distance in:			
	m	AU	Light-years	Parsec
1 AU	1.496×10^{11}	1	0.000016	0.000005
1 light-year	9.461×10^{15}	63240	1	0.3066
1 parsec	3.086×10^{16}	206265	3.2616	1

HUBBLE CONSTANT

Combining his data on the distance to galaxies and their redshift, Hubble produced a graph (Figure 13.1.8) which made it clear that there was a relationship between the properties of the galaxies he had measured. Hubble found that the further away the galaxy, the faster it seemed to be moving. When Hubble graphed his data, he found that the 'speed of recession' was actually proportional to the distance away. This has become known as **Hubble's law** and is expressed by the simple equation:

$v = H_0 d$

where

v is the speed of recession ($km\,s^{-1}$)

H_0 is Hubble's constant ($km\,s^{-1}\,Mpc^{-1}$)

d is the distance away (Mpc)

H_0 is the **Hubble constant** and corresponds to the gradient of the graph. It is usually given in units of $km\,s^{-1}$ per megaparsec. It has a value of around $70\,km\,s^{-1}\,Mpc^{-1}$. So, for example, a 'typical' galaxy at a distance of 100 Mpc would be rushing away from Earth at $7000\,km\,s^{-1}$. That's about the same as travelling from Perth to Sydney in half a second!

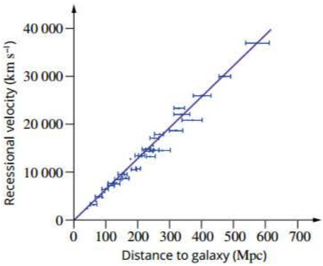


FIGURE 13.1.8 Hubble discovered that the galaxies were receding (rushing away) from Earth and that their recessional velocities were proportional to their distance from Earth.

SKILLBUILDER N

DETERMINING THE SLOPE FROM A LINE OF BEST FIT

There are many examples in science where there is a linear relationship between two variables measured in an experiment. Plotting this data, finding a line of best fit through the data and determining the slope of the line of best fit is often useful.

1 Decide on a scale to use for each axis that will allow all data points to be plotted.
2 Clearly label each axis with the name of the quantity and its unit.
3 Plot each data point on clearly labelled, unbroken axes.
4 Identify and label but otherwise ignore any suspect data points.
5 Draw, by eye, the line of best fit for the points. The points should be evenly scattered either side of the line. Recall that a straight line graph fits the equation $y = mx + c$.
6 Locate the vertical axis intercept and record its value as c.
7 Choose two points on the line of best fit to calculate the gradient, m. Do not use two of the original data points, as this will not give you the gradient of the line of best fit. The gradient is equal to $\frac{\text{rise}}{\text{run}} = \frac{\Delta y}{\Delta x}$. If you are using a computer or graphics calculator, it will be able to calculate the gradient for you.
8 Write the equation for your straight line using $y = mx + c$, replacing x and y with appropriate symbols, and use this equation for any further analysis.

Note: Do not force your line of best fit through the origin (0,0) or even consider it as a data point as this is rarely valid.

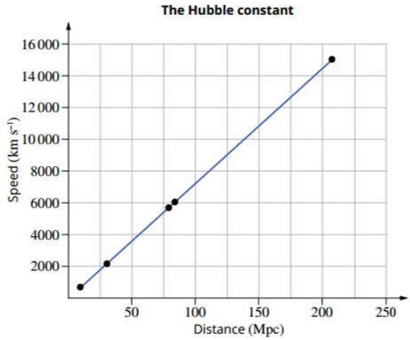
Worked example 13.1.1

CALCULATING THE HUBBLE CONSTANT

Astronomers have found the following data for the speed of distant galaxies:

Distance (Mpc)	Velocity ($km\,s^{-1}$)
10	730
31	2200
80	5600
84	6000
207	15000

a Analyse the data to find a value for the Hubble constant.

Thinking	Working
Construct a graph of the distance versus the velocity and draw a line of best fit.	
Interpret the graph.	The gradient of the graph is equal to 72.6. This means that for this data the Hubble constant is equal to 73 (to the correct number of significant figures). $H_0 = 73\,km\,s^{-1}\,Mpc^{-1}$

b Using the Hubble constant found in part a, calculate the recessional speed of a galaxy that is 450 Mpc from the Earth.

Thinking	Working
State Hubble's law.	$v = H_0 d$
Substitute the given values into Hubble's law.	$v = 73 \times 450$
Calculate the speed.	$v = 33\,000\,km\,s^{-1}$ (to two significant figures)

Worked example: Try yourself 13.1.1

CALCULATING THE HUBBLE CONSTANT

Astronomers have found the following data for the speed of distant galaxies:

Distance (Mpc)	Velocity (km s^{-1})
20	1300
50	3400
65	4300
120	8100
160	11 000

a Analyse the data to find a value for the Hubble constant.

b Using the Hubble constant found in part a, calculate the recessional speed of a galaxy that is 700 Mpc from the Earth.

WS 8.2

THE EARLY UNIVERSE

Logically, if space is expanding, then at some time in the past it must have occupied a much smaller space. If this is true then all the matter and energy in the universe must have been contained in this small space. This tells us that the early universe must have been very hot, with all the energy and matter in the universe packed into a much smaller space. If we take this further there must have been a time when the universe was created. We call this creation event the big bang.

Throughout the 1930s and 1940s physicists developed the basis for the modern big bang theory. George Gamow proposed the first serious theory that the universe began as a point, which expanded outwards in a hot big bang and which has been expanding ever since. He suggested that initially the universe was full of energy. When it expanded and cooled, nuclei and atoms formed and this sowed the seeds for the universe we see today.

Two competing theories

Not everyone agreed with the ideas put forward by Gamow. In 1949 astronomer Fred Hoyle made this very clear and soon put forward what became known as the **steady state theory**.

He suggested that the universe is:

- infinite—the 'outer' stars would never reach infinity and so could go on moving away from us forever.
- expanding—matter is being created all the time at just the right rate to keep the density of the universe constant.

The **big bang theory** was the main competitor to Hoyle's steady state theory. It stated that galaxies outside the Milky Way were moving away from us. This implied that if time is run backwards, at some instant in time, the universe and everything in it was contained in a single point. In other words, everything was created from nothing in one 'big bang'.

It was in fact Hoyle who first used the expression 'big bang' to describe a universe that started off from a 'primordial atom'. At the time, he was intending to be derogatory, but the name stuck.

Searching for the evidence

The astrophysicists of the day really were caught in a dilemma. Either way they had to accept the unacceptable: that matter just came into being out of nothing—either continuously or in one 'big bang'.

It seemed impossible to resolve this dilemma. If Hoyle was right, then the new matter being created should be able to be found somewhere. And as for the big bang, the only observational evidence—that the universe was expanding—was consistent with either theory.

PHYSICSFILE CCT

Time, space, matter and...

It is important to understand that the big bang was not seen as some sort of explosion from a small point in space. It was more an explosion *of* space, or, more correctly, the creation of space–time.

Einstein had already shown that time and space were not separate entities. So the big bang would not have occurred at some point in time any more than it would have occurred at a point in space. Time, space and matter were all created together in the one event we call the 'big bang'.

If matter was being created at a steady rate, it was not too difficult to calculate that about two or three atoms of hydrogen would need to be created every day in a volume about the size of a large sports arena to account for the observed density and expansion. That didn't seem too impossible, but there was not much point in trying to look for it. The universe is so big, it would be hard to know where to look.

If the big bang theory was true, the universe would have been denser at an earlier time. It seemed that only some means of looking back in time could resolve the problem.

EVIDENCE FAVOURS THE BIG BANG

In the 1960s physicists at Princeton University were trying to resolve the question of the steady state theory versus the big bang theory. They decided that if the big bang did in fact happen, it must have been extraordinarily hot at the beginning. If so, it would have produced large amounts of radiation, as does any hot object. (More detail about how objects radiate heat can be found in Chapter 11.)

GO TO ➤ Section 11.1 page 284

They also proposed that, as the universe expanded, this radiation would have been stretched by the expansion. They calculated that the wavelength would have increased to around a millimetre by now. That means that, in effect, the apparent temperature of the radiation would have fallen from billions of degrees to just a few degrees above absolute zero.

As it happened, two other physicists, Robert Wilson and Arno Penzias, at the nearby Bell Telephone labs, had discovered that their new antennas, designed to pick up radio signals from satellites, seemed to be picking up 'interference' from space. It turned out that this radiation was at just the sort of wavelength that Princeton physicists were looking for. Furthermore, it had the characteristics of heat radiation of just 2.7 K (that is, 2.7 degrees above absolute zero) as predicted. This was solid evidence in favour of the big bang theory at last, as there was no mechanism by which the steady state theory could produce this sort of radiation.

Wilson and Penzias were not sure at first where the interference picked up by the antenna was coming from. They thought it could have come from the Sun, but then ruled this out after some investigation. They made sure the interference wasn't from sources in the cities. They even ruled out any effect from pigeons by removing them and their droppings from around the antenna. Their realisation of what actually caused the interference was unexpected and absolutely groundbreaking.

This leftover radiation came to be called the cosmic microwave background radiation, or CMB radiation for short. Arno Penzias and Robert Wilson won the Nobel Prize for this discovery. Special satellites have mapped it carefully (Figure 13.1.9), as it is, in effect, a radio fingerprint of the early universe.

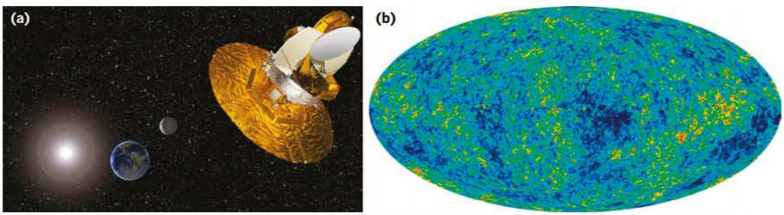


FIGURE 13.1.9 (a) The WMAP satellite mapped the cosmic microwave background radiation and discovered tiny fluctuations in it, which probably marked the beginning of the formation of galaxies. (b) The cosmic microwave background radiation.

The CMB radiation is very uniform from all directions in the universe. This is to be expected, because it filled the early universe and still does today. However, it is not totally uniform. The tiny variations in the CMB radiation represent differences in temperature of only about 0.0003 K, but they are very significant because it means the early universe was not totally uniform. Had the early universe been uniform, it would not have been able to create stars and galaxies. The small variations meant that matter was slightly clumped together in some places and so its unbalanced gravity could gradually pull it together to form the stars and, on a larger scale, galaxies.

Inflation

In order to explain the initial stages of the big bang, astrophysicists found that there must have been an incredible period of '**inflation**' right at the beginning. This period only lasted about 10^{-24} s, but during that time the size of the universe expanded to about 10^{50} times its original size.

This short period of inflation was necessary in order to prevent a very rapid collapse of the initial universe back into a black hole. Had the inflation lasted any longer, however, the greater expansion would have meant that atoms would have never formed.

Matter and antimatter

To understand where the matter in the early universe came from, a brief introduction to the Heisenberg uncertainty principle is needed (this will be discussed in greater detail in Chapter 15). It says that there is a fundamental uncertainty in the position and momentum of any object. The more accurately one variable is known, the less accurately we can know the other variable.

GO TO ➤ Section 15.2 page 409

It can also be interpreted as an uncertainty between mass and time. So in any very short time there is a fundamental uncertainty about mass. This is proportional: the shorter the time, the greater the uncertainty. The actual meaning of this is that in a very short time, mass can come in and out of existence.

Furthermore, when particles are created, they always come in pairs. For every matter particle there is an **antimatter** particle. But because they are so close when they pop into existence, they immediately **annihilate** each other (convert to energy). So matter and antimatter particles are actually popping in and out of existence everywhere all the time (Figure 13.1.10). All of this sounds like science fiction, but we know these pairs of particles are created and annihilated as we can measure it in the lab. Particle physicists have been studying them for almost 100 years.

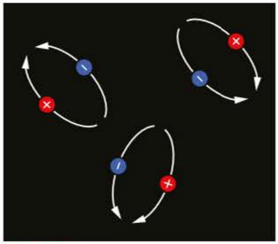


FIGURE 13.1.10 Electron–positron pairs were rapidly going in and out of existence in the early universe, but as they met they annihilated.

Creation of matter

To return to the origin of our universe, the two ideas just discussed can be combined—the inflation at the very beginning, and the continual creation and annihilation of matter. As two opposite particles are always produced in this creation, the process is called **pair production**. Due to annihilation, normally pair production doesn't result in the creation of any lasting matter. However, during the period of inflation, because of the extremely rapid expansion of space, pairs of particles rapidly became separated and didn't get a chance to annihilate. And so in that tiny fraction of a second of inflation, huge amounts of matter were created.

In those first moments after inflation, the universe was absolutely chaotic, with particles and antiparticles annihilating each other and producing high-energy gamma photons (electromagnetic radiation). Those photons themselves again collided with others and their energy formed new particle pairs (Figure 13.1.11).

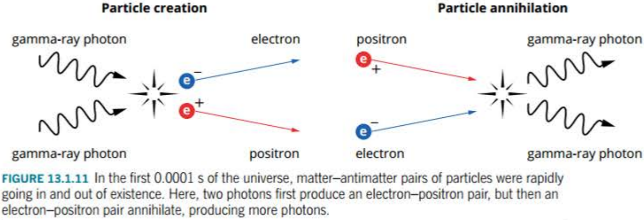


FIGURE 13.1.11 In the first 0.0001 s of the universe, matter–antimatter pairs of particles were rapidly going in and out of existence. Here, two photons first produce an electron–positron pair, but then an electron–positron pair annihilate, producing more photons.

As the temperature of the expanding universe fell to about 10^{12} degrees and became too low for the creation of new particles, there was a great amount of annihilation of matter with antimatter and a huge reduction in the total amount of matter. The annihilation also produced an enormous increase in the amount of radiation. This radiation filled all of space and dominated the universe for the next few hundred thousand years. It was, of course, the origin of the CMB radiation seen today.

PHYSICSFILE L

Antimatter

Antimatter particles are similar to ordinary matter particles, but have the opposite sign. An 'antielectron' is called a positron and is the same as an electron except that it is positively charged. An antiproton has the same mass and other properties as a proton, but it has a negative charge. When a matter particle and an antimatter particle come together they annihilate each other, releasing all their mass as high-energy photons (electromagnetic radiation). Theoretically we could construct ordinary objects out of antimatter, but as soon as they came into contact with normal matter the two would annihilate with the release of absolutely huge amounts of energy.

Formation of simple elements

Since the 1930s physicists have worked to develop theoretical predictions of the creation and evolution of matter based on the idea of a hot big bang origin of the universe. One example is that the energetic early universe should have produced a composition of approximately 25% helium and the remainder hydrogen (Figure 13.1.12). Spectroscopic studies in the 1960s and 1970s showed that stars and galaxies were composed of approximately 24% helium. Numerous other predictions from theory were slowly supported by evidence as better observations were made. Today, many of the predictions are tested directly in particle accelerators like the Large Hadron Collider at CERN. The energetic conditions of the early universe are recreated billions of times each day, and particle physicists probe the conditions that existed in the universe only a tiny fraction of a second after the big bang. All of this has led to what we think is the history of the universe as told by the big bang theory.

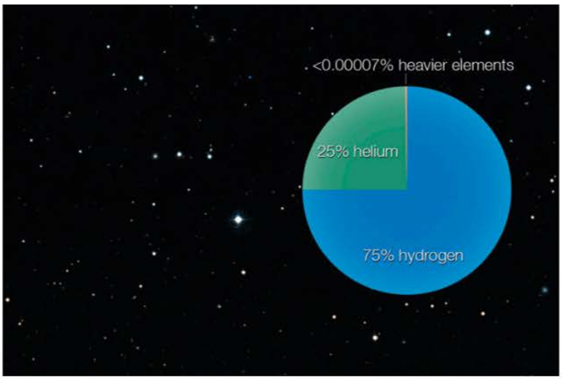


FIGURE 13.1.12 The percentages of the elements measured in one of the universe's oldest stars. It has approximately 75% hydrogen, 25% helium and only a tiny fraction of other elements, which is in good agreement with predictions made by the big bang theory.

In the first few seconds of the universe, while the temperature remained over a billion degrees or so, elementary particles known as **quarks** combined to form protons and neutrons. Protons and neutrons were forced close enough to fuse (join) together, forming hydrogen, helium and lithium nuclei. After a few more minutes the temperature dropped below that needed for this **fusion** and no further nuclei were formed (Figure 13.1.13).

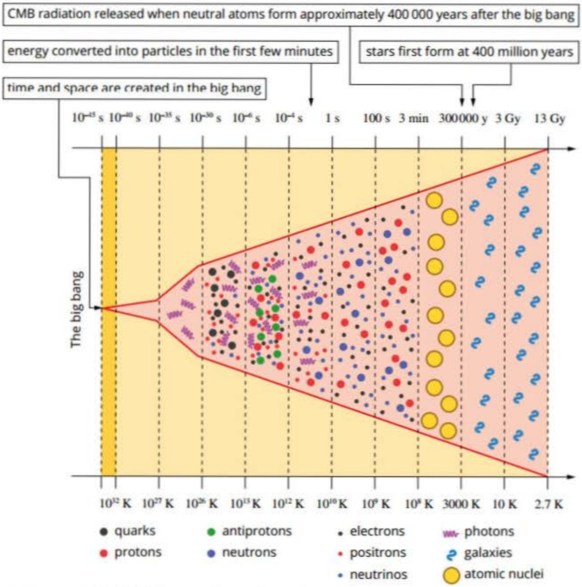


FIGURE 13.1.13 The 13.8 billion year history of the universe.

So how was any matter left after all the annihilation of matter and antimatter? Strangely, there was not quite an even balance of matter and antimatter in the early universe and so after the initial rapid annihilation, there was actually matter left over. The imbalance between matter and antimatter has been estimated at only about one extra matter particle in a billion particles—but that one in a billion now makes up all the matter in our universe.

Without that very slight imbalance between matter and antimatter, the universe would be empty—except for the CMB radiation.

For around 300 000 years the universe was too hot for atoms to form. Any electron captured by a proton was soon knocked away by an energetic photon. But by that time, the expansion of the universe resulted in the temperature dropping to about 3000 K. This meant photons were now in the red or infrared part of the spectrum. They did not have enough energy to ionise (strip electrons from) atoms so they stopped being scattered by atoms and became what we now detect as the CMB radiation. Now, atoms of hydrogen, helium and lithium could form as these nuclei captured free electrons. Only these three types of nuclei had formed in the first few minutes while the universe was hot enough for fusion to occur.

MATTER FORMS STARS, GALAXIES AND EVENTUALLY US

For the next billion years or so the universe was expanding and cooling, and was very dark. However, due to very slight irregularities in the original 'quantum foam' in which all those early particles came into existence, matter was starting to clump together. This process accelerated as the matter came closer and gravitational attraction increased.

The matter eventually started to form huge clumps, which are now the galaxies (refer back to Figure 13.1.13 to put this stage into context). Within the galaxies matter collapsed into many smaller clumps. As these smaller clumps collapsed, energy was released when the matter (mostly hydrogen and helium) crashed together. This raised the temperature within the 'clumps' to millions of degrees. At these temperatures, the fusion (joining) of atoms was reignited—with hydrogen nuclei fusing to form new helium nuclei and releasing yet more heat. Stars were born.

These early stars were huge and very hot. They burnt out relatively quickly, and after using up all their hydrogen fuel they collapsed. This collapse, however, resulted in the release of yet more energy in huge supernovae explosions (Figure 13.1.14). These explosions were so energetic that they resulted in the fusion of lighter elements into the heavier ones, a crucial process for the eventual formation of life.

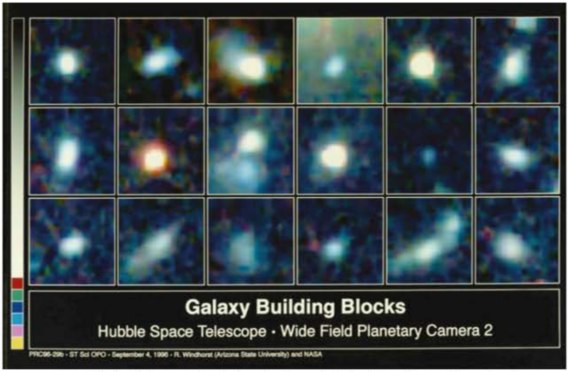


FIGURE 13.1.14 Early galaxies contained short-lived very hot (blue) stars. They collapsed as huge supernovae which resulted in the creation of the heavier elements.

Our Sun and solar system formed from the gravitational collapse of the 'dust' from some of these earlier supernova explosions and so contain heavier elements. Fortunately for us, the heavier elements were concentrated in the terrestrial (rocky) planets, including the Earth. And thus, eventually, life evolved.

PHYSICS IN ACTION ICT

Cosmological modelling

Studying the evolution of the universe is known as cosmology. Without being able to make direct observations of the very early universe, physicists use computer modelling to make predictions.

Utilising supercomputers with extremely fast processing power, scientists create models by inputting different initial conditions, and allowing their model to evolve. In Figure 13.1.15, this computer model shows the clumping together of matter where galaxies would form. If the output of the model matches the conditions that can be observed in the universe today, it can be assumed that the model works.

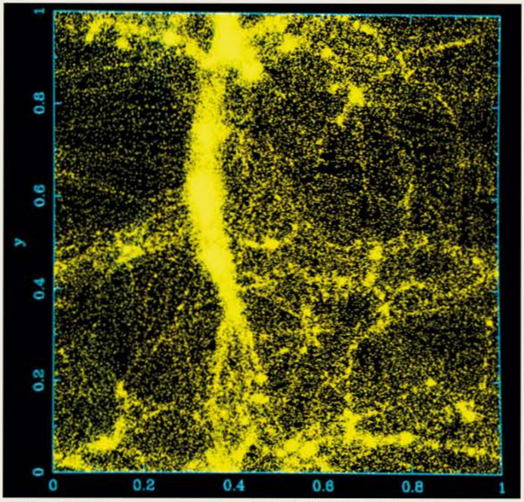


FIGURE 13.1.15 Physicists use computer modelling to study the evolution of the universe.

13.1 Review

SUMMARY

- General relativity predicted an expanding universe and Edwin Hubble found the evidence that supported this prediction. He saw that the further away a galaxy was, the faster it seemed to be receding from us.
- Last century, two different theories of the evolution of the universe seemed possible. Either the universe was in a 'steady state' with matter being continuously created or matter was all created with a 'big bang'.
- The big bang theory predicts many features of the early universe and there are numerous pieces of evidence that support its predictions; for example, the existence of the cosmic microwave background radiation and the abundance of simple elements in the universe.
- Space and time were created in the big bang.
- Energy was transformed into matter in an extremely short time called inflation. The expansion was so rapid that matter and antimatter pairs of particles could not annihilate until the expansion slowed.
- In the early universe there was slightly more matter than antimatter, and that excess is the matter that makes up the universe today.
- Galaxies and stars formed from matter as the result of very small irregularities in the early universe. The irregularities allowed gravitational attraction to 'pull' matter together.
- The heavy elements necessary for life came later with the explosive supernova of heavy stars.

KEY QUESTIONS

1 The steady state theory assumes the universe has a constant density. How does this theory take into account the expansion of the universe?

2 Why is the discovery of cosmic microwave background radiation seen as supporting evidence for the big bang theory and not the steady state theory?

3 Describe the formation of matter in the early universe.

4 The inflation period in the early universe lasted only for an incredibly short time. What might have been a consequence had it lasted for a longer time?

5 What happens when an electron meets a positron?

6 At only a few thousandths of a second after the big bang, matter and antimatter annihilated almost completely. What is the reason there was some matter left over?

7 All matter in the universe was created in the first few hundred thousand years and yet few of the elements essential for life existed then.

a What elements did exist at that time and why?

b How were the elements needed for life created?

8 Explain how Hubble's findings from studying galaxy redshifts support an expanding universe.

13.2 The life cycle of a star

Astrophysicists use spectroscopy to analyse light from distant stars. Spectroscopy involves spreading out the spectrum of light (both visible and invisible light) received from a star and then analysing patterns in the data collected. The spectra are compared with others from Earth that are known to scientists. From spectral analysis, the properties and even the life cycle of stars can be determined. Many of these properties were introduced in Chapter 9.

GO TO ➤ Section 9.2 page 249

The absorption spectra from other stars are quite similar to the solar spectrum in most respects. This confirms that the basic chemistry and physics understood on the Earth seems to apply throughout the universe. No strange new lines have ever been discovered in any other stellar spectra. There are subtle differences, however, and it is these differences that tell astronomers an enormous amount about other stars.

ℹ The spectral lines observed in emission and absorption spectra are generated through the transitions of electrons between energy levels within atoms. This is discussed in greater detail in Section 15.1.

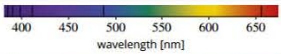
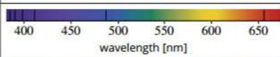
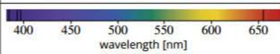
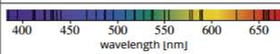
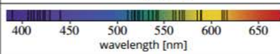
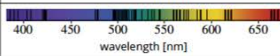
CLASSIFYING STARS

The method of classifying stars based on the presence or absence of certain lines in the spectra (the sequence OBAFGKM) was introduced in Chapter 9 and is repeated in Table 13.2.1.

The differences in the stars' spectra also correspond to different temperatures, with the O stars being the hottest and the M stars being comparatively cooler. This can be found by analysing the black-body spectrum from stars. The peak intensity of a black-body spectrum coincides with a specific wavelength that is inversely proportional to the temperature.

ℹ A star emits light across a range of wavelengths and it can be described as a black body. A black body does not reflect any radiation.

TABLE 13.2.1 The spectral type or spectral class classification of stars based on the absorption lines present in their spectra.

Spectral type	Approximate temperature (K)	Main characteristics	Spectrum	Colour
O	50000–28000	relatively few lines; the lines of ionised helium		blue
B	28000–10000	the lines of neutral helium		blue–white
A	10000–7500	very strong hydrogen lines		white
F	7500–6000	strong hydrogen lines; ionised calcium lines; numerous metal lines		yellow–white
G	6000–4900	strong ionised calcium lines; numerous strong lines of ionised and neutral iron and other metals		yellow
K	4900–3500	strong lines of neutral metals		orange
M	3500–2000	titanium oxide streaks		red

It was shown that the changes between the classes were associated with the fact that different atoms become ionised (lose electrons) at different temperatures. At cooler temperatures the light may not have enough energy to excite the atoms sufficiently to create some of the lines. So it was realised that some lines may appear over a certain temperature range but not at higher or lower temperatures, as shown in Figure 13.2.1.

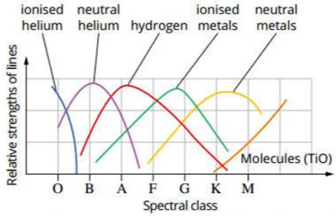


FIGURE 13.2.1 The relative strength of common spectral lines is shown for each spectral type or spectral class. By comparison with Table 13.2.1 the relationship between dominant spectral lines and temperature can be deduced.

A careful analysis of a spectrum reveals much important information about the star, not just its temperature. The mind map in Figure 13.2.2 sums up the relationship between some of these measurements and quantities.

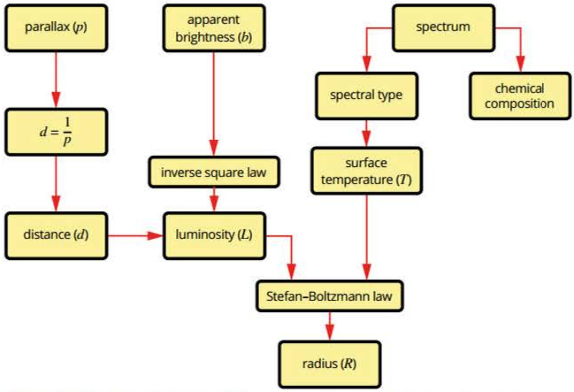


FIGURE 13.2.2 A mind map illustrating what particular aspects of stars' light can tell astronomers about a star.

If the surface temperature of a star is known from the spectrum, a basic law of physics (the Stefan–Boltzman law) can be used to determine the amount of energy given off each second by each unit area of the surface. Knowing this and the **luminosity** (the total energy given off by the star) enables the total surface area and hence the radius of the star to be calculated. The luminosity is found using the apparent brightness and the distance determined by stellar parallax.

THE HERTZSPRUNG–RUSSELL DIAGRAM

Once data for many stars had been collected, it was natural to look for patterns in the data. In 1911 and 1913, the astronomers Ejnar Hertzsprung in Denmark and Henry Russell in America independently discovered a pattern that brought order to the apparent chaos. The graph they produced is now known as the **Hertzsprung–Russell diagram**, or H–R diagram for short. In effect they plotted the luminosity of a star against its temperature and found a distinct pattern.

Figure 13.2.3 shows the common way of arranging the H–R diagram with luminosity on the vertical axis and temperature on the horizontal axis. Each axis is labelled with two different scales. Temperature is derived from spectral type, so these are both plotted on the horizontal axes. Luminosity is derived from the **absolute magnitude**, so these are plotted together on the vertical axes.

You'll notice from the H–R diagram that the stars are not randomly distributed. The band of stars from the lower-right corner to the upper-left corner is called the **main sequence** and includes 90% of the stars in the sky. Main-sequence stars go from dull, cool stars at the bottom right to bright, hot stars at the upper left of the diagram. There is a clear relationship between the brightness (luminosity) and temperature of a star. Generally, it was found that hot stars in the top-left corner of the H–R diagram are brighter and bigger, but there were important exceptions to this rule as you can see.

In one sense this is hardly surprising—if two stars are the same size it is reasonable to expect the hotter star to be brighter. However, stars in the upper right of the diagram, such as Betelgeuse, are very large. They need to be large in order to be so bright, given that their temperatures are relatively low. This group of stars are called **giants** or supergiants and are over one thousand times the size of the Sun. On the other hand, stars in the opposite lower-left corner must be small as they are not bright, despite being very hot. These stars are only about one-thousandth the size of the Sun and are called **white dwarfs**.

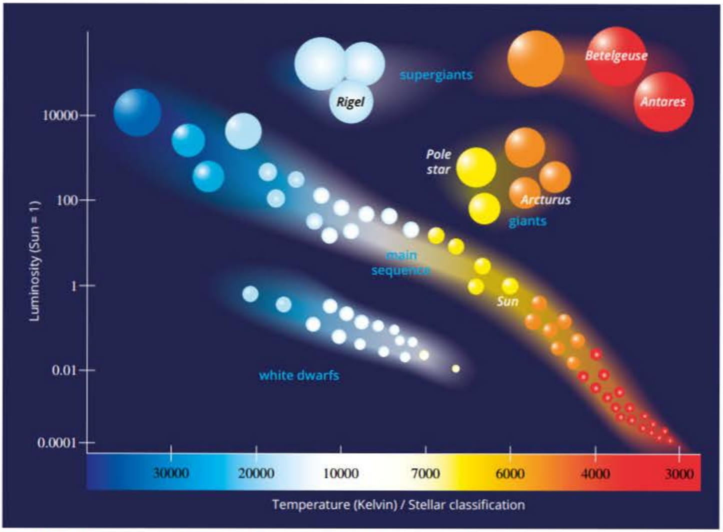


FIGURE 13.2.3 The Hertzsprung–Russell (H–R) diagram is one of the astronomer's basic tools. The luminosity scale takes the Sun's value as 1, which corresponds to an absolute magnitude of about +4.8. The temperatures corresponding to the spectral types are plotted along the bottom. Hotter stars are to the left, indicated also by the colours of the stars throughout the diagram.

Worked example 13.2.1

DETERMINING THE LUMINOSITY OF A STAR USING THE H–R DIAGRAM

A main-sequence star is observed to have a surface temperature of 4000 K. What is its approximate luminosity?

Thinking	Working
Determine where on the H–R diagram this star would sit.	The main sequence is a band that runs from bottom right to top left on the H–R diagram, and a temperature of 4000 K should be at the bottom right of this band.
Draw a vertical line from the required temperature on the *x*-axis to intersect with the luminosity curve. Draw a horizontal line from the intersection of the luminosity curve to the *y*-axis.	The star's luminosity is approximately 0.01 times that of the Sun.

Worked example: Try yourself 13.2.1

DETERMINING THE LUMINOSITY OF A STAR USING THE H–R DIAGRAM

A main-sequence star is observed to have a surface temperature of 20 000 K. What is its approximate luminosity?

Most stars are on the main sequence, and the hotter they are, the brighter they are. The Sun is approximately in the middle of the main sequence. By adding up the stars in the main sequence on the H–R diagram (Figure 13.2.3), you can see that most stars are both cooler and dimmer than the Sun—there is a greater density of stars on the lower right of the main sequence. Only about 1% of stars are giants or supergiants (Figure 13.2.4), and a further 9% are white dwarfs.

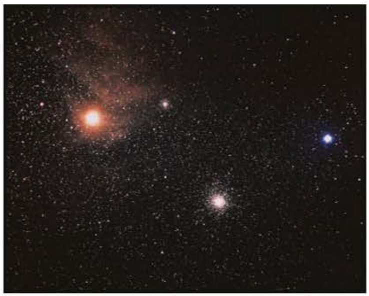


FIGURE 13.2.4 The brightest star, at centre left of this image, is Antares (Alpha Scorpii), a red supergiant star 500 light-years from the Earth.

Placing stars on the H–R diagram

If measurements of a star are made, it can be plotted on the H–R diagram. If the star is close enough to Earth, luminosity and temperature can be determined directly. If a star is much further away, its temperature can still be determined from its spectrum but its luminosity can't be determined directly. This means its position on the vertical axis of the H–R diagram is not known. Luckily the spectral lines in the star's spectrum provide additional information that allows luminosity to be determined.

+ ADDITIONAL

The brightness of stars

Astronomers measure the apparent brightness of stars on a scale that actually originated with Hipparchus in the second century BCE! He called the brightest stars he could see 'first-magnitude' stars (+1), those about half as bright 'second-magnitude' (+2) and so on to those barely visible, which were 'sixth-magnitude' (+6).

When astronomers sailed into the Southern Hemisphere they discovered brighter stars, and so the scale had to be extended 'upwards' to 0 magnitude and then −1 magnitude and so on.

When the telescope was invented the scale went 'downwards' to +7 and beyond.

The scale is referred to as the apparent magnitude scale. Don't be confused by the fact that the scale seems 'backwards'—dimmer stars have a numerically higher magnitude.

In the nineteenth century when astronomers were better able to measure the brightness of stars, they defined apparent magnitude more precisely by saying that a difference in magnitude of 5 corresponds exactly to a factor of 100 times in apparent brightness—which agreed roughly with the old values. In other words, it would take 100 stars of magnitude +6 to equal the brightness of one star of magnitude +1. Mathematically, this means that each level of magnitude represents a change in brightness of about 2.5 times (instead of Hipparchus's 'double'). Therefore it would take 2.5 stars of magnitude +6 to equal the brightness of a single +5 star, for example, or it would take $2.5 \times 2.5 = 6.3$ stars of magnitude +6 to equal the brightness of a +4 star. Figure 13.2.5 shows the scale with some examples. You might like to check with your calculator that it would take about 13 billion Sirius A stars to equal the apparent brightness of the Sun! (What we call 'Sirius' is actually two stars very close together—Sirius A and the much dimmer Sirius B.)

Clearly there must be a relationship between the apparent brightness of a star, its distance from Earth and its **intrinsic brightness** (actual brightness). Once the actual distances of a number of stars were known, it was possible to determine their intrinsic brightness. It was soon found that stars vary enormously in intrinsic brightness. For example, Sirius and Canopus are the two brightest stars in the sky; however, although Canopus appears almost as bright, it is actually about 36 times further away than Sirius. To make up for this big difference in distance, it can be calculated that Canopus must be about 3000 times brighter than Sirius. The very closest star, Proxima Centauri, is not even visible to the naked eye and so must be a very dim star. In fact, it is intrinsically only about $\frac{1}{30000}$ as bright as Sirius. The intrinsic brightness of stars varies over a huge range.

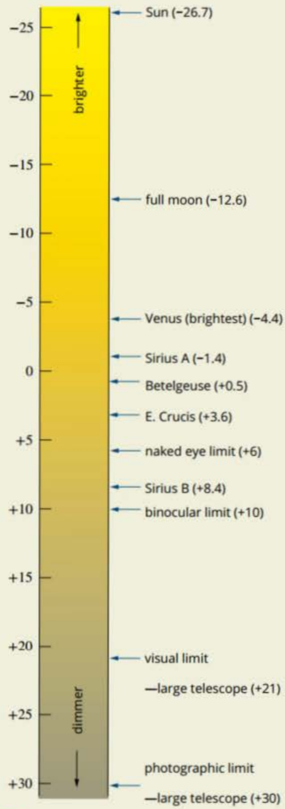
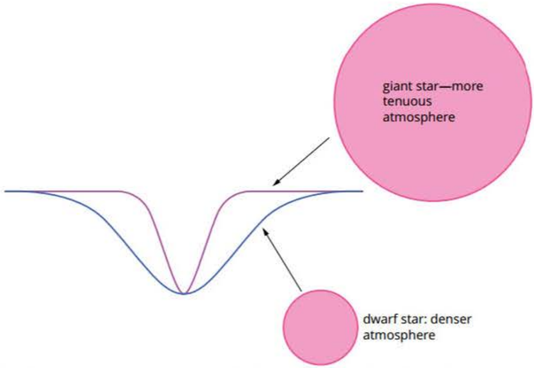


FIGURE 13.2.5 The apparent magnitude scale is based on an ancient Greek scale with nineteenth-century mathematical corrections. Visible stars range from −1.44 (Sirius A) to about +6 for stars barely visible under the best conditions.

WS 8.3

Spectral lines in stellar spectra tell us more than just the elements present. The pressure of the gas from which the light originated has a subtle effect on the width of spectral lines as they are sensitive to the number of collisions the atoms are experiencing. The collisions tend to slightly increase or decrease the energy and hence the wavelength of the emitted light. This in turn slightly broadens the spectral line (Figure 13.2.6). A giant star will tend to have a less-dense photosphere, and hence the atoms will experience fewer collisions. This means that the spectral lines are less broadened than those of a main-sequence or dwarf star.



13.2.6 The pressure in the atmosphere of a star is proportional to the broadening of the absorption lines seen in their spectra. The higher the pressure in the atmosphere, the more broadened the absorption lines are.

Known relationships between the amount of line broadening due to pressure and the luminosity of stars have led to a classification system using luminosity classes as shown in Figure 13.2.7. This system uses Roman numerals I to V. A numerical way of classifying a star from its spectral features then consists of its spectral class and luminosity class, e.g. our Sun has a spectral class G2 and a luminosity class V. This allows the position of the Sun to be located on the H–R diagram on both axes.

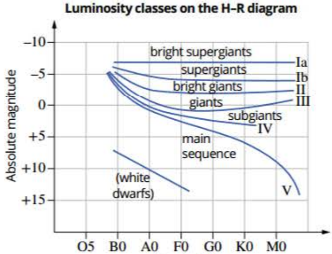


FIGURE 13.2.7 Luminosity class provides an estimation of the vertical position of a star on the H–R diagram when a direct measurement of luminosity is not available.

PHYSICSFILE AHC DD

Indigenous astronomy

Well before Western science understood the processes of how stars formed, their movement across the sky influenced the knowledge and understanding of the natural world in many different cultures. For example, Greek astronomers studied the cosmos in depth, and Indigenous Australians observed the night sky. Indigenous observations over thousands of years informed Indigenous knowledge systems to navigate, find food and predict tidal movements. These systems of knowledge aimed to explain and predict natural processes which have been passed down through generations.

In general, much of the astronomical information has informed knowledge and understanding of seasonal calendars. For example, the beginning of the dingo breeding season for the Pitjantjatjara people is signalled when the Pleiades star cluster rises in the morning. Similarly for Torres Straight Islanders, when the stars of the Big Dipper reach the horizon in the evening it tells people to steer clear of the water as it is now shark-breeding season.

13.2 Review

SUMMARY

- The continuous spectrum of a star provides information about the temperature of its surface, and the absorption lines tell us what elements are present and under what conditions.
- The spectra of stars are quite similar to the spectrum of the Sun, indicating that they have similar characteristics and processes.
- Stars are classified according to a system which uses the absorption lines visible in stellar spectra to organise them according to surface temperature (OBAFGKM; hot to cool respectively).
- The Hertzsprung–Russell (H–R) diagram plots stars according to surface temperature and colour on the horizontal axis. Surface temperature increases to the left as colour goes from red to orange then yellow to white and finally blue. Luminosity or absolute magnitude is on the vertical axis and increases upwards.
- All stars fall into three main groups: main sequence; giants and supergiants; and white dwarfs.
- Moving up the main sequence from right to left on the H–R diagram, the stars become hotter, brighter and larger.

KEY QUESTIONS

1 Which properties of a star are plotted on an H–R diagram?

2 According to the H–R diagram, how does the luminosity of a main-sequence star relate to its temperature?

3 What information does the continuous spectrum emitted by a star and its absorption spectrum reveal about the star?

4 Compare emission and absorption spectra with a continuous black-body spectrum.

5 A new star has been discovered and astronomers wish to place it on the H–R diagram. It's observed that the spectral lines on the spectra of the star are significantly less broadened than those from the Sun and the spectrum has a peak intensity towards the red end of the visible spectrum. What type of star has been observed?

6 In general, which of the following is true for most stars?

A luminosity increases with the surface temperature
B luminosity decreases with the surface temperature
C luminosity is not related to the surface temperature
D luminosity is directly proportional to the surface temperature

7 How do the luminosity and surface temperature of an A-type main-sequence star compare with those of the Sun?

8 Which of the following best describes what is plotted on the Hertzsprung–Russell diagram?

A absolute magnitude against the luminosity of the stars
B apparent magnitude against the temperature of the stars
C temperature against the spectral type of the stars
D luminosity against the spectral type of the stars

13.3 The life and death of stars

THE SUN: EARTH'S CLOSEST STAR

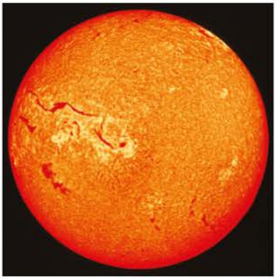


FIGURE 13.3.1 The apparent movement of sunspots and other features across the Sun's surface make its rotation clear. Sunspots and flares reveal the active and ongoing processes within the Sun. Nuclear fusion reactions deep inside the Sun release huge amounts of energy that stream from the Sun, resulting in a conversion of about 4 million tonnes of mass into energy every second.

It was only after the time of Galileo and Newton that scientists realised that the stars were actually Sun-like objects a very long way away. The best way to learn about stars, then, was to look at the closest one.

Galileo's discovery of sunspots led to the understanding that the Sun rotates on its axis (Figure 13.3.1). However, the equator of the Sun was seen to have a period of about 25 days, while regions of higher latitude took several more days for a full cycle, indicating that the Sun is not a solid body like the Earth. Because it is obviously so hot it was assumed to be gaseous, but just how it produced so much heat wasn't understood.

It was known that life had existed on the Earth for at least several hundred million years and so that meant that the Sun must have been radiating energy at fairly much the same rate for at least that time. The problem was that any known mechanism for producing heat could not possibly have generated so much energy over such a long period of time.

The English physicist Lord Kelvin (who also determined the temperature of absolute zero) and the German physicist Hermann von Helmholtz proposed that vast amounts of heat would be generated from the enormous weight of gas collapsing into the Sun. As it fell, the potential energy would be converted into kinetic energy and would produce heat. This was good physics, but again it could not possibly last for the billions of years that the Sun has been producing energy. Scientists now know their theory was wrong, but the Kelvin–Helmholtz contraction does turn out to be important in the formation of new stars—and in the death of old ones.

THE EQUIVALENCE OF MASS AND ENERGY

Calculations based on the mass and energy output of the Sun showed that the amount of energy being produced for each atom in the Sun was around a hundred million times greater than the energy produced by each atom in chemical reactions. Something very different was going on in the Sun.

The clue came with Einstein's theory of special relativity. Einstein's equation $E = mc^2$ is well known, but the real meaning of this equation is not so simple. A reasonable interpretation for our present purposes is that there is a huge amount of energy locked up in mass. The factor c^2 is the speed of light squared and so is a very large number (approximately 9×10^{16} m^2 s^{-2}). This means that if this energy can be released, a small loss of mass could produce a huge amount of energy.

It was thought that the Sun was somehow tapping into this so-called 'mass–energy'. An enormous amount of energy is being radiated from the Sun. Assuming the energy was being produced mainly in the core, it was calculated that the temperature at the centre of the Sun must be millions of degrees. This would mean that the atoms at the centre would be stripped of their electrons and there would be just a frenzied mass of charged nuclei and electrons flying around at enormous speeds. Since the density of the Sun was relatively low, it was assumed that it must be composed mostly of hydrogen with perhaps some helium.

In the 1920s, British astrophysicist Robert Atkinson put all of this together and suggested that in the conditions at the centre of the Sun the hydrogen nuclei may 'fuse' together, creating heavier helium nuclei. If this were the case, huge amounts of energy would be released. This process is called nuclear fusion and results in the formation of the nuclei of heavier elements, so it is often called **nucleosynthesis**.

GO TO ➤ Section 12.3 page 326

The process of fusion was discussed in greater detail in Chapter 12.

Worked example 13.3.1

MASS–ENERGY EQUIVALENCE

The Sun is producing about 4.0×10^{26} J of energy every second as visible and invisible radiation. At what rate is the Sun losing mass due to this total energy loss? (Use $c = 3.0 \times 10^8\,\text{m s}^{-1}$.)

Thinking	Working
The energy comes from the fusion of hydrogen into helium with a corresponding loss in the potential energy of the nuclei. This loss of energy will correspond to a mass loss given by Einstein's equation $E = mc^2$.	$E = mc^2$ $E = 4.0 \times 10^{26}\,\text{J s}^{-1}$ $c = 3.0 \times 10^8\,\text{m s}^{-1}$ $m = ?$
Rearrange $E = mc^2$ in terms of mass and solve.	$E = mc^2$ $m = \frac{E}{c^2}$ $= \frac{4.0\times10^{26}}{(3.0\times10^8)^2}$ $= 4.4 \times 10^9\,\text{kg s}^{-1}$ So the Sun is losing mass due to visible and invisible radiation at a rate of $4.4 \times 10^9\,\text{kg s}^{-1}$.

Worked example: Try yourself 13.3.1

MASS–ENERGY EQUIVALENCE

The visible portion of the energy the Sun is producing each second is approximately equal to $5.0 \times 10^{25}\,\text{J s}^{-1}$. At what rate is the Sun losing mass due solely to this energy loss?

(Use $c = 3.0 \times 10^8\,\text{m s}^{-1}$.)

PHYSICSFILE L

Einstein's $E = mc^2$

It is often stated that in nuclear reactions 'mass is converted into energy'. This is actually an oversimplification of Einstein's ideas. There is no mass 'shaved off' the particles and somehow mysteriously turned into energy. What is true is that the total mass of the particles in a helium nucleus is a little less than the total mass of the equivalent particles in hydrogen nuclei. They are exactly the same particles as before, except that they now have a little less total potential energy.

Einstein showed that mass is actually a property not of the individual particles, but of the system of particles including the energy bound up in the forces between them. As some of the energy has been released, so also has the equivalent mass.

FUSION SHAPES THE STRUCTURE OF THE SUN

The combination of Einstein's theory of relativity and an improved model of the atom (Bohr's quantum atom was suggested in 1913) enabled astrophysicists to understand the processes occurring in the Sun. Bohr's model of the atom will be discussed in detail in Chapter 15. Today, computer models have enabled considerable insight into the mechanisms that keep the Sun 'burning' in the sky.

GO TO ➤ Section 15.1 page 392

The basic principle used in modelling the Sun is that any part of it must be in what is called **hydrostatic equilibrium** (Figure 13.3.2 on page 364). That means that on any 'piece' of Sun, the inward pressure from the weight of all the material above it must be balanced by the outward thermal pressure and radiation pressure originating from the nuclear reactions in the core below.

Computer modelling suggests a number of theories about the Sun's structure and production of energy. Nuclear fusion occurs in a zone that extends from the centre out to about 0.25 of the Sun's radius ($R_\odot$). The temperature in this region is above 10 million degrees, the density is about 160 000 kg m^{-3} and the pressure is about 340 billion times the Earth's atmospheric pressure.

The energy flows outwards from this zone by a combination of convection and radiative diffusion. Radiative diffusion is a process in which light bounces around, transferring energy in the process. It is the main mechanism for energy transfer out to about $0.7R_\odot$, where convection takes over. At this point the temperature is down to around one million degrees and the protons and electrons come together to form hydrogen atoms, which absorb the light more effectively and so the radiative diffusion becomes less effective. The density is only 80 kg m^{-3} (much less than water, but about 60 times that of air) and the pressure is down to about 10 million times Earth's atmospheric pressure. Most (99%) of the Sun's mass is below this level.

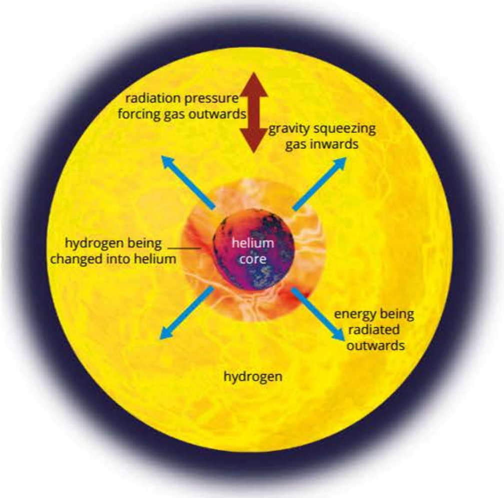


FIGURE 13.3.2 The Sun is currently in a state of hydrostatic equilibrium, where at any radius from the centre the inward pull of gravity is balanced by the pressure pushing outwards. The outward pressure is mostly the thermal pressure within the plasma with some contribution from radiation pressure.

At the surface, the temperature is down to approximately 5800 K. There is actually no real 'surface', just a layer where the churning hot gases start to sink again as they lose their energy by radiation into space. This layer, called the **photosphere**, is a thin layer from where the Sun's visible light is emitted (Figure 13.3.3). The whole process may sound rapid, but it has been calculated that energy produced by the nuclear fusion in the Sun's core takes about 170 000 years to travel out to the surface. It then takes just over eight minutes to reach the Earth. The energy received from the Sun today was actually generated inside the Sun when humans were just beginning to diverge from the apes! Some properties of the Sun are listed in Table 13.3.1.

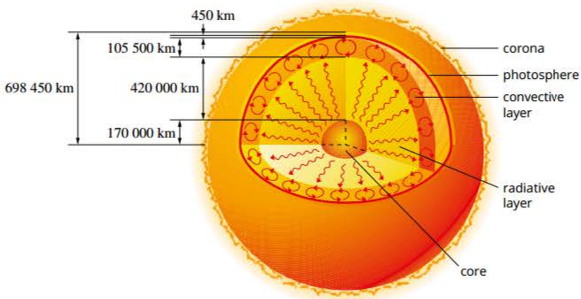


FIGURE 13.3.3 The photosphere is a thin layer from where the Sun's visible light is emitted.

TABLE 13.3.1 The physical properties of the Sun.

Average distance between the Earth and the Sun (1 AU)	1.496×10^{11} m (390 times as far as the Moon)
Angular diameter from Earth	0.5° (same as the Moon)
Mass of the Sun	1.99×10^{30} kg (300 000 times the mass of the Earth)
Diameter of the Sun	1.4×10^{9} m (109 Earth diameters)
Surface temperature	5780 K
Average density of the Sun	1.4×10^{3} kg m^{-3} (Earth is 5.5×10^{3} kg m^{-3})
Energy output of the Sun	3.86×10^{26} W (J s^{-1})

MODELS AND OBSERVATIONS UNCOVER STELLAR EVOLUTION

Before the twentieth century it had generally been thought that the stars were basically permanent features of an unchanging eternal universe. Scientists now understand the processes that fuel the nuclear reactions inside stars. As this understanding has developed, it has become obvious that while stars last a very long time, they do not last forever. The Sun, for example, will not last forever. The changes it will experience through its life cycle are depicted in the artwork in Figure 13.3.4.

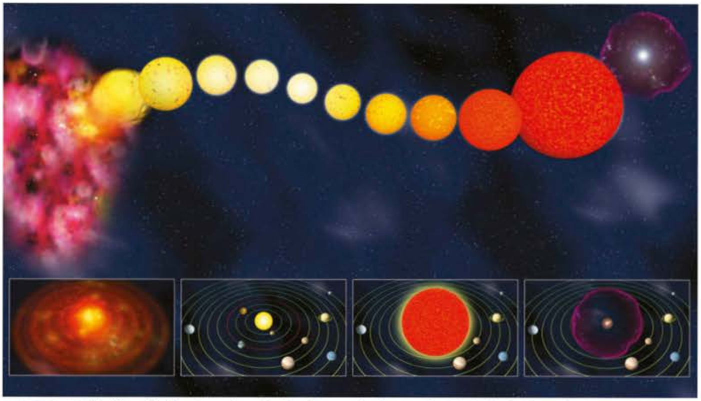


FIGURE 13.3.4 The life cycle of a star, such as the Sun, from molecular gas cloud in the beginning to the predicted red giant which will eventually swallow up the inner planets. Once the Sun is on the red giant branch it may then turn into a yellow giant before forming a planetary nebula and finally becoming a white dwarf.

The story of how astrophysicists have come to interpret the H–R diagram and to build up a picture of the life and death of stars is one of the most fascinating in physics. At the current rate at which the Sun is fusing its hydrogen, it should run out of fuel in about a hundred billion years. However, models of the nuclear processes occurring and the changing conditions with its core throughout its life predict a much shorter life span.

As the heavier helium nuclei build up in the core of a star, the nuclear fusion reaction zone moves outwards. Once around 10% of the hydrogen is consumed, the star becomes unstable and the outer layers expand until the star becomes about ten times the size. In other words, the star becomes a giant. The Sun is due to do this in around five billion years and, when it does, it will engulf the Earth.

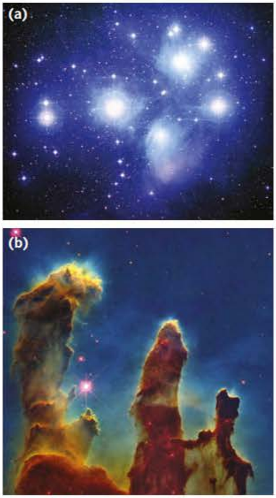


FIGURE 13.3.5 (a) The Pleiades star cluster, also known as the Seven Sisters or M45. It includes over 500 young stars, including the seven brightest visible to the naked eye. It can be seen in the constellation Taurus. (b) This famous photograph of the Eagle Nebula taken by the Hubble telescope shows huge gaseous pillars several light-years high. The pillars are evaporating gaseous globules and are thought to be regions in which stars are forming. The nebula can be found near the constellation of Sagittarius.

For this reason, it was originally thought that the stars are 'born' on the main sequence of the H–R diagram and then eventually moved into the giant phase. Heavier stars were expected to burn their fuel more rapidly than lighter stars. Although they have more fuel, computer models predict heavier stars will run out of fuel sooner and so they should have a shorter life span.

Confirming these models' predictions can't be achieved by watching what happens to a particular star through its lifetime as the time involved is millions or billions of years. Scientists have to infer, from the different types of stars that can be seen, which stars are examples of which different stages within their life cycles. Then they need to check whether their properties fit the predictions of the models.

There are groups of stars, called star clusters, which seem to have all been born at about the same time out of the same cloud of 'dust' and gas. The Pleiades group, or Seven Sisters as it is more commonly known, is one such cluster. The famous 'Starbirth' photograph shows another cloud where clusters of stars could be forming. Images of both are shown in Figure 13.3.5.

These clusters provide the opportunity to look for differences between stars of similar ages but different masses. If the modelling that suggests more massive stars will have shorter lifetimes is correct, there should be evidence of this in clusters. The bluer, hotter and more massive stars should be moving off the main sequence and becoming giants sooner than the less massive, cooler stars.

Many clusters were studied in this way in the 1950s and this is just what was found. Stars at the upper end of the main sequence were more likely to have moved towards the giant area than those further down.

The number of stars seen in the various areas on the H–R diagram is presumably an indication of the amount of time stars spend in that area. Because there are many more stars in the main sequence than in the giant phase it is assumed that stars spend most of their life as reasonably stable main-sequence stars. There are few stars between the main sequence and the giant phase, and so it can be deduced that after reaching the end of their main-sequence life, they fairly rapidly expand to become red giants.

NUCLEOSYNTHESIS AND STELLAR EVOLUTION

With the aid of computer modelling, a fairly complete picture of the birth, life and death of stars has been built up that agrees well with observational evidence.

A star begins its life as a **protostar**, a large mass of gas and dust that has come together as a result of gravitational attraction. The heat generated in this gravitational collapse (the so-called Kelvin–Helmholtz contraction mentioned earlier) causes the protostar to become very hot. As the gravitational collapse continues, the gravity becomes stronger, and accelerates the collapse even further (Figure 13.3.6).

FIGURE 13.3.6 The Horsehead Nebula in the constellation Orion. This nebula is part of the Orion Nebula complex, an enormous starbirth region some 1500 light-years from Earth.

Eventually the interior of the new star becomes hot enough, at about 10 million kelvin, to ignite the nuclear fusion reactions. The extra heat generated by these reactions eventually stops the gravitational collapse. This is as a result of the outward thermal pressure from the plasma and the radiation pressure from the radiation streaming outwards from the core. At this point, the protostar has become a new main-sequence star and remains in a fairly stable condition for a time that is dependent on the mass of the star. This time will be measured in millions of years for very massive stars, billions of years for stars like the Sun, and tens of billions of years for the majority of stars lower down the main sequence.

During the main-sequence phase of the life of a star the energy generated within a star comes from hydrogen fusion in its core. The mass of the star determines the temperature within the core and therefore which fusion reactions can occur. Figure 13.3.7 shows the relative rates of energy production via the two dominant fusion reactions that occur within main-sequence stars. These reactions are the proton–proton (PP) chain and carbon–nitrogen–oxygen (CNO) cycle reactions.

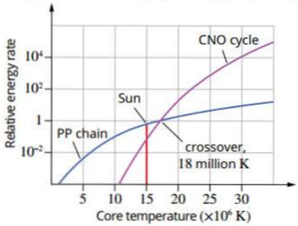


FIGURE 13.3.7 The relationship between temperature and the rate of energy release is shown on this graph for the two hydrogen fusion processes that are dominant in main-sequence stars. Higher-mass stars generate higher core temperatures, so as you move to the top-left corner of the main sequence on the H–R diagram the CNO cycle dominates energy production within the core.

Both the PP chain (Figure 13.3.8) and CNO cycle achieve the same result of combining four hydrogen nuclei to form one helium nucleus. The energy that holds a star up is released due to the mass defect between the products and the reactants. The series of reactions in the proton–proton chain dominates energy production in main-sequence stars with core temperatures below 18 million kelvin.

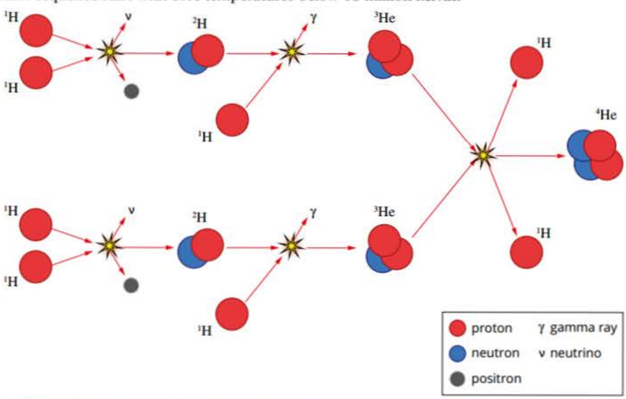


FIGURE 13.3.8 The proton–proton chain takes four hydrogen nuclei and creates one helium nucleus.

In more massive stars the more complex CNO cycle is dominant. The series of reactions involved can be seen in Figure 13.3.9. Carbon atoms act as a catalyst in this cycle as they are converted into nitrogen, then oxygen and back to carbon again. These more massive stars have much shorter lives than their lighter counterparts as the rate of reactions occurring increases drastically as temperature increases. Even though these more massive stars contain much more hydrogen, the rate of fusion far exceeds this increase in fuel and they use it up in a much shorter time.

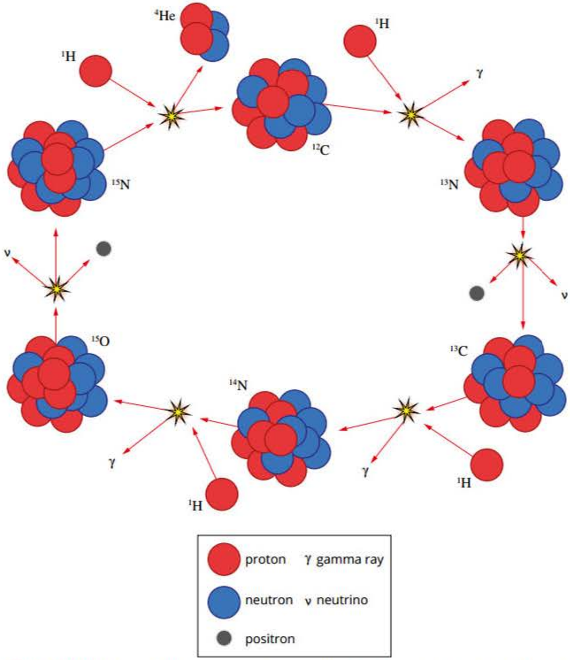


FIGURE 13.3.9 In the CNO cycle, carbon atoms are converted into nitrogen, then oxygen and back to carbon again.

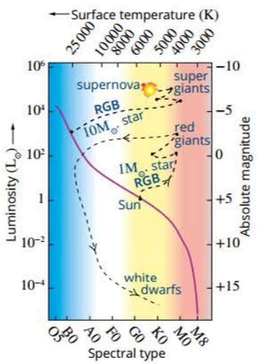


FIGURE 13.3.10 The life cycles of 1 and 10 solar mass stars are plotted on this H–R diagram. Once stars reach the end of their life on the main sequence they move up into the giants along the red giant branch (RGB). Depending on their mass, they may fade and produce a white dwarf or explode in a supernova and leave behind a neutron star or black hole.

As a main-sequence star ages it builds up non-fusing helium in its core. This causes the energy output to decrease and the star contracts. This contraction causes heating which increases the temperature and therefore the energy output. This process causes the star to slowly increase in luminosity over its main-sequence lifetime. Eventually the core contains so much non-fusing helium that contraction and heating cause a shell of hydrogen fusion to form around a non-fusing helium core. This marks the end of a star's life on the main sequence, and a shift up and to the right on the H–R diagram along the red giant branch (RGB) can be seen in Figure 13.3.10. This part of a star's life can be classified as post-main-sequence.

Once a star becomes a red giant a new process starts to occur in the core. The further gravitational collapse of the heavy core heats it to even higher temperatures. At around one hundred million degrees, new nuclear reactions start to occur: three helium nuclei combine to form carbon, as shown in Figure 13.3.11.

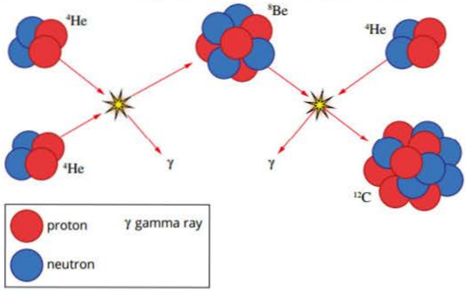


FIGURE 13.3.11 The two fusion reactions resulting in the formation of a carbon nucleus are together known as the triple alpha process.

This is the origin of elements that have made our existence possible. The general trend of the reactions inside stars is to fuse lighter nuclei to produce heavier ones. This nucleosynthesis is responsible for producing the elements that make up most of the mass in your body. Since the only elements formed in the early universe were hydrogen and helium, you could not exist unless stars produced carbon, oxygen, nitrogen, phosphorus and other elements that you require for life.

Stars less than eight solar masses

Once a star has reached the red giant phase, the fate of stars more or less than eight times the mass of the Sun ($8M_{\odot}$) is very different. Further development of a hydrogen fusing shell and a non-fusing core cause another series of changes to the stellar structure. A star with a mass of less than $8M_{\odot}$ will become unstable and will lose a large portion of its mass during pulsations of its diameter, forming a **planetary nebula** (Figure 13.3.12). What is left behind is a white dwarf; then the nebula dissipates and the white dwarf cools, as shown on the bottom left of the H–R diagram.

FIGURE 13.3.12 The Helix Nebula is classed as a planetary nebula and lies around 700 light-years from the Earth in the constellation Aquarius.

At this stage, the star has collapsed into electron degenerate matter that basically means that the atoms have collapsed into a mass of protons, neutrons and electrons with a density about a million times that of water. They gradually radiate their remaining heat away, become cooler and redder, and slide down and out of the H–R diagram at the lower right.

Stars greater than eight solar masses

The previous picture of stellar evolution is appropriate for most stars, but a more spectacular fate awaits some. A **supernova** is the explosive end of a massive star. There have been a number recorded in history, and some have been visible to the naked eye as a bright star that appears and then disappears weeks later. There was a notable supernova in the constellation Cassiopeia in 1572 that was observed by Tycho Brahe, who was so fascinated he went on to become one of the great astronomers of the period. A 1604 supernova had a similar effect on another great astronomer, Johannes Kepler. In 1987, a new supernova (shown in Figure 13.3.13) suddenly appeared in the Large Magellanic Cloud, a fuzzy area near the Southern Cross. The Large Magellanic Cloud is actually a satellite galaxy of the Milky Way.

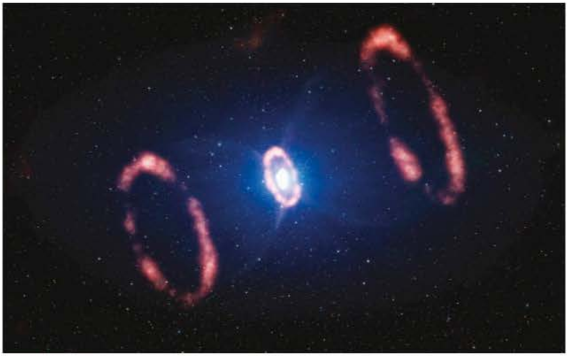


FIGURE 13.3.13 The remnants of Supernova 1987a. The rings are thought to be glowing gas that was ejected when the star was a red giant, and which was ionised by the intense UV light from the explosion.

Huge pressure is created inside more-massive stars. After the initial hydrogen-burning phase, the temperatures created by the contracting star can reach 600 million degrees, at which point new nuclear fusion reactions can take place. In these reactions, the carbon produced by the helium fusion is fused to produce oxygen, neon, sodium and magnesium. Depending on the mass of the star, various heavier elements are produced and burnt in a series of reactions that occur at temperatures up to billions of degrees. The effect of these reactions is to move the star around on the H–R diagram as it changes brightness and temperature with each new set of reactions. Stars over about $25M_{\odot}$ eventually become so large and bright that they form the supergiants at the top right of the H–R diagram. Betelgeuse and Rigel in Orion are examples of these supergiants.

There is an upper limit to the nucleosynthesis that can occur in stars. Once silicon has fused and iron has been produced, no further fusion reactions will occur as they would require a net input of energy. The star begins gravitational collapse again, generating more heat.

Neutron stars and pulsars

In an amazing sequence of events not totally understood, the core, largely composed of iron by this stage, rapidly skyrockets to many billions of degrees in a fraction of a second. This results in the protons and electrons in the core being forced together with such ferocity that they combine to form neutrons and small almost-massless particles called neutrinos.

The core by this stage is about 20 km in diameter but with a mass many times that of the Sun. In this state the core can collapse no further, but the rest of the mass of the star is still collapsing in on it, building up enormous pressure which creates an absolutely huge 'bounce back'. This sends the outer layers of the star flying off into space at enormous speeds, partly propelled by the neutrinos trying to escape from the core. The energy released in this explosion is an incredible 10^{46} J, which is far more than the Sun will produce in its entire life.

What is left of the star after this gigantic explosion is an extremely dense core made of 'neutron matter'. This is matter in which all the electrons and protons have collapsed to form neutrons. A **neutron star** with a diameter of 30 km would have the same mass as the Sun. A few cubic centimetres would have a mass of hundreds of millions of tonnes.

In the 1960s a number of **pulsars** were found. These were objects producing extremely regular pulses of radio emissions. Later, it was found that these pulses could also be seen as flashes in visible light. They are now thought to emanate from rotating neutron stars about 20 km in diameter but with about $1.4M_{\odot}$. The frequency of rotation of a star is greatly increased as it collapses, in just the same way that skaters or dancers can spin faster by pulling their arms closer to their body as they spin.

The periods of pulsars are measured in seconds or even fractions of a second—they are spinning at an incredible rate for their size. As they spin, leftover charged matter interacts with their powerful magnetic fields and produces a beam of intense radiation that sweeps around the sky like the beam from a lighthouse. If observers happen to be in line with the beam, a flash is seen each time it sweeps past.

The other remains from supernovae are nebulae (dust and gas clouds). An example of such a nebula is the Crab Nebula (seen in Figure 13.3.14). The Crab Nebula is the remnant of a supernova that was first seen from Earth in 1054. Nebulae are the ejected gases that glow for possibly hundreds of years after the supernova. Gradually they fade and disperse into space to form more of the dust and gas clouds out of which new stars will be born. Many of these remnants can only be seen by radio telescopes, as their temperature drops and the wavelength of their radiation becomes longer than that of visible light.

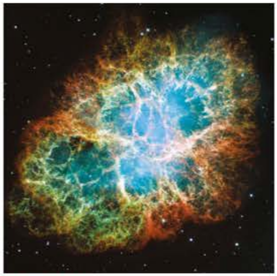


FIGURE 13.3.14 An image from the Hubble Space Telescope of the Crab Nebula, denoted M1. The nebula is the remnants of a supernova that was visible from Earth in 1054 and was bright enough to be seen during the day.

Black holes

What is most fascinating is what happens to really massive stars, i.e. over $20M_{\odot}$, once they run out of fuel. There is a point at which even the neutrons in a super-heavy neutron star will collapse. It seems that this further collapse has no end point at all and matter simply collapses into what is called a **black hole**. More particularly it is referred to as a 'singularity'—a point of infinitely small volume and infinite density. The force of gravity around such an object is so strong that nothing, including light, can escape. A radius exists around the singularity where the escape velocity exceeds the speed of light. This is called the event horizon and it is essentially the point of no return for anything, even light.

Black holes are not visible like stars but there have been many observations of gas and stars in the vicinity of an object that can only be a black hole due to its small size and enormous gravity. Our galaxy alone may contain as many as a million black holes. They are the remnants of the largest stars our galaxy has ever produced. Galaxies themselves are all thought to have huge black holes at their centres which have been growing since the galaxies were formed. The one at the centre of the Milky Way is about 4 million times the mass of the Sun.

13.3 Review

SUMMARY

- The Sun, like other main-sequence stars, produces energy by the fusion of hydrogen nuclei to form helium nuclei.
- Calculations based on the mass and energy output of the Sun showed that the amount of energy being produced for each atom in the Sun was around a hundred million times greater than the energy produced by each atom in chemical reactions. This can be explained by mass–energy equivalence and can be found by Einstein's equation $E = mc^2$.
- Most stars begin their lives on the main sequence of the H–R diagram, eventually become giants and then move downwards to become dwarf stars.
- As stars age, the products of nuclear fusion within the star change. Young stars fuse hydrogen to form helium through the proton–proton (PP) chain or carbon–nitrogen–oxygen (CNO) cycle. As stars age they produce increasingly heavy elements. For example, red giants fuse helium into carbon through the triple alpha process. Once silicon starts to be fused to produce iron, further nuclear fusion reactions would require a net input of energy and thus do not occur. The star begins to collapse under gravitational forces.
- Some very large stars end their lives as supernovae, which create the heavier elements, allowing the creation of new stars, planets and life itself.
- Some massive stars collapse to form super-dense neutron stars, pulsars and black holes.

KEY QUESTIONS

1 Why can the nuclear reactions involving hydrogen in the Sun produce so much more energy than hydrogen burning in oxygen on the Earth?

2 A star is losing mass at the rate of $6 \times 10^9\,\mathrm{kg\,s^{-1}}$. How much energy is being produced per second in total radiation if this is the case? (Use $c = 3.0 \times 10^8\,\mathrm{m\,s^{-1}}$.)

3 What is the order of stellar evolution for most stars?

4 The equation for the fusion of two isotopes of hydrogen (deuterium and tritium) is shown below.

$$^{2}_{1}\mathrm{H} + {}^{3}_{1}\mathrm{H} \rightarrow {}^{4}_{2}\mathrm{He} + {}^{1}_{0}\mathrm{H}$$

Explain why energy is released during this process.

5 Near the end of their lives, some stars, with masses less than $8M_{\odot}$, shed some of their mass as a series of shells of gas. What are these shells of gas called?

6 The most likely scenario for the end of the Sun once it has exhausted its fuel supply is to finish as what kind of astronomical object?

7 Describe the forces that are balanced in the Sun in its current state of hydrostatic equilibrium.

8 Explain how the nucleus of a carbon atom can be converted during a fusion reaction into the nucleus of a nitrogen atom.

Chapter review

13

KEY TERMS

absolute magnitude
annihilate
antimatter
big bang theory
black hole
fusion
giant
Hertzsprung–Russell diagram
Hubble constant
Hubble's law
hydrostatic equilibrium
inflation
intrinsic brightness
luminosity
main sequence
neutron star
nucleosynthesis
pair production
parallax movement
photosphere
planetary nebula
protostar
pulsar
quark
steady state theory
supernova
white dwarf

REVIEW QUESTIONS

1 Describe the evidence that Hubble found to support the idea of an expanding universe.

2 Name the process responsible for the initial formation of matter from energy after the big bang and the stage in the evolution of the universe characterised by rapid expansion.

3 Sketch a graph showing Hubble's law and write an expression illustrating how the slope is related to Hubble's law.

4 What evidence is there that the big bang theory correctly describes the origin of the simplest elements on the periodic table?

5 Explain how an expanding universe implies there was a hot dense beginning for the universe—what we now refer to as the big bang.

6 Describe the effect a larger Hubble constant would have on our estimate of the age of the universe.

7 Name the radiation that was predicted and subsequently observed that is strong evidence supporting the big bang theory.

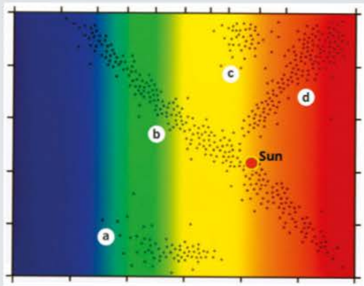
8 Which characteristics of a star can you directly determine from its spectrum?

9 As the stars are so far away, how can scientists be so sure that they are not made of totally new elements never seen before on Earth?

10 Referring to Figure 13.2.3 on page 357, order the following stars from smallest to largest.
Betelgeuse
Arcturus
Rigel
Polaris

11 Justify the presence of strong absorption lines from molecules in the spectrum of a star classified as spectral type M.

12 Below is a representation of the Hertzsprung–Russell (H–R) diagram in which stars are grouped by luminosity and temperature. Identify the types of stars shown by the labels a–d on the diagram.



13 Rigel is a blue supergiant star located in the middle-top of the H–R diagram. Based on this information, describe briefly, in point form, the main stages of its expected life cycle.

14 In terms of the H–R diagram, where are stars 'born'?

15 The Sun is an 'average' star. In what sense is this statement true in regard to the H–R diagram, and in what sense is it not true?

16 A star is losing mass at the rate of $4 \times 10^9\,kg\,s^{-1}$. How much energy is being produced per second in total radiation?

17 Compare the main fusion reactions that occur in main-sequence and red giant stars.

18 Explain how stars are responsible for forming most of the elements on the periodic table.

19 Assess the big bang theory's accuracy in explaining the origin of the elements using examples of available evidence.

20 In relation to the formulation of the big bang theory, what role does experimental evidence and observation play in the scientific method?

21 Describe the change within the interior of a star that marks the end of its main-sequence life.

22 Compare neutron stars to black holes.

23 After completing the activity on page 342, reflect on the inquiry question: What evidence is there for the origins of the elements?

CHAPTER

14 Structure of the atom

In this chapter you will develop an understanding of the theoretical models and experimental studies that explored and explained the nature of the atom. The chapter follows the development of the concept of the atom from a single, hard, indivisible component of matter to a model of a nucleus composed of protons and neutrons surrounded by a cloud of electrons.

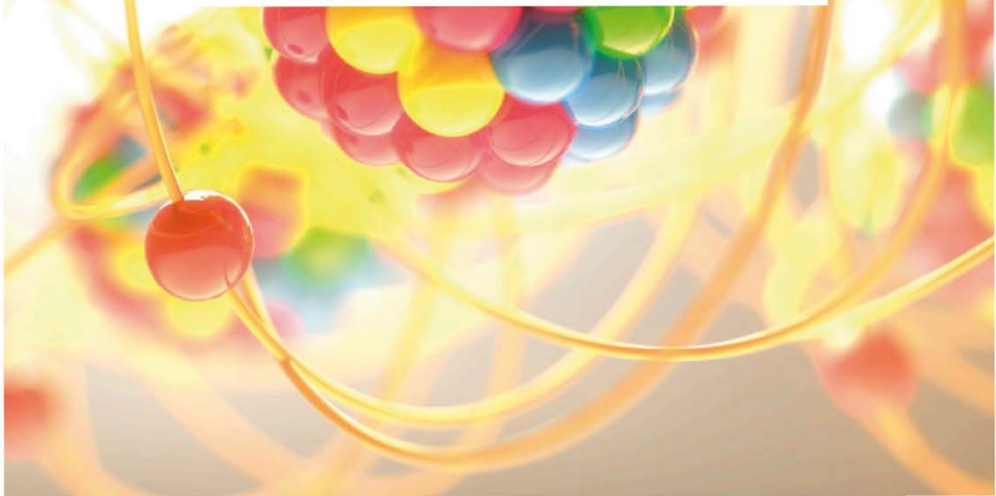
Content

INQUIRY QUESTION

How is it known that atoms are made up of protons, neutrons and electrons?

By the end of this chapter you will be able to:

- investigate, assess and model the experimental evidence supporting the existence and properties of the electron, including: ICT
  - early experiments examining the nature of cathode rays
  - Thomson's charge-to-mass experiment
  - Millikan's oil-drop experiment (ACSPH026)
- investigate, assess and model the experimental evidence supporting the nuclear model of the atom, including: ICT
  - the Geiger–Marsden experiment
  - Rutherford's atomic model
  - Chadwick's discovery of the neutron (ACSPH026).



14.1 The electron

PHYSICS INQUIRY N CCT

Chilled atoms

How is it known that atoms are made up of protons, neutrons and electrons?

COLLECT THIS...

- large mixing bowl
- two half-spherical jelly moulds or other container to set the models in
- clear gelatine
- shiny metal beads that reflect light
- hundreds and thousands
- laser pointer
- white paper

DO THIS...

1 Choose an element to make. Research the number of protons, neutrons and electrons in the element. In this model the electrons will be represented by hundreds and thousands, and the protons will be represented by metal beads. Count out two lots of electrons and one lot of protons in preparation for making two models of the element.

2 Mix gelatine according to the directions on the container. Place the gelatine in the refrigerator until it is half set. Oil the sides of the moulds.

3 Create a plum pudding atomic model by pouring half of the gelatine into one of the moulds and stirring in the hundreds and thousands to represent electrons. Place in the fridge to finish setting. To unmould, heat the outside of the mould with a warm wet towel or by placing the mould in warm water. Upturn the mould onto a plate.

4 Create a Rutherford atomic model by pouring some gelatine into the other mould until it is half full. Gently place the beads into the centre of the mould. Pour the remaining gelatine in carefully to cover the beads. Place in the fridge to finish setting. Turn out onto a plate. When the edge is still soft, gently press the hundreds and thousands into the side.

5 Set up the gelatine models of the atom with white paper behind to be a detection screen. Shine the laser into each model along different paths and record the path of the light. Only Class 1 and Class 2 laser pointers are permitted in schools. Discuss with your teacher any safety precautions around using lasers.

RECORD THIS...

Describe how observing the beam path can provide evidence of atomic structure.

Present your models and observations, highlighting the different results between the two models.

REFLECT ON THIS...

How is it known that atoms are made up of protons, neutrons and electrons?

What knowledge of the atom was needed before the evidence from the Geiger–Marsden experiment was able to be interpreted?

What does the gelatine represent in each of the models?

From observations of lightning to the temporary electric charges created by rubbing certain substances together, electrical phenomena have been studied for many years. During 1874, the Irish physicist George Stoney estimated the unit charge of electricity from the application of Faraday's laws, and later named the 'atom of electricity' the **electron** (from the electric ion). At the time it was unclear if there were both positive and negative charges to carry electricity, and if these charges could be separated from atoms. Two key discoveries from the Cavendish Laboratory of Cambridge University helped to further our understanding of electrons and atomic structure.

Prior to these discoveries, the atom was thought to be indivisible, known as the solid-ball model. In fact the word atom comes from the Greek word *atomos*, which means indivisible. These new discoveries showed that atoms were actually composed of smaller components.

CATHODE RAYS

A cathode ray tube (Figure 14.1.1) is a sealed glass tube from which most of the air has been evacuated, with two electrodes inside arranged as an electron gun

(Figure 14.1.2). Electrons are released from a negative terminal, or **cathode**, in a vacuum, and accelerated across a potential difference towards a positive terminal, or **anode**.

In Chapter 5 it was shown how a charged particle will act within an electric potential. A negative charge will move towards a positive plate, while a positive charge will move towards a negative plate (Figure 14.1.3).

GO TO ➤ Section 5.1 page 147

In 1897 Joseph John (J. J.) Thomson was studying electric discharges from cathode ray tubes. Thomson found that the tube glowed more when the gas was at a lower pressure. This showed that something was exciting the gas. When the tube was painted with phosphorescent paint, the paint opposite the negatively charged cathode sparked and glowed.

Thomson placed two electrically charged plates above and below the cathode ray tube, and the cathode ray deflected towards the positively charged plate, indicating the beam was negatively charged. He also changed the material of the cathode, and the same beam emerged. Thomson concluded that the beam was composed of a stream of negatively charged particles, but he could not be certain if these were charged molecules, atoms, or something smaller.

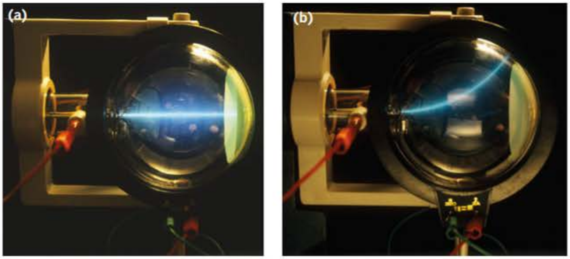


FIGURE 14.1.1 A cathode ray tube, formed from a glass ball with a low-pressure gas. (a) The electron gun produces a beam, visible as it ionises the gas which emits photons. The beam excites the phosphor. (b) When a magnetic field is applied, the cathode ray deviates at some angle.

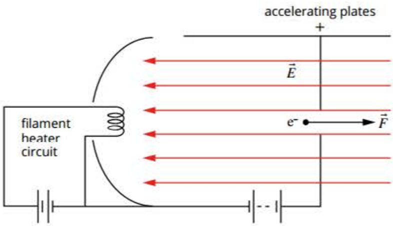


FIGURE 14.1.2 An electron gun. The negatively charged heater circuit emits electrons, which accelerate towards the anode. The electrons emerge through the gap in the anode.

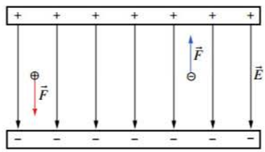


FIGURE 14.1.3 The direction of the electric field *E* indicates the direction in which a force would act on a positive charge. The positive charge is attracted to the negative plate; the negative charge is attracted to the positive plate.

CHARGE-TO-MASS RATIO

In 1897, Thomson demonstrated that cathode ray particles are fundamental constituents of every atom. To indicate their importance, cathode rays were renamed electrons.

Thomson's experiment was to try to measure the mass of the particle (Figure 14.1.4). He could not directly measure its mass or charge, but the degree of deflection of the beam would reveal the ratio of the charge of the particle to its mass, e/m. (In this case the charge is represented by e, as the charge of the electron, rather than the usual symbol for charge q.)

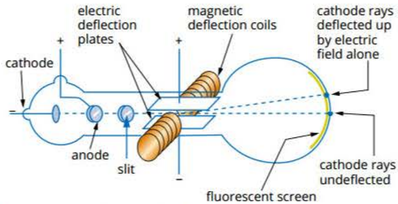


FIGURE 14.1.4 Thomson's apparatus for finding the charge-to-mass ratio for cathode rays. Thomson used an electron gun and cathode ray tube to measure the degree of deflection of the beam.

The experiment was performed in two stages. At first the forces on a beam of electrons were balanced using an electric and a magnetic field, as shown by the central dotted line striking the fluorescent screen in Figure 14.1.4. This enabled Thomson to find the speed of the electrons. Then the magnetic field was switched off, and the beam was deflected under the influence of the electric field alone, as shown by the upper dotted line in Figure 14.1.4. The deflection of the beam was measured, allowing Thomson to find the charge-to-mass ratio for the cathode rays. Thomson repeated the experiment with a variety of different cathodes to show that all cathode rays yielded the same value. His result produced a value for the ratio of about 1×10^{11} C kg^{-1}; the accepted value today is 1.76×10^{11} C kg^{-1}.

This was astonishing. In the previous year the German physicist Emil Wiechert had measured the e/m ratio for charged hydrogen atoms (what we would now call H^+ or a **proton**), and Thomson's result for the ratio of the cathode particles was more than one thousand times larger. Either cathode rays carried an enormous charge or they were very light compared to their charge.

MOTION OF PARTICLES IN FIELDS

In Chapter 5, the motion of charged particles in electric and magnetic fields was analysed. The following equations were derived, which are also useful when investigating Thomson's charge-to-mass experiment.

The force on a charge in an electric field: $\vec{F} = q\vec{E}$

The electric field strength between two charged plates a distance d apart: $E = \frac{V}{d}$

The final velocity of a charge accelerated from rest through an electric field is given by the electron-gun equation: $K = \frac{1}{2}mv^2 = qV$

The force on a charge moving within a magnetic field: $F = qvB\sin\theta$

If a moving charge experiences a force of constant magnitude that remains at right angles to its motion, it will experience circular motion. Thus, the electrons will follow a curved path of radius r. In this case, the net force acting on the charge is given by Newton's second law: $\vec{F}_{net} = m\vec{a}$. This is equivalent to the magnetic force on the charge, so that

$$qvB = ma$$

The acceleration in this situation is centripetal (towards the centre of the circular path) and has the magnitude

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

Substituting this relationship into the previous equation and rearranging gives an expression that predicts the radius of the path of an electron travelling at right angles to a constant magnetic field. This relationship can be used to calculate the radius of the path followed by a charged particle travelling at right angles to any magnetic field.

It is assumed here that the electrons are not travelling at relativistic speeds. The effects of relativity on the motion of particles are discussed in greater detail in Chapter 12.

A charged particle moving in a magnetic field turns with a radius, r, given by:

$$r = \frac{mv}{qB}$$

where

r is the radius of the path (m)

m is the rest mass of the particle (kg; for an electron $m_e = 9.109 \times 10^{-31}$ kg)

v is the speed of the charge (m s^{-1})

q is the charge on the particle (C; for an electron $q_e = -1.602 \times 10^{-19}$ C)

B is the strength of the magnetic field (T)

Worked example 14.1.1

MOTION OF A CHARGED PARTICLE

A 500.0 V electrical field is applied to an electron that is initially at rest.

a Calculate the final speed of the electron.

Thinking	Working
Write out the information relevant to the final velocity. Note that the particle starts at rest, so its initial velocity is zero.	$q = 1.602 \times 10^{-19}$ C $V = 500.0$ V $m = 9.109 \times 10^{-31}$ kg $u = 0\text{ m s}^{-1}$ $v = ?$
Select the equation that best fits the information you have.	$\frac{1}{2}mv^2 = qV$
Rearrange, substitute values and solve for v.	$v = \sqrt{\frac{2qV}{m}}$ $= \sqrt{\frac{2 \times 1.602 \times 10^{-19} \times 500}{9.109 \times 10^{-31}}}$ $= 1.326 \times 10^7\text{ m s}^{-1}$

b The electron enters a magnetic field of 30.0 mT and moves in a curved path. Calculate the radius of curvature of the electron's path.

Thinking	Working
Write out the information relevant to the final radius.	$q = 1.602 \times 10^{-19}$ C $B = 30.0\text{ mT} = 0.030$ T $m = 9.109 \times 10^{-31}$ kg $v = 1.33 \times 10^7\text{ m s}^{-1}$
Select the equation that best fits the information you have.	$r = \frac{mv}{qB}$
Substitute values and solve for r.	$r = \frac{9.109 \times 10^{-31} \times 1.326 \times 10^7}{1.602 \times 10^{-19} \times 0.030}$ $= 0.00251$ $= 2.51 \times 10^{-3}$ m

Worked example: Try yourself 14.1.1

MOTION OF A CHARGED PARTICLE

The Australian synchrotron accelerates electrons using an electron gun at 90 kV.

a Calculate the final velocity of the electrons.

b The electrons enter a circular storage ring with a magnetic field of strength 30 μT. Calculate the radius of curvature of the electrons' path.

PHYSICSFILE WE ICT

Water-drop experiment

Before Millikan's oil-drop experiment, physicist J. J. Thomson had completed a similar investigation instead using water droplets.

Water is a more volatile substance than oil, i.e. it evaporates more easily, so that the mass and size of the droplets change rapidly. In order to improve on the accuracy of the results, the mass of the droplets needed to be measured to a high degree of precision, so water was replaced with oil.

THE OIL-DROP EXPERIMENT

In 1909 at the University of Chicago, Robert Millikan and his student Harvey Fletcher created the oil-drop experiment to measure the charge on the electron.

A spray of oil drops was introduced between two metal plates (Figure 14.1.5). When an electrical potential was applied to the metal plates, charged drops became suspended between them. To do this, the drop must have been in equilibrium—the magnitude of the upward force (due to the electrostatic force experienced by the charge in the electrical field) balanced the downward force (due to gravity). This gives the equation $qE = mg$.

Some drops would become inadvertently charged due to friction from the oil spray device. By varying the degree of ionisation or irradiating the air with X-rays to induce more ionisation, the drops could be charged to higher levels, and therefore a weaker electrical field would be needed to suspend the same sized drop.

However, Millikan found that the charge on a drop was always a multiple of 1.6×10^{-19} C. This showed the smallest unit of charge must be equal to the charge of a single electron. (The modern value of the charge on the electron is -1.602×10^{-19} C, less than a 1% difference from Millikan's calculation).

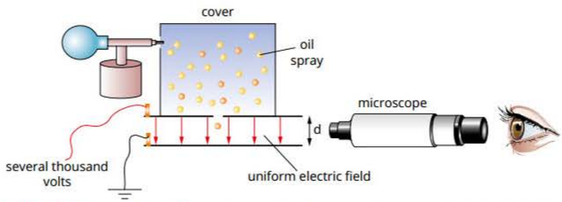


FIGURE 14.1.5 Schematic of Millikan's experiment. As the charged drops enter the electric field, the upward electrostatic force balances out the downward weight force due to gravity.

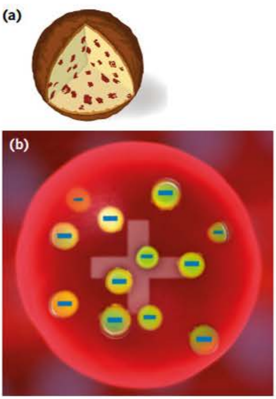


FIGURE 14.1.6 (a) The archaic dessert, the plum pudding. (b) Thomson's atomic model. Discrete elections are distributed in a diffuse, positively charged space. This model met with the observations, although the nature of the positive-charged part of the atom remained poorly defined.

Physicists were now able to combine the experimental result for the charge on an electron with its charge-to-mass ratio so that its mass could be calculated and is now known to be $m_e = 9.109 \times 10^{-31}$ kg.

AN ATOMIC MODEL

Thomson also knew that atoms were electrically neutral, and therefore there should also be a positively charged component within atoms to balance out the negative charge from electrons. In 1904 he proposed an **atomic model** to address this idea. It consisted of a diffuse ball of positive charges with negatively charged particles floating within, and Thomson named it the plum pudding model (Figure 14.1.6).

In the same year, Hantaro Nagaoka at the University of Tokyo proposed the Saturnian model in which the negative charged particles orbited around the outside of a positively charged centre, like the rings around the planet Saturn.

+ ADDITIONAL

The Cavendish Laboratory

The Cavendish Laboratory was a key institution for early research in atomic physics. The Cavendish Laboratory is a part of the University of Cambridge, and opened in 1874. Cambridge is a collegiate university, effectively an alliance of teaching and research institutions.

Up until the early 19th century, physics was mainly theoretical and was regarded as the province of mathematicians. The rise of an industrial society meant there was a need for a focal point for training scientists and engineers in experimental physics, rather than research by individuals or small collaborations.

James Clerk Maxwell, of electromagnetism fame, was the university's first Cavendish Professor of Experimental Physics. He was succeeded in 1879 by John Strutt, Lord Rayleigh, then the long tenure of Thomson (Figure 14.1.7a) from 1884 to 1919. The early years of the Cavendish produced a number of the next generation of researchers, such as Ernest Rutherford (Figure 14.1.7b; atomic structure) and Charles Wilson (the cloud chamber). The Australian William Bragg (X-ray crystallography) succeeded Rutherford as director, and during Bragg's directorship Francis Crick and James Watson determined the structure of DNA.

Known discoveries from Cavendish include the electron (1897), nuclear disintegration (1900), X-ray diffraction (1912), chemical isotopes (1920), the ionsphere (1924), electron diffraction (1927), the neutron (1932), superfluidity (1937), DNA structure (1953), haemoglobin structure (1959), neutron stars (1968), polymer LEDs (1990), and more. Practical companies and devices spun out of the Cavendish include early television and radio manufacturers (1920s), aircraft glues (1938) and CT scanners (1956). There are few better examples for the benefit of 'blue sky' research to bring a range of often unexpected practical outcomes than the Cavendish.

Of course the Cavendish was not the only group in early atomic research. The universities of Manchester (UK), Göttingen and München and the Kaiser Wilhelm Institute (now the Max Planck Institute) in Berlin (Germany), and the universities of Paris (France), Chicago (USA) and Copenhagen (Denmark), and—later into the 1930s following the flight of many scientists from Europe—Cornell, Caltech and Princeton (USA) all included researchers who made significant contributions. At these handful of institutions scientists with the most advanced thinking could easily share their ideas and collaborate.

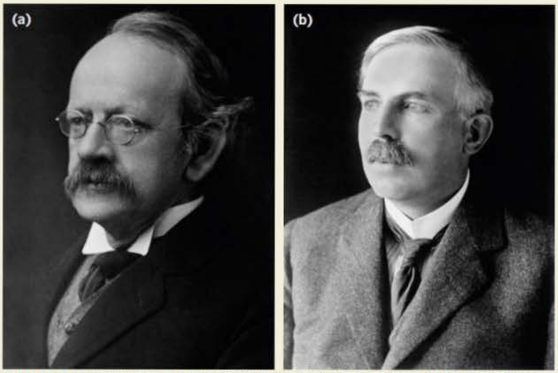


FIGURE 14.1.7 (a) Thomson and (b) Rutherford, two leaders in early studies of the atom and two directors of the Cavendish Laboratory.

14.1 Review

SUMMARY

- A cathode ray tube creates a negatively charged beam, whose component electrons have a kinetic energy and velocity described by the electron-gun equation, $K = \frac{1}{2}mv^2 = qV$.
- Charged particles in a magnetic field move in a circular path described by $r = \frac{mv}{qB}$.
- Thomson demonstrated an atom is composed of negatively and positively charged components, and that the negatively charged component is substantially lighter in mass than the positively charged component.
- Millikan's oil-drop experiment quantified the charge on the electron as -1.602×10^{-19} C. This led to the calculation of the magnitude of the positive charge (i.e. the charge of a proton) and the mass of the electron.
- The mass of an electron, m_e, is 1.109×10^{-31} kg.

KEY QUESTIONS

To answer these questions use $q_e = -1.602 \times 10^{-19}$ *C and* $m_e = 1.109 \times 10^{-31}$ *kg.*

1 An electron-gun assembly emits electrons with energies of 10.0 keV.

a Calculate the magnitude of the predicted exit velocity of the electrons. (Ignore any effects of relativity.)

b Upon exiting the electron-gun assembly, the electrons enter a uniform magnetic field of 1.50 mT oriented perpendicular to their motion. Calculate the predicted radius of the electron beam.

2 A muon, a particle of mass 1.88×10^{-28} kg and the same charge as an electron, encounters the same 10.0 kV accelerating voltage and 1.50 mT magnetic field as in Question 1. What is its exit velocity and the radius of curvature of its path?

3 An earlier version of Millikan's oil-drop experiment used water, which is a more volatile substance. Why does this make it less preferable than oil for this experiment?

4 Research online or in other appropriate sources in order to complete the following table.

Atomic model	Description
solid-ball model	
plum pudding model	
nuclear model	
planetary model	

14.2 Nuclear model of the atom

After New Zealander Ernest Rutherford visited the Cavendish Laboratory and contributed to Thomson's research on the electron, his research focus shifted to radioactivity. His work on radioactivity was to win him the Nobel Prize in Chemistry in 1908.

GOLD FOIL EXPERIMENT

Rutherford had found earlier that **alpha particles** (later found to be composed of two protons and two neutrons) were produced from the decay of many radioactive isotopes, and could be deflected in electric and magnetic fields. Rutherford conducted an experiment where he measured this deflection by finding the final positions of the alpha particles on a piece of photographic film. If the alpha particles first passed through a thin slice of mica (a mineral), they were deflected and the images on the photographic film were blurred.

This result made sense from Thomson's plum pudding model. If a fast-moving alpha particle collided with an atom in the mica slice, it would be affected only by the atom's electrical field. As the alpha particle was much heavier than an electron and moving rapidly, the particle would be scattered (deflected) by a small amount, enough to blur the image.

Rutherford's associate Hans Geiger improved this experiment by replacing the mica slice with gold foil (Figure 14.2.1). Because gold could be beaten into a film comprising only 400 atoms in width, the experiment would produce a larger angle of deflection. Geiger also added a fluorescent screen to detect the particles. When a particle hit the screen it gave a tiny flash of light, so that these experiments needed to be conducted in the dark. Geiger next developed a device that converted the radiation into an electrical signal, which meant the experiment could be done under normal light. In any event most of the alpha particles were scattered by less than 2°.

In 1909 Geiger and his student Ernest Marsden explored the idea of looking for alpha particles scattered at large angles. Rutherford didn't think it likely that there would be any scattering; in science a negative result can be just as helpful to test a theory. After just a few days Geiger and Marsden found evidence of large-angle scattering, with some particles bouncing back nearly to the source! About one in 10 000 particles were found to have an angle of deflection greater than 90°.

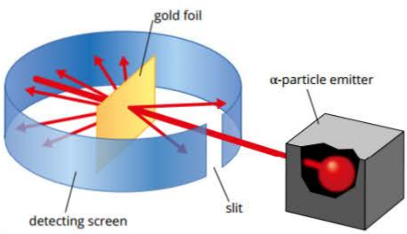


FIGURE 14.2.1 Schematic of the gold foil experiment.

RUTHERFORD'S ATOMIC MODEL

The gold foil experiment made sense if the mass and electric charge were concentrated in the centre of the atom, known as an atomic **nucleus**. Rutherford proposed an atomic model of a charged nucleus surrounded by a cloud of electrons.

PHYSICSFILE ICT

The Geiger–Müller counter

In 1928, some 20 years after the nuclei experiments, Geiger and his student Walther Müller improved on Geiger's original idea to measure radioactivity, and developed the Geiger–Müller tube (G–M tube), known commonly today as the Geiger counter and the principle tool to measure radioactivity.

A Geiger counter consists of a Geiger–Müller tube filled with argon gas, as shown in Figure 14.2.2.

A voltage of about 400 V is maintained between the positively charged central electrode and the negatively charged aluminium tube. When radiation enters the tube through the thin mica window, the argon gas becomes ionised and releases electrons. These electrons are attracted towards the central electrode and ionise more argon atoms along the way. For an instant, the gas between the electrodes becomes ionised enough to conduct a pulse of current between the electrodes. This pulse is registered as a count. The counter is often connected to a small loudspeaker so that the count is heard as a 'click'.

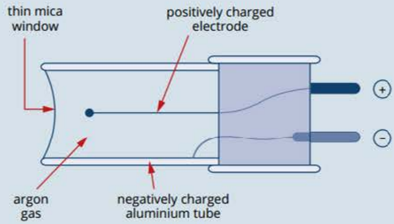


FIGURE 14.2.2 A schematic diagram of a Geiger counter used for detecting ionising radiation.

Accompanying the model, Rutherford developed an equation to predict the scattering angles, which was supported by further experimental testing by Geiger and Marsden (Figure 14.2.3). Using different metals, they showed that the ability of the nucleus to block the alpha particle was not greatly affected by the atomic mass. They had found that the nucleus occupied less than 10^{-14} m; that is, the atom was mostly empty space. This is what enables the alpha particles to pass through the foil undetected.

By 1913 the scattering experiments also demonstrated that the nucleus was positively charged. None of these observations matched the Thomson model, and the plum pudding model could then be rejected.

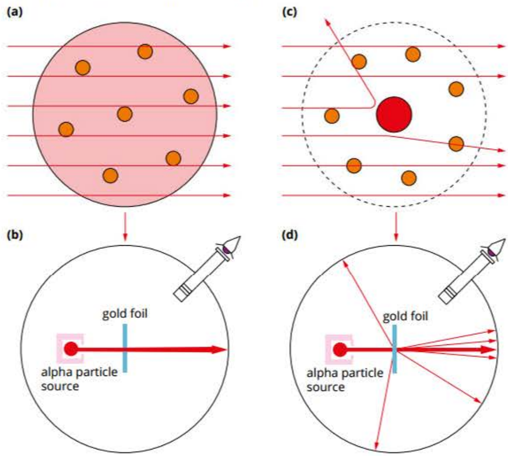


FIGURE 14.2.3 Explaining the Geiger–Marsden experiment. (a) In Thomson's atomic model the alpha particles travel through the atom, affected only weakly by electrical forces, and (b) only low angle scattering is possible. (c) In Rutherford's atomic model most alpha particles travel through the atom, but some impact the nucleus and scatter, (d) which is observed in the gold foil experiment.

All scientific models reflect the evidence available at the time, and Rutherford's atomic model was no exception. While the model explained scattering and the location of the protons, it did not address what stops the negatively charged electrons from combining with the positively charged nucleus. The model assumes that the electrons orbit the nucleus, like planets around the Sun. Newtonian mechanics requires that as the electrons would lose energy they would eventually spiral into the nucleus. The energy that the electrons emit should be detectable as a continual outflow of electromagnetic radiation. The model could also not explain the atomic absorption and emission spectra. The answer to these would come with Niels Bohr's atomic model, discussed in greater detail in Chapter 15.

GO TO ➤ Section 15.1 page 392

DISCOVERY OF THE NEUTRON

Rutherford had recognised another problem with the nuclear model. If the nucleus contains positively charged protons, electrical repulsion should not allow stable atomic nuclei. While collaborating with Niels Bohr during the early 1920s, he developed two hypotheses to explain this. The first hypothesis was that an uncharged particle must be exerting an attractive force within the nucleus to bind the protons. He dubbed this theoretical particle the '**neutron**'. The second hypothesis was that there existed electrons constrained in the atomic nucleus, 'nuclear electrons', to shield and reduce the electrical repulsion between the protons. How electrons could enter the nucleus would need to be explained. However, during some radioactive processes, electrons were known to be emitted, meaning the nuclear electron hypothesis was the stronger choice.

Further developments in quantum theory provided several objections to the nuclear electron hypothesis. The ideas of wave–particle duality implied that electrons could not be constrained to a space as small as an atomic nucleus without requiring impossibly high energy barriers. Wave–particle duality will be discussed in greater detail in Chapter 15.

In 1931, Walther Bothe and Herbert Becker in Berlin observed that when light elements such as lithium, beryllium and boron were bombarded by alpha particles, a new kind of radiation emerged. This radiation was unusually penetrating and was unaffected by an electric field. The assumption was that the emissions were gamma rays. In the next year, Irène Joliot-Curie and Frédéric Joliot in Paris used alpha particles from a polonium source on hydrocarbons, such as paraffin wax, to generate Bothe's radiation, and found that the new radiation caused high-energy protons to be ejected with energies of 5.3 MeV. They captured photographic evidence for the proton ejection, but unfortunately did not interpret the data correctly (Figure 14.2.4).

FIGURE 14.2.4 The experiment and first photographic evidence for the existence of the neutron, taken by Irène Joliot-Curie and Frédéric Joliot.

James Chadwick, one of Rutherford's former students and then his collaborator at the Cavendish Laboratory, investigated collisions between alpha particles and the element beryllium. The principle of conservation of momentum was used to interpret the data from the Curie–Joliot experiments.

Chadwick had studied gamma rays as a student and knew that protons were too heavy to be dislodged by gamma rays. He calculated that the radiation consisted of particles with a mass close to the proton's mass, but without an electric charge—the neutron. Subsequent investigations confirmed his experiments. With no electric charge the neutron could not be investigated through the interactions of charged particles with electric fields.

With Chadwick's announcement, models for atomic nuclei composed of protons and neutrons soon followed. Proposed models included arranging the protons and neutrons packed in together in the nucleus like in Figure 14.2.5; or a shell model, in which the particles are arranged within the atomic nucleus in a specific order akin to the defined arrangements of electrons around an atomic nucleus.

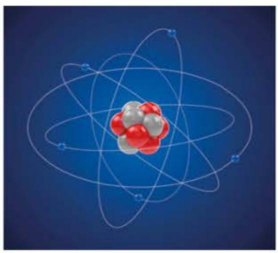


FIGURE 14.2.5 From Chadwick's discovery, the model of the atomic nucleus now included both protons and neutrons.

PHYSICSFILE

Neutron decay

In 1934 Enrico Fermi finally excluded the nuclear electron hypothesis. He explained that the beta particles observed from radioactive processes were from the radioactive decay of a neutron. Theory suggested that a neutron would decay into a positively charged proton, a negatively charged electron, sometimes gamma radiation, and a then-theoretical particle (called the neutrino) to preserve energy and momentum considerations; experimental proof of the idea took another 20 years.

PHYSICS IN ACTION ICT

Uses of the neutron

Neutrons are very penetrating because they are uncharged, and they can be fired into an atomic nucleus without being repelled.

A neutron can even be absorbed by an atomic nucleus. If sufficient neutrons enter a nucleus, the nucleus can become destabilised and fission (split), in a process known as atomic fission. When the nucleus fissions, energy is released. If at least one neutron is released, a second fission reaction may be triggered. This is the basis for nuclear reactors. If a substantial number of free neutrons are produced, these neutrons can be used for other purposes. Uses of neutrons include the following:

- *Neutron diffraction* to probe the structure and arrangements in materials, much like X-ray diffraction. Neutron diffraction studies provide similar structural information from X-ray diffraction; however, lighter atoms like hydrogen that are invisible to X-ray scattering can be detected.
- *Neutron spectroscopy* to probe nuclear vibrations.
- *Neutron chemical analysis* to detect elements in a sample. Neutron activation analysis characterises the atomic nucleus and ignores the chemical nature of an element (bonding, material state). No sample preparation is needed, and it is useful in 'messy' industrial, geological or forensic samples.
- *Neutrons in medicine* provide more energy to a focused area than electromagnetic radiation treatments for cancers. A beam of low-energy neutrons can also be used once a patient has been treated with boron-10—the boron captures the neutrons and decays to lithium-7 with the production of an alpha particle with enough energy to kill the cancer.

There are, of course, limits to the use of neutrons. The production of neutrons generally requires a nuclear reactor. In Australia the OPAL research reactor at ANSTO remains the only source of research neutrons. Neutron irradiation has a tendency to render the samples somewhat radioactive if certain less-stable elements are present.

+ ADDITIONAL

Discovery of the positron

Following the discovery of cosmic rays—high-energy protons and atomic nuclei—in the 1930s, a young physicist called Carl Anderson built an improved version of the cloud chamber that was being used by researchers to study the tracks of ionising radiation like cosmic rays. By placing his cloud chamber in a magnetic field, Anderson photographed the curved path of charged ionising particles (like protons and electrons) that were created by bombarding cosmic rays.

Perhaps due to the improvements he had made to the design and the composition of the vapour, Anderson could collect very clear photographic evidence (Figure 14.2.6). In one of his photos, Anderson noticed that one particle, with the same mass as an electron, curved in the opposite direction to what was expected due to the direction of the magnetic field it was passing through.

After further refinements of his experiment to exclude the possibility that the particle was a proton, Anderson and his colleagues concluded that it was indeed a positively charged equivalent of an electron. They dubbed it the positron. The antimatter positrons that Anderson discovered were created when a cosmic ray spontaneously created an electron–positron pair. For this discovery, Anderson was awarded the 1936 Nobel Prize in Physics at the age of just 31.

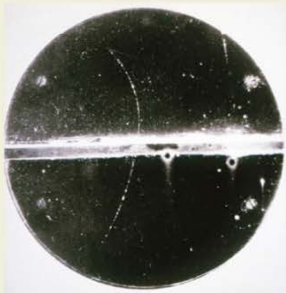


FIGURE 14.2.6 The first photographic evidence for the existence of the positron, taken by Carl Anderson. The particle is known to be positive because of the direction of the curve, and the radius of the curve shows that its mass is the same as that of an electron.

14.2 Review

SUMMARY

- The Geiger–Marsden scattering experiment showed that the mass of an atom is not distributed throughout the space occupied by the atom but is rather constrained to the nucleus.
- Rutherford proposed a nuclear model of the atom to replace Thomson's plum pudding model.
- Rutherford's nuclear model did not explain electron arrangement or how positive charges in the nucleus did not mutually repel.
- Chadwick used conservation of momentum to interpret experimental results to discover the neutron.

KEY QUESTIONS

1 Predict the paths that would be taken by alpha particles that were fired at atoms with a Thomson's plum pudding model structure. Explain why you would expect the alpha particles to take these paths.

2 If alpha particles of higher energy than those in Question 1 are fired at plum pudding atoms, predict how their paths will differ from the lower-energy alpha particle paths.

3 The Geiger–Marsden experiment determined that the nucleus has a radius of less than 10^{-14} m. Why could the experiment determine a maximum size and not an actual size of the nucleus?

4 Which has a greater mass: a neutron or a proton?

Chapter review

14

KEY TERMS

alpha particles
anode
atomic model
cathode
electron
neutron
nucleus
proton

REVIEW QUESTIONS

To answer the following questions, use electron mass $m_e = 9.109 \times 10^{-31}$ kg and charge $q_e = -1.602 \times 10^{-19}$ C.

1 Why are atoms electrically neutral?

2 In a cathode ray tube the electrons come off the cathode and travel towards the anode. Why do the electrons not simply stop at the anode?

3 In a cathode ray tube, what provides the centripetal acceleration to change the direction of the electrons in the cathode ray?

4 Electrons in a cathode ray tube are accelerated through a potential difference of 2.5 kV. Calculate the speed at which they hit the screen of the cathode ray tube.

5 Electrons in a cathode ray tube are accelerated through a potential difference from a cathode to a screen. Calculate the speed at which they hit the screen if the potential difference between the electrodes is 4.5 kV.

6 **a** Calculate the force exerted on an electron travelling at a speed of $6.4 \times 10^6\,\text{m s}^{-1}$ at right angles to a uniform magnetic field of strength 9.1×10^{-3} T.

b Calculate the radius of the electron's path.

7 An electron with a speed of $4.3 \times 10^6\,\text{m s}^{-1}$ travels through a uniform magnetic field and follows a circular path of diameter 8.4×10^{-2} m. Calculate the magnitude of the magnetic field through which the electron travels.

8 In an experiment similar to Thomson's for determining the charge-to-mass ratio (e/m) of cathode rays, electrons travel at right angles through a magnetic field of strength 1.50×10^{-4} T. Given that they travel in an arc of radius 6 cm and that $e/m = 1.76 \times 10^{11}\,\text{C kg}^{-1}$, calculate the speed of the electrons.

9 In the Millikan oil-drop experiment, the magnitude of the upward force of electrical attraction is balanced by the downward force of gravity, at which $qE = mg$ (where q is the charge on an electron, V is the electric potential, E is the magnitude of the electric field strength, m is the mass of the oil drop and g is the magnitude of the acceleration due to gravity).

a Derive an expression for the charge in terms of the variables d (the distance between the two charged plates), m, g and V.

b The mass, m, of an oil drop is difficult to measure, but its size (radius r) is not. Given oil of density ρ, express the above expression for q in terms of ρ, r, d, g and V. (Assume the drops of oil are spherical.)

c Oil droplets, all of density $900\,\text{kg m}^{-3}$ and radius $0.50\,\mu\text{m}$, are injected into an oil-drop experiment with plates 5.0 cm apart. What difference in electrical potential will suspend the droplets between the plates? (Assume the droplets all have the same charge q_e).

d In Millikan's oil-drop experiment, the charges on the drops were not all equal. Calculate the electrical potential needed to suspend a droplet with a charge of $3q_e$.

10 Order the work of the following scientists from earliest to latest.

A Ernest Rutherford

B Niels Bohr

C Joseph John Thomson

11 What did Ernest Rutherford discover about the arrangement of mass in an atom?

12 Of the three common particles—electrons, protons, neutrons—which one could not be explained by Rutherford's atomic model? What was the key evidence?

13 What might Rutherford have concluded if Geiger and Marsden had not detected any backward scattering?

14 Using the Sydney Cricket Ground as a model for an atom, consider a capacity crowd of 51 436 fans that are either supporting Team Proton or Team Electron. If the fans are apportioned according to the relative masses of the electron and proton, how many 'Electron' and 'Proton' fans are there?

15 While atoms are now known to be divisible, what changes to the structure of the atom occur when it is divided?

16 What is the main difference between protons and neutrons?

17 An electron travels through a uniform magnetic field of 9.4×10^{-4} T and follows a circular path of diameter 9.2×10^{-2} m. Calculate the tangential velocity of the electron.

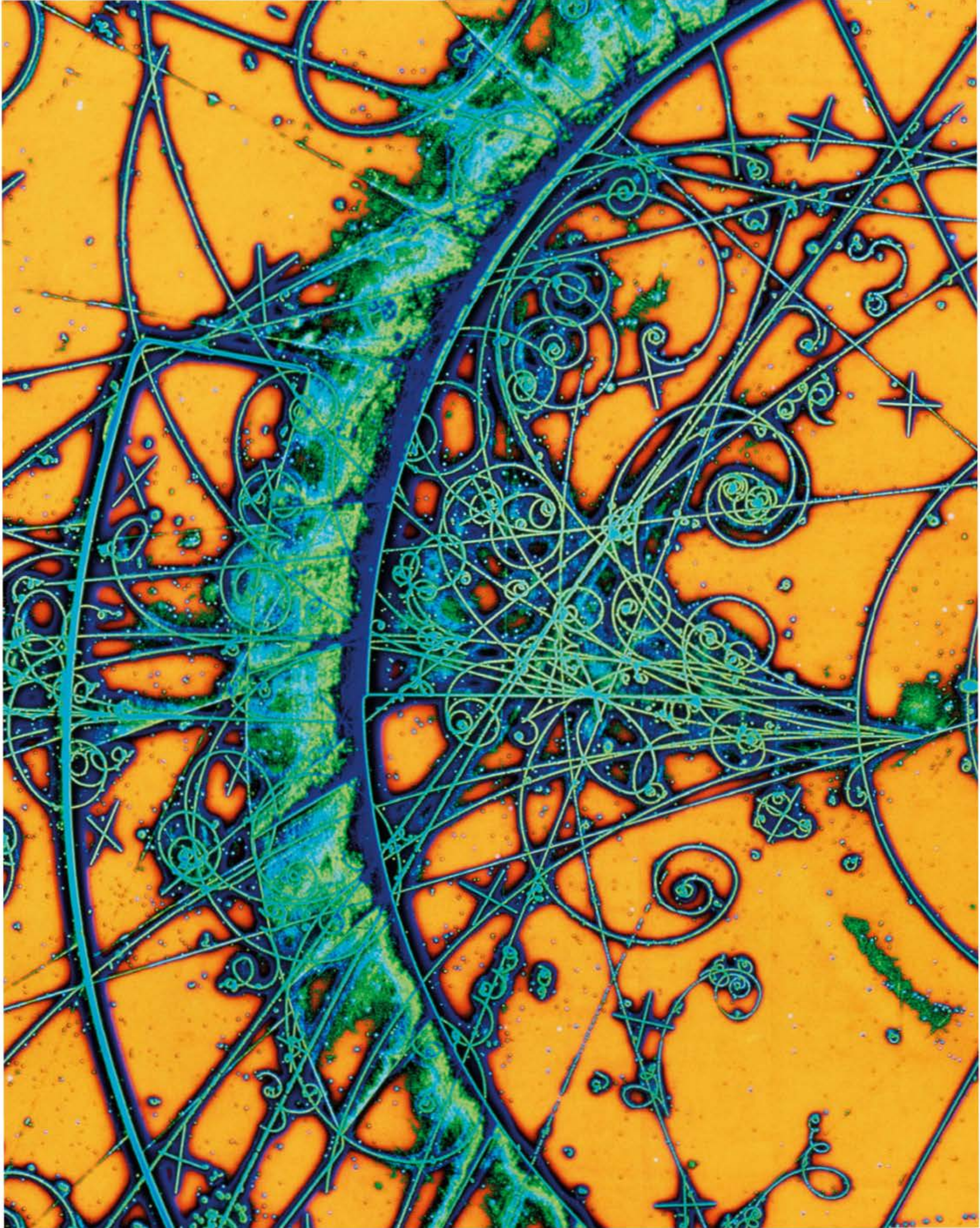
18 A muon, a particle of mass 1.88×10^{-28} kg and the same charge as an electron, is accelerated through a 5.00 kV potential difference before it encounters a magnetic field of strength 1.55 mT. What is its velocity as it enters the magnetic field and its resultant radius of curvature?

19 A proton 'gun' accelerates protons by means of a potential difference of 3.5 kV. The protons follow a circular arc with a radius of curvature of 5.0 cm. ($m_p = 1.67 \times 10^{-27}$ kg, q_p 1.60×10^{-19} C).

a Find the speed at which the protons enter the cyclotron.

b Determine the strength of the magnetic field in the cyclotron.

20 After completing the activity on page 376, reflect on the inquiry question: How is it known that atoms are made up of protons, neutrons and electrons?



CHAPTER 15

Quantum mechanical nature of the atom

The nuclear model of the atom suggested that a positive nucleus was surrounded by electrons. This work was pioneered by Thomson and Rutherford. Through further studies and investigations, Bohr, de Broglie and Schrödinger showed that a better method for understanding the structure of the atom was through utilising quantum mechanics.

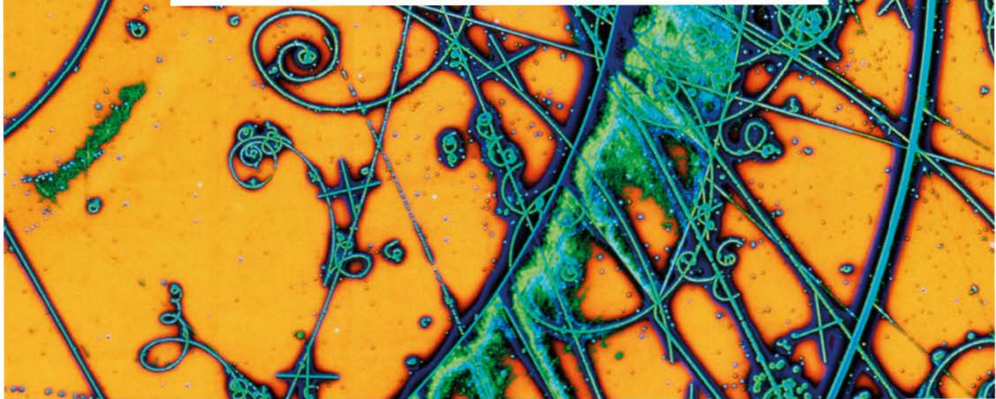
Content

INQUIRY QUESTION

How is it known that classical physics cannot explain the properties of the atom?

By the end of this chapter you will be able to:

- assess the limitations of the Rutherford and Bohr atomic models ICT
- investigate the line emission spectra to examine the Balmer series in hydrogen (ACSPH138) ICT
- relate qualitatively and quantitatively the quantised energy levels of the hydrogen atom and the law of conservation of energy to the line emission spectrum of hydrogen using:
  - $E = hf$
  - $E = \frac{hc}{\lambda}$
  - $\frac{1}{\lambda} = R\left[\frac{1}{n_f^2} - \frac{1}{n_i^2}\right]$ (ACSPH136) ICT N
- investigate de Broglie's matter waves, and the experimental evidence that developed the following formula:
  - $\lambda = \frac{h}{mv}$ (ACSPH140) ICT N
- analyse the contribution of Schrödinger to the current model of the atom.



15.1 Bohr model

PHYSICS INQUIRY N CCT

Electron probability

How is it known that classical physics cannot explain the properties of the atom?

COLLECT THIS...

- marker pen
- A3 plain paper
- exercise book or other item to cushion the fall

DO THIS...

1. Draw a circle of radius 1 cm in the centre of the paper. Around this circle draw a series of concentric circles with their radii increasing by 1 cm each. This will be the target for the activity.
2. Place the paper on the exercise book.
3. Stand above the target with your arm outstretched holding the felt tip marker directly above the centre circle.
4. Drop (do not throw) the marker so it lands on the target and makes a mark. Repeat to create 100 marks.
5. Count the number of marks in each circle for both drop heights. If a mark lands on the border, include it in the circle that most of it sits in. If it is exactly on the line, include the mark in the count for the smaller circle.

RECORD THIS...

Describe the distribution of the position of the marks for both drop heights, relating it to the probability of the pen being in that position.

Present a table of the number of dots per cm^3 in each circle. Plot a graph with the ring number on the *x*-axis and the probability on the *y*-axis.

REFLECT ON THIS...

How is it known that classical physics cannot explain the properties of the atom?

How does the two-dimensional area of your probability compare with the three-dimensional orbital of hydrogen?

What could you do to improve your probability distribution?

In the previous chapter, the Rutherford atomic model was introduced. His nuclear model of the atom suggested that electrons surround the nucleus of the atom, and that the electrons revolve around the nucleus in circular paths which he named **orbits**. Although Rutherford's model was based on experimental observations, it failed to explain certain things.

Rutherford's model proposed that electrons revolve around the nucleus in circular orbits. According to the law of conservation of energy, the orbiting electrons should emit energy due to their motion. That energy loss would in turn cause the circular orbit to shrink. While Rutherford's model of circular orbits did explain how electrons could stay away from the nucleus, the stability of an atom could not be explained. Further limitations of Rutherford's model were related to the splitting of spectral lines (the Zeeman effect), the existence of faint spectral lines, and the fact that the model could only really be accurately applied to single-electron atoms (e.g. hydrogen). For these reasons, a more complex quantum approach was required.

EMISSION SPECTRA AND ENERGY LEVELS IN ATOMS

Towards the end of the nineteenth century, scientists had devised a variety of ways of making atoms produce light. These methods included:

- heating substances until they glowed
- applying high voltages to gases in glass tubes causing the gas to glow
- burning salts in gas flames causing the salt to produce a bright flash of light.

When the light that was emitted from the atoms was analysed using a spectroscope, a distinctive emission spectrum was observed for each different atom. Thus the emission spectrum of an atom became a unique property of the atom. See, for example, the two emission spectra in Figure 15.1.1, which can be used to distinguish between sodium and mercury.

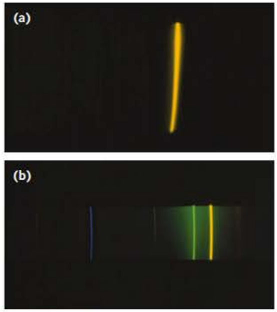


FIGURE 15.1.1 The emission spectra of (a) sodium and (b) mercury.

Recall that white light is actually made up of an infinite number of different frequencies (or wavelengths) of light. If white light from an incandescent light globe is passed through a spectroscope, a continuous rainbow of colours is seen. The rainbow will contain all the shades of the visible light spectrum from red to violet. The emission spectra for sodium and mercury in Figure 15.1.1 do not show continuous rainbows, just some specific colours (frequencies). For many years, scientists could not explain why atoms emitted only discrete frequencies of light rather than continuous spectra.

In 1912, Niels Bohr devised a sophisticated model of electron **energy levels** for the atom. He was later awarded a Nobel Prize in Physics for this work. Energy levels can be shown as horizontal lines on a graph. The graph in Figure 15.1.2 shows the energy levels for sodium gas.

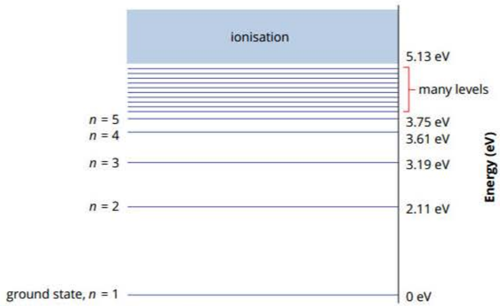


FIGURE 15.1.2 Energy levels in a sodium atom

In Figure 15.1.2, when an electron receives enough energy so that ionisation occurs (i.e. $E = 5.13$ eV for the sodium atom), this means that the electron will have enough energy to escape the atom.

Bohr's main ideas were as follows.

- The electron moves in a circular orbit around the nucleus of an atom.
- The force keeping the electron moving in a circle is the electrostatic force of attraction (the positive nucleus attracts the negative electron).
- A number of allowable orbits of different radii exist for each atom and are labelled $n = 1, 2, 3$ etc. The electron may occupy only these orbits.
- An electron ordinarily occupies the lowest energy orbit available.
- An electron can jump to a higher energy level by absorbing some energy. The absorbed energy must be exactly equal to the difference between the electron's initial and final energy levels.
- Electromagnetic radiation is emitted by an excited atom when an electron falls from a higher energy level to a lower energy level. The energy of the emitted light will be exactly equal to the energy difference between the electron's initial and final levels.

These ideas are shown in Figure 15.1.3. In this particular example, the electron absorbs energy from light that strikes the atom. The energy absorbed is just the right amount for the electron to make the jump from its ground state to a higher level. In this diagram, the light is labelled as a photon.

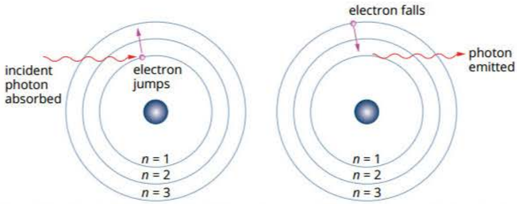


FIGURE 15.1.3 (a) If the incident photon (light) carries an amount of energy equal to the energy difference between two levels, the photon's energy can be absorbed, allowing the electron to jump to the higher level. The photon ceases to exist. (b) An atom will remain in an excited state for less than a millionth of a second. The electron will then fall to its ground state. The electron may fall in one step, or in a number of stages, emitting a photon or photons as it falls.

PHYSICSFILE CCT

Labelling energy levels

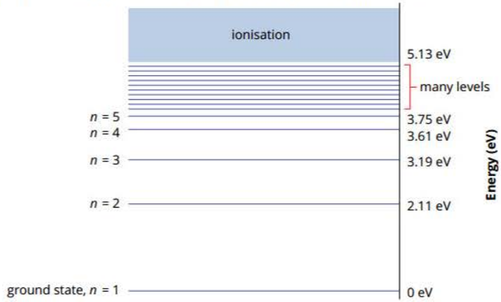
There are two systems in use for labelling the energy levels of an atom. Sometimes the ground state ($n = 1$) is allocated 0 eV and therefore the higher levels have positive values. Alternatively, sometimes the ground state is allocated a negative energy value and the ionisation energy level has a value of 0 eV.

Each possible electron transition (jump) produces light of different energy. The energy corresponds to a different coloured line in the emission spectrum for that atom. The greater the energy emitted, the higher the frequency of the light.

Worked example 15.1.1

ENERGY LEVELS

The energy levels for sodium gas are shown. Work out the energy of the light that is produced as an electron drops from the $n = 4$ to the $n = 2$ state.

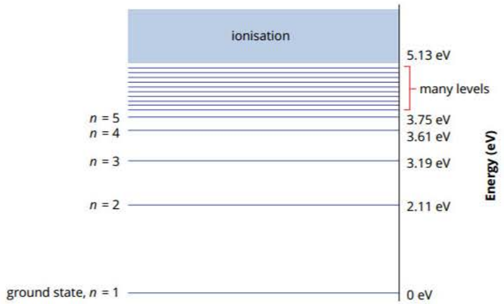


Thinking	Working
Using the figure, find the energy (in eV) of each level involved.	$n = 4$, $E_4 = 3.61$ eV $n = 2$, $E_2 = 2.11$ eV
Calculate the difference between these levels.	$\Delta E = E_4 - E_2$ $= 3.61 - 2.11$ $= 1.50$ eV

Worked example: Try yourself 15.1.1

ENERGY LEVELS

The energy levels for sodium gas are shown. Work out the energy of the light that is produced as an electron drops from the $n = 4$ to the $n = 3$ state.



SPECTRAL ANALYSIS

Recall from Chapter 9 that an emission spectrum is the result of electrons absorbing energy (the electrons become 'excited') and then releasing energy in the form of a photon. An emission spectrum can be analysed in terms of the energy of the photons produced. In his work on the photoelectric effect (see Chapter 11), Einstein used Planck's equation for the energy of a photon:

GO TO ➤ Section 9.2 page 250

GO TO ➤ Section 11.2 page 290

$$\Delta E = hf = \frac{hc}{\lambda}$$

where

ΔE is the energy of the photon produced (J)

h is Planck's constant (6.626×10^{-34} J s or 4.14×10^{-15} eV s)

f is the frequency of the photon (Hz)

c is the speed of light ($3.00 \times 10^{8}\ \text{m s}^{-1}$)

λ is the wavelength of the photon (m)

Notice that ΔE has been used in this equation instead of E. ΔE corresponds to the difference in energy between the excited state and the ground state of the electron that released the photon, and so it is used to represent the energy of the photon.

Worked example 15.1.2 relates to the emission spectra of **metal vapour lamps**. Metal vapour lamps produce light as their atoms are excited and then emit a photon as they return to their ground state. The emitted photons have wavelengths characteristic of the metals whose atoms are being excited in the lamp. A common type of metal vapour lamp is the sodium lamp. These are often used in street lighting and emit a distinctive yellow colour (Figure 15.1.4).

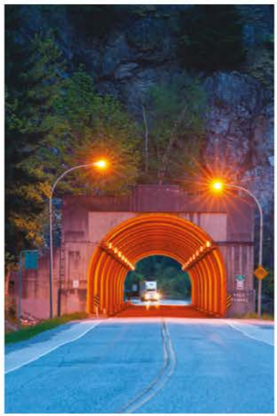


FIGURE 15.1.4 Sodium vapour lamps are commonly used as street lights and have a distinctive yellow colour due to the yellow wavelengths of the sodium emission spectrum.

Worked example 15.1.2

SPECTRAL ANALYSIS

The emission spectrum of a sodium vapour lamp is analysed and shows that most of the light is emitted with a frequency of around 5.1×10^{14} Hz. Calculate the energy of these photons in joules.

Thinking	Working
Recall Planck's equation.	$\Delta E = hf$
Substitute in the appropriate values and solve for ΔE.	$\Delta E = 6.626 \times 10^{-34} \times 5.1 \times 10^{14}$ $= 3.4 \times 10^{-19}$ J

Worked example: Try yourself 15.1.2

SPECTRAL ANALYSIS

In the Sun's absorption spectrum, one of the dark Fraunhofer lines corresponds to a frequency of 6.9×10^{14} Hz. Calculate the energy (in joules) of the photon that corresponds to this line. Use $h = 6.626 \times 10^{-34}$ J s.

The energy of the photon emitted or absorbed can also be expressed in electron-volts (eV). See the SkillBuilder in Chapter 11 to refresh yourself on converting between joules and electron-volts. To simplify calculations, Planck's constant can be restated in terms of electron-volts ($h = 4.14 \times 10^{-15}$ eV s), and this figure can be used to calculate the energy directly in electron-volts.

GO TO ➤ SkillBuilder page 288

Planck's constant $h = 4.14 \times 10^{-15}$ eV s $= 6.626 \times 10^{-34}$ J s

HYDROGEN'S ABSORPTION SPECTRUM

In the late 19th century, the emission and absorption spectra of hydrogen were of particular interest to scientists as it had been recognised that lines in the absorption spectrum of hydrogen matched lines in the solar spectrum (Figure 15.1.5).

FIGURE 15.1.5 In the absorption spectrum of hydrogen (a) there is a background of continuous white light (broken into a spectrum of colours), with black lines that correspond to wavelengths of the radiation absorbed by the hydrogen atoms. In the emission spectrum of hydrogen (b) there is a black background against which lines corresponding to the wavelengths emitted by the hydrogen atoms can be observed.

Although some scientists were able to come up with an empirical (based on experimental data) formula that predicted the wavelength of the lines in the hydrogen spectra, no one was able to provide a theoretical explanation for the production of these lines using a wave model for light.

BOHR MODEL OF THE ATOM

In 1913, the Danish physicist Niels Bohr proposed an explanation for the emission spectrum of hydrogen that drew on the quantum ideas proposed by Planck and Einstein, including Planck's quantum relation equation, $\Delta E = hf$. Bohr realised that:

- the absorption spectrum of hydrogen showed that the hydrogen atom was only capable of absorbing a small number of different frequencies of light and therefore energies of very specific values; that is, the absorbed energy was quantised
- the emission spectrum of hydrogen showed that hydrogen atoms could only emit quanta with the same exact energy value that the atoms were able to absorb
- hydrogen atoms have an ionisation energy of 13.6 eV; light of this energy or greater can remove an electron from a hydrogen atom, creating a positive ion
- photons of light with all energies above the ionisation value for hydrogen are continuously absorbed.

Bohr labelled the possible electron orbits for the hydrogen atom with a quantum number (n), and he was able to calculate the energy associated with each quantum number. Remember that the wavelengths (Figure 15.1.6) of all of the lines of the hydrogen emission spectrum can be calculated using Planck's equation:

$$\Delta E = \frac{hc}{\lambda}$$

Figure 15.1.7 shows the energy levels for the hydrogen atom. These energies are expressed in terms of how strongly the electron is bound to the nucleus. The ground level ($n = 1$) represents the orbit that is closest to the nucleus, i.e. the unexcited state. An electron in this orbit has an energy of (–)13.6 eV, which means that it would need to gain 13.6 eV of energy for it to escape the atom. Higher energy levels represent orbits that are further from the nucleus.

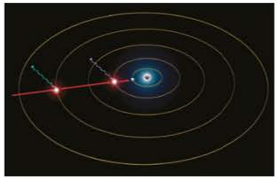


FIGURE 15.1.6 A diagram showing hydrogen spectrum emission levels based on the Bohr model of the atom. Electrons may only orbit in specific energy orbits, shown by the concentric circles. Electrons absorb energy to move to higher levels in their excited states and emit light in specific wavelengths characteristic of the element when returning to the ground state.

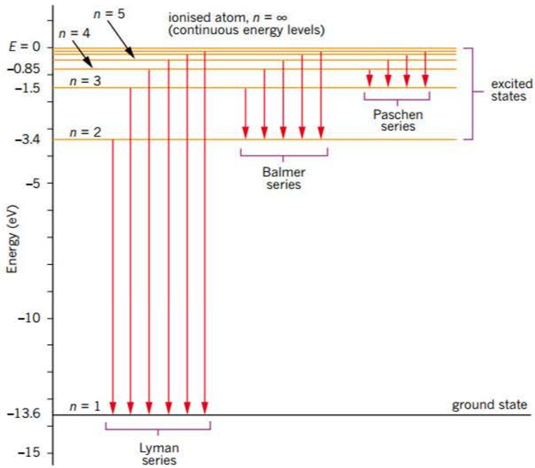


FIGURE 15.1.7 An energy level diagram for hydrogen. An electron in the ground state ($n = 1$) has an energy of (–)13.6 eV. For higher energy levels ($n > 1$), the energy levels can be seen to crowd together.

When a hydrogen atom gains energy, either by heating or from an electrical current, its electron moves from the ground state to one of the higher energy levels—the atom becomes excited. Eventually, the electron will drop from the higher energy level to one of the lower levels and will emit a photon with an energy equal to the difference in energy between the levels.

In Figure 15.1.7 the energy levels within the atom are negative in value. A free electron (at $n = \infty$) must have zero potential energy as it has escaped the electrostatic attraction of the proton in the nucleus. To raise an electron from one energy level to another, the appropriate amount of energy must be delivered. As an electron then falls back to its previous energy level, its energy value decreases. That is, it becomes a larger negative number.

Figure 15.1.7 also shows that the spectral lines of hydrogen can be explained in terms of electron transitions. The different series shown on the diagram (Lyman, Balmer, Paschen) represent specific transitions. The Balmer series, for example, shows transitions back to $n = 2$ from various excited energy levels. These transitions represent wavelengths of the visible lines of the hydrogen emission spectrum.

Worked example 15.1.3

USING THE BOHR MODEL OF THE HYDROGEN ATOM

Calculate the wavelength (in nm) of the photon produced when an electron drops from the $n = 4$ energy level of the hydrogen atom to the $n = 2$ energy level. Identify the spectral series to which this line belongs.

Use Figure 15.1.7 to calculate your answer.

Thinking	Working
Identify the energy of the relevant energy levels of the hydrogen atom.	$n = 4$, $E_4 = -0.85\,\text{eV}$ $n = 2$, $E_2 = -3.4\,\text{eV}$
Calculate the change in energy.	$\Delta E = E_4 - E_2$ $= -0.85 - (-3.4)$ $= 2.55\,\text{eV}$
Calculate the wavelength of the photon with this amount of energy.	$E = \frac{hc}{\lambda}$ $\therefore \lambda = \frac{hc}{E}$ $= \frac{4.14 \times 10^{-15} \times 3.0 \times 10^{8}}{2.55}$ $= 4.87 \times 10^{-7}\,\text{m}$ $= 487\,\text{nm}$ $= 490\,\text{nm}$ (to two significant figures)
Identify the spectral series.	The electron drops down to the $n = 2$ energy level. Therefore, the photon must be in the Balmer series.

Worked example: Try yourself 15.1.3

USING THE BOHR MODEL OF THE HYDROGEN ATOM

Calculate the wavelength (in nm) of the photon produced when an electron drops from the $n = 3$ energy level of the hydrogen atom to the $n = 1$ energy level. Identify the spectral series to which this line belongs.

Use Figure 15.1.7 to calculate your answer.

PHYSICS IN ACTION ICT

How lasers work

Lasers use the principle of stimulated emission of radiation. In fact, laser is an acronym for Light Amplification by Stimulated Emission of Radiation. Lasers produce a coherent and very intense light beam that is the result of a chain reaction, where photons cause identical photons to be produced.

There are a number of conditions necessary for laser light to be produced:

1. There must be more atoms in the excited state than in the ground state. This is not the normal condition of matter and it is achieved by 'pumping' the lasing material with energy from an external source (bombarding electrons (electrical pumping) or bombarding photons (optical pumping)).
2. The excited state used must be metastable (a relatively stable energy level where the electron remains excited for longer than normal and takes longer to drop to a lower level).
3. The photons emitted as the electrons jump back down to ground state must be used to continue the chain reaction of photon emissions. This is achieved by placing mirrors at the end of the laser tube: one mirror reflects all the photons back; the other mirror is partially silvered, making it reflect some, but not all, of the photons. The mirrors cause the photons to be reflected back and forth through the lasing medium, stimulating further emissions. This is the 'light amplification' part of the process. At the same time, some light escapes through the partially reflective mirror. This is the laser beam.

The entire process is illustrated in Figure 15.1.8.

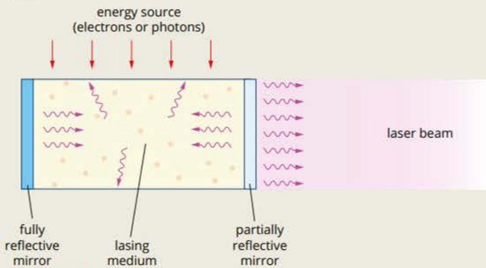


FIGURE 15.1.8 As the emitted photons reflect back and forth between the mirrors, further stimulated emissions occur and a laser beam is produced.

Balmer and Rydberg—empirical equations

In 1885, the Swiss mathematician Johann Balmer found an empirical equation that predicted the wavelength of the visible lines of the hydrogen emission spectrum:

$$\lambda = \frac{hm^2}{m^2 - n^2}$$

where λ is the wavelength of light (nm)

h is a constant with a value of 365 nm

m could take values of 3, 4, 5 or 6

$n = 2$.

When Balmer put $m = 7$ into the equation, it gave an answer of 397 nm, which corresponded to a spectral line that had been independently observed by Anders Ångström. Consequently, this set of spectral lines in the visible part of the electromagnetic spectrum came to be known as the Balmer series.

In 1888, Johannes Rydberg realised that Balmer's formula was a special case of the more general formula:

$$\frac{1}{\lambda} = R\left[\frac{1}{n_f^2} - \frac{1}{n_i^2}\right]$$

where

R is the Rydberg constant for hydrogen ($1.097 \times 10^7\ m^{-1}$)

n_f and n_i are any two integers where $n_i > n_f$

This equation predicted that there should be spectral lines in other parts of the electromagnetic spectrum. The ultraviolet series was later observed by Theodore Lyman and two different infrared series were observed by Friedrich Paschen and Frederick Brackett.

Worked example 15.1.4

USING THE RYDBERG FORMULA

Using the Rydberg formula, calculate the wavelength (in nm) of the photon produced when an electron drops from the $n = 2$ energy level of the hydrogen atom to the $n = 1$ energy level. Identify the spectral series to which this line belongs.

Thinking	Working
Identify the known variables.	$n_f = 1$ $n_i = 2$ $R = 1.097 \times 10^7\,m^{-1}$
Recall the Rydberg formula.	$\frac{1}{\lambda} = R\left[\frac{1}{n_f^2} - \frac{1}{n_i^2}\right]$
Solve for the wavelength, λ.	$\frac{1}{\lambda} = R\left[\frac{1}{n_f^2} - \frac{1}{n_i^2}\right]$ $= 1.097 \times 10^7 \times \left[\frac{1}{1^2} - \frac{1}{2^2}\right]$ $= 1.097 \times 10^7 \times 0.75$ $\lambda = 1.21 \times 10^{-7}$ $= 121\,nm$
Identify the spectral series.	The electron drops down to the $n = 1$ energy level. Therefore, the photon must be in the Lyman series.

Worked example: Try yourself 15.1.4

USING THE RYDBERG FORMULA

Using the Rydberg formula, calculate the wavelength (in nm) of the photon produced when an electron drops from the $n = 4$ energy level of the hydrogen atom to the $n = 1$ energy level. Identify the spectral series to which this line belongs.

ABSORPTION OF PHOTONS

The Bohr model also explains the absorption spectrum of hydrogen (Figure 15.1.5a on page 396).

You have already seen that the missing lines in absorption spectra correspond to the energies of light that a given atom is capable of absorbing. This is due to the energy differences between the atom's electron orbits. Only incident light carrying just the right amount of energy to raise an electron to an allowed level can be absorbed.

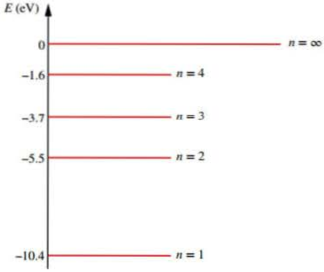
An electron ordinarily occupies the lowest energy orbit. Incident light that does not carry enough energy to raise an electron from this lowest energy level to the next level cannot be absorbed by the atom. Incident light below a certain energy value would simply pass straight through. If light with greater energy than the ionisation energy of an atom is incident, then the excess energy provided by the photon will simply translate to extra kinetic energy for the released electron.

For hydrogen, then, if a hydrogen atom absorbs a photon with 13.6 eV or more, as this is the energy required for the electron to escape the atom completely, the hydrogen atom is said to be **ionised**.

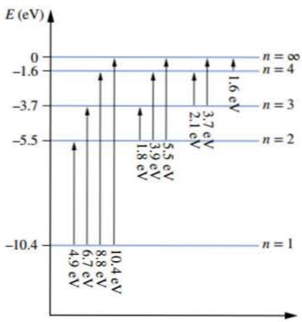
Worked example 15.1.5

ABSORPTION OF PHOTONS

Some of the energy levels for atomic mercury are shown in the diagram below.



Ultraviolet light with photon energies 4.9 eV, 5.0 eV and 10.50 eV is incident on some mercury gas. What could happen as a result of the incident light?

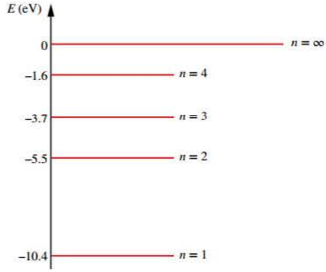
Thinking	Working
Check whether the energy of each photon corresponds to any differences between energy levels by determining the difference in energy between each level.	
Compare the energy of the photons with the energies determined in the previous step. Comment on the possible outcomes.	A photon of 4.9 eV corresponds to the energy required to promote an electron from the ground state to the first excited state ($n = 1$ to $n = 2$). The photon may be absorbed. A photon of 5.0 eV cannot be absorbed since there is no energy level above the ground state that corresponds exactly to 5.0 eV. A photon of 10.5 eV may ionise the mercury atom. The ejected electron will leave the atom with 0.1 eV of kinetic energy.

WS 8.5

Worked example: Try yourself 15.1.5

ABSORPTION OF PHOTONS

Some of the energy levels for atomic mercury are shown in the diagram below.



Light with photon energies 6.7 eV, 9.0 eV and 11.0 eV is incident on some mercury gas. What could happen as a result of the incident light?

LIMITATIONS WITH BOHR'S MODEL

Bohr's model of the hydrogen atom applied a quantum approach to the energy levels of atoms to explain a set of important, previously unexplained phenomena—the emission and absorption spectra of hydrogen. In principle, Bohr's work on the hydrogen atom could be extended to other atoms and, in 1914, the German scientists James Franck and Gustav Hertz demonstrated that mercury atoms contained energy levels similar to hydrogen atoms. Bohr's model signified an important conceptual breakthrough.

However, Bohr's model was limited in its application, for the following reasons:

- It could only really be accurately applied to single-electron atoms—hydrogen and ionised helium—as it modelled inner-shell electrons well but could not predict the higher-energy orbits of multi-electron atoms.
- It could not explain the varying intensity and thickness of the spectrum lines of hydrogen (thought, by some scientists, to be caused by the elliptical instead of circular orbit of electrons).
- It could not explain why spectral lines were split in the presence of a magnetic field (known as the Zeeman effect). Some of the observed emission lines could be resolved into two very close spectral lines, and Bohr's model could not explain this.
- It could not explain why some spectral lines split into a series of spectral lines with a spacing proportional to the magnetic field strength (known as the anomalous Zeeman effect).
- It could not explain the discovery of the continuous spectrum emitted by solids. As a result, a more complex quantum approach was required.

PHYSICSFILE CCT

The special case of hydrogen

The hydrogen atom was a relatively simple place to begin the development of the field that would come to be known as 'quantum mechanics'. The hydrogen atom contains two charged particles—the positively charged nucleus (which usually contains a single proton) and the electron (Figure 15.1.9). This means that only one electrical interaction (i.e. between the electron and the nucleus) needs to be considered.

In more complex atoms, such as helium, electrical interactions between the electrons are also significant. This makes the construction of mathematical models for these atoms vastly more complicated than for hydrogen.

FIGURE 15.1.9 The hydrogen atom contains only two particles: the proton in the nucleus and the electron.

15.1 Review

SUMMARY

- An emission spectrum is produced by energised atoms as electrons undergo transitions between energy levels. The spectrum for each element is unique.
- Atoms that followed Rutherford's model of the atom would be unstable, so a new model—a quantum model—was needed.
- Niels Bohr suggested that electrons in atoms orbit the nucleus in specially defined energy levels, and no radiation is emitted or absorbed unless the electron can jump from its energy level to another.
- An electron in an atom which drops between energy levels emits a photon (light) of energy equal to the difference between the energy levels. The energy of the photon determines the colour of the light.
- Bohr suggested that electrons in atoms orbit the nucleus in specially defined energy levels. No radiation is emitted or absorbed unless the electron can jump from one energy level to another. Electron energies are said to be quantised, since only certain values are allowed.
- The wavelength of the spectral lines for the hydrogen atom are given by the Rydberg formula $\frac{1}{\lambda} = R\left[\frac{1}{n_f^2} - \frac{1}{n_i^2}\right]$.
- The frequency of a photon emitted or absorbed by a hydrogen atom can be calculated from the difference between the energy levels involved, i.e. $E_2 - E_1 = hf = \frac{hc}{\lambda}$.
- The Bohr model of the atom is limited in its application, but was a significant development at the time as it took a quantum approach to the energy levels of atoms and incorporated the quantum nature of electromagnetic radiation.

KEY QUESTIONS

To answer these questions, use $h = 6.63 \times 10^{-34} = 4.14 \times 10^{-15}$ eVs and $R = 1.097 \times 10^{7}\,\text{m}^{-1}$.

1 Describe a typical emission spectrum for an element.

2 What is the link between an atom's emission spectrum and its structure?

3 How are the energy levels within an atom commonly represented on a graph?

4 An emission line of frequency 6.0×10^{14} Hz is observed when looking at the emission spectrum of a particular elemental gas. What is the energy, in joules, of photons corresponding to this frequency?

5 Photons of energy 0.42 eV are emitted by a particular atom as it returns from the excited to the ground state. What is the corresponding wavelength of these photons?

6 Calculate the energy of the photon required to move an electron in a hydrogen atom from its ground state ($n = 1$) to the $n = 4$ energy level. Refer to Figure 15.1.7 on page 397.

7 Using the Rydberg formula, calculate the wavelength of the emitted photon when an electron transitions from the $n = 5$ energy level to the ground state ($n = 1$).

15.2 Quantum model of the atom

GO TO ➤ Section 11.2 page 290

The photoelectric effect and the quantum model for light were introduced in Chapter 11 and the first part of this chapter. In order to explain the photoelectric effect, Einstein used the photon concept that Planck had developed. However, like many great discoveries in science, the development of the quantum model of light raised almost as many questions as it answered. It has already been well established that a wave model was needed to explain phenomena such as diffraction and interference. How could these two contradictory models be reconciled to form a comprehensive theory of light?

Answering this question was one of the great scientific achievements of the 20th century and led to the extension of the quantum model to matter as well as energy. It led to a fundamental shift in the way the universe is viewed.

WAVE–PARTICLE DUALITY

FIGURE 15.2.1 An artist's attempt to represent wave–particle duality.

In many ways, the wave and particle models for light seem fundamentally incompatible. Waves are continuous and are described in terms of wavelength and frequency. Particles are discrete and are described by physical dimensions such as their mass and radius.

In order to understand how these two sets of ideas can be used together, it is important to remember that scientists describe the universe using models. Models are analogies that are used to illustrate certain aspects of reality that might not be immediately apparent.

Physicists have come to accept that light is not easily compared to any other physical phenomenon. In some situations, light has similar properties to a wave; in other situations, light behaves more like a particle. This understanding is called wave–particle duality (Figure 15.2.1). Although this may seem somewhat paradoxical and counter-intuitive, in the century since Einstein did his work establishing quantum theory, many experiments have supported this duality and no scientist has (yet) come up with a better explanation.

De Broglie's wave–particle theory

In 1924, the French physicist Louis de Broglie proposed a groundbreaking theory. He suggested that since light (which had long been considered to be a wave) sometimes demonstrated particle-like properties, then perhaps matter (which was considered to be made up of particles) might sometimes demonstrate wave-like properties.

He quantified this theory by predicting that the wavelength of a particle would be given by the equation:

$$\lambda = \frac{h}{p}$$

where

λ is the wavelength of the particle (m)

p is the momentum of the particle ($kg\,m\,s^{-1}$)

h is Planck's constant (6.63×10^{-34} J s or 4.14×10^{-15} eV s)

This is also commonly written as:

$$\lambda = \frac{h}{mv}$$

where

m is the mass of the particle (kg)

v is the velocity of the particle ($m\,s^{-1}$)

The wavelength that de Broglie described, λ, is referred to as the **de Broglie wavelength** of matter.

Worked example 15.2.1

CALCULATING THE DE BROGLIE WAVELENGTH

The Davisson–Germer experiment was a famous experiment that showed that electrons have wave-like properties. The electrons in this experiment travelled at about $4.0 \times 10^6\,\text{m}\,\text{s}^{-1}$.

Calculate the de Broglie wavelength of these electrons if the mass of an electron is $9.109 \times 10^{-31}\,\text{kg}$.

Thinking	Working
Recall de Broglie's equation.	$\lambda = \frac{h}{mv}$
Substitute the appropriate values into the equation and solve it.	$\lambda = \frac{h}{mv}$ $= \frac{6.63\times10^{-34}}{9.109\times10^{-31}\times4\times10^{6}}$ $= 1.8 \times 10^{-10}\,\text{m}$ or $0.18\,\text{nm}$

Worked example: Try yourself 15.2.1

CALCULATING THE DE BROGLIE WAVELENGTH

Calculate the de Broglie wavelength of a proton travelling at $7.0 \times 10^5\,\text{m}\,\text{s}^{-1}$. The mass of a proton is $1.67 \times 10^{-27}\,\text{kg}$.

Worked example 15.2.2

CALCULATING THE DE BROGLIE WAVELENGTH OF A MACROSCOPIC OBJECT

Calculate the wavelength of a cricket ball of mass $m = 160\,\text{g}$ travelling at $150\,\text{km}\,\text{h}^{-1}$.

Thinking	Working
Convert mass and velocity to SI units.	$m = 160\,\text{g} = 0.16\,\text{kg}$ $v = \frac{150}{3.6} = 42\,\text{m}\,\text{s}^{-1}$
Recall de Broglie's equation.	$\lambda = \frac{h}{mv}$
Substitute the appropriate values into the equation and solve it.	$\lambda = \frac{h}{mv}$ $= \frac{6.63\times10^{-34}}{0.16\times42}$ $= 9.9 \times 10^{-35}\,\text{m}$

Worked example: Try yourself 15.2.2

CALCULATING THE DE BROGLIE WAVELENGTH OF A MACROSCOPIC OBJECT

Calculate the de Broglie wavelength of a person with $m = 66\,\text{kg}$ running at $36\,\text{km}\,\text{h}^{-1}$.

It can be seen from worked examples 15.2.1 and 15.2.2 that the wavelength of an electron is smaller than that of visible light, but is still large enough to be measurable. However, the wavelength of an everyday object such as a cricket ball is extremely small ($9.9 \times 10^{-35}\,\text{m}$). Hence, you will never notice the wave properties of everyday objects. To illustrate this, consider the observable wave behaviour of diffraction. Recall that for diffraction to be noticeable, the size of the wavelength needs to be comparable to the size of the gap or obstacle. Therefore for an everyday object, with its tiny wavelength, to produce a noticeable diffraction, it would need to pass through a gap much smaller than a fraction of a proton diameter!

ELECTRON DIFFRACTION PATTERNS

De Broglie's prediction that matter could exhibit wave-like behaviour was controversial. However, it was experimentally confirmed by the Americans Davisson and Germer in 1927 when they observed diffraction patterns being produced when they bombarded the surface of a piece of nickel with electrons (Figure 15.2.2).

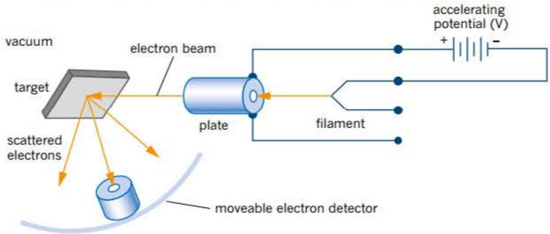


FIGURE 15.2.2 The Davisson and Germer apparatus to show electron scattering.

They used an electron 'gun' which provided a beam of electrons. The speed of the electrons was known because they had been accelerated through a known voltage. The detector could be swung around on an axis so that it could intercept electrons scattered from the nickel target in any direction in the plane shown.

Davisson and Germer found that as they moved their detector through the different scattering angles, they encountered a sequence of maximum and minimum intensities (Figure 15.2.3).

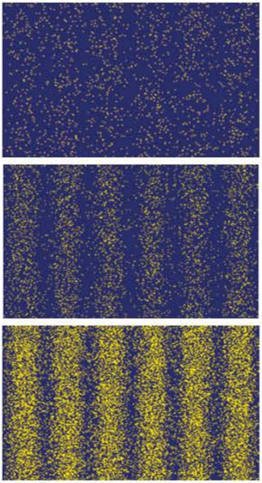


FIGURE 15.2.3 An electron diffraction pattern like the one observed by Davisson and Germer can be built up over time from repeated observations.

Clearly, the electrons were being scattered by the different layers within the crystal lattice of the target (Figure 15.2.4) and were undergoing interference. When Davisson and Germer analysed the diffraction pattern to determine the wavelength of the 'electron waves', they calculated a value of 0.14 nm, which was consistent with de Broglie's hypothesis.

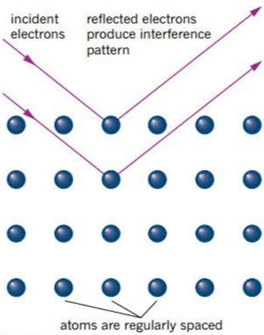


FIGURE 15.2.4 Electrons reflecting from different layers within the crystal structure create an interference pattern like those produced by a diffraction grating.

Worked example 15.2.3

WAVELENGTH OF ELECTRONS FROM AN ELECTRON GUN

Find the de Broglie wavelength of an electron that has been accelerated from rest through a potential difference of 75 V. The mass of an electron is 9.11×10^{-31} kg and the magnitude of the charge on an electron is 1.6×10^{-19} C.

Thinking	Working
Calculate the kinetic energy of the electron from the work done on it by the electric potential. Recall from earlier chapters that $W = qV$.	$W = qV$ $= 1.6 \times 10^{-19} \times 75$ $= 1.2 \times 10^{-17}$ J
Calculate the velocity of the electron.	$K = \frac{1}{2}mv^2$ $v = \sqrt{\frac{2K}{m}}$ $= \sqrt{\frac{2 \times 1.2 \times 10^{-17}}{9.11 \times 10^{-31}}}$ $= 5.1 \times 10^6\ \text{m s}^{-1}$
Use de Broglie's equation to calculate the wavelength of the electron.	$\lambda = \frac{h}{mv}$ $= \frac{6.63 \times 10^{-34}}{9.11 \times 10^{-31} \times 5.1 \times 10^6}$ $= 1.4 \times 10^{-10}$ m $= 0.14$ nm

Worked example: Try yourself 15.2.3

WAVELENGTH OF ELECTRONS FROM AN ELECTRON GUN

Find the de Broglie wavelength of an electron that has been accelerated from rest through a potential difference of 50 V. The mass of an electron is 9.11×10^{-31} kg and the magnitude of the charge on an electron is 1.6×10^{-19} C.

Comparing the wavelengths of photons and electrons

In the same year that Davisson and Germer conducted their experiment, other supporting evidence came from G. P. Thomson (son of J. J. Thomson, discoverer of the electron). Rather than scatter an electron beam from a crystal, Thomson produced a diffraction pattern by passing a beam of electrons through a tiny crystal. Thomson then repeated his experiment, using X-rays of the same wavelength in place of the electrons. The X-ray diffraction pattern was almost identical to the one made with electrons, as shown in Figure 15.2.5.

As the diffraction patterns obtained for the X-ray photons and electrons were the same, and as both were passed through the same 'gaps' to obtain this diffraction pattern, then an important conclusion could be made. The electrons must have a similar wavelength to the X-rays. Since their wavelengths are similar, the momenta of the electrons and the X-ray photons must also be comparable (but not their speeds).

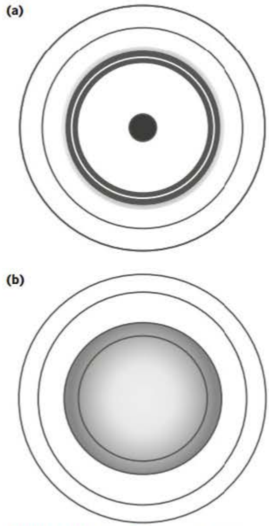


FIGURE 15.2.5 These diffraction patterns were taken by using (a) X-rays and (b) a beam of electrons with the same target crystal. Their similarity suggests a wave-like behaviour for the electrons and a similar electron de Broglie wavelength to that of X-rays.

PHYSICS IN ACTION ICT

Electron microscopes

The discovery of the wave properties of electrons had an important practical application in the invention of the electron microscope. Just as an optical microscope makes use of the wave properties of photons to magnify tiny objects, so too can the wave properties of electrons be used to create magnified images (Figure 15.2.6).

One of the limitations of an optical microscope is that it can only create a clear image of structures that are similar in size to the wavelength of the light being used. This is because the light diffracts around these structures. So a light microscope is only useful for seeing things down to about 390 nm, the lower wavelength end of the visible-light spectrum.

However, the wavelength of a beam of electrons is often smaller than the wavelength of a beam of visible light. This means that electron microscopes can create images with much finer detail than optical microscopes.

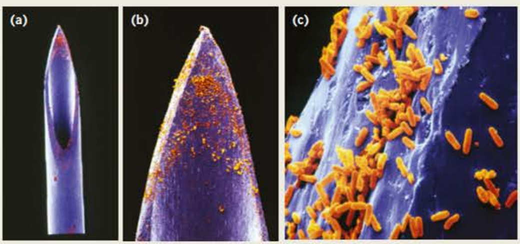


FIGURE 15.2.6 Images formed by an electron microscope: rod-shaped bacteria (orange) clustered on the point of a syringe used to administer injections. The magnifications are (a) ×9, (b) ×36 and (c) ×560 at 35 mm size.

STANDING WAVES AND THE DUAL NATURE OF MATTER

Wave behaviour can be used to indicate the probability of the path of a particle. If particles can be thought of as matter waves, then these matter waves must be able to maintain steady energy values if the particles are to be considered stable.

De Broglie applied his approach to the discussion of Bohr's model for the hydrogen atom. He viewed the electrons orbiting the hydrogen nucleus as matter waves. He suggested that the electron could only maintain a steady energy level if it established a **standing wave**.

De Broglie reasoned that if an electron of mass m were moving with speed v in an orbit with radius r, this orbit would be stable if it matched the condition

$$mvr = n\frac{h}{2\pi}$$

where n is an integer.

This can be rearranged to

$$2\pi r = n\frac{h}{mv}$$

Since $2\pi r$ is the circumference, C, of a circle, and using the de Broglie equation, $\lambda = \frac{h}{mv}$, this equation can be rewritten as $C = n\lambda$.

In other words:

> **The stable orbits of the hydrogen atom are those where the circumference is exactly equal to a whole number of electron wavelengths.**

This can be visualised by imagining a conventional standing-wave pattern, like that of a vibrating string, being looped around on itself in three dimensions as shown in Figure 15.2.7.

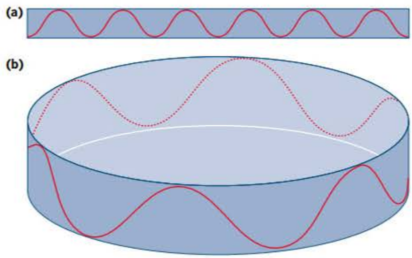


FIGURE 15.2.7 A standing-wave pattern (a) can be looped around on itself to form (b) if the circumference of the circle is equal to a whole number of wavelengths.

If the circumference of the circle is not equal to a whole number of wavelengths, then destructive interference occurs, a standing-wave pattern cannot be established and the orbit cannot represent an energy level (Figure 15.2.8).

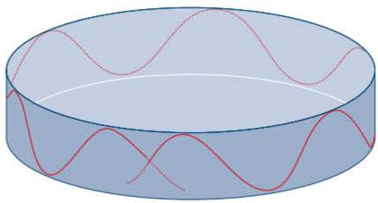


FIGURE 15.2.8 A circular standing-wave pattern cannot be established if the circumference of the circle is not equal to a whole number of wavelengths.

HEISENBERG'S UNCERTAINTY PRINCIPLE

In the early 20th century, scientists struggled to interpret the evidence of the dual wave–particle nature of energy and matter. Waves and particles are fundamentally different—waves are extended and continuous, whereas particles are discrete. How two such different models could be combined to describe the fundamental building blocks of nature was a serious puzzle. As scientists delved into this mystery, they discovered fundamental limitations to their ability to explore the 'quantum' universe.

A QUANTUM INTERPRETATION OF THE ELECTRON

WS 8.6

In 1925, the Austrian physicist Erwin Schrödinger built on the work of Niels Bohr by developing a mathematical equation that could describe the wave behaviour of electrons in situations other than the simple hydrogen atom.

In Schrödinger's model, the wave properties of electrons are interpreted as representing the probability of finding an electron in a certain location.

Quantum mechanics is the name now given to the area of physics in which the wave properties of electrons are studied. Schrödinger's equation (Figure 15.2.9) has been used to calculate the regions of space in which an electron can be found in a hydrogen atom. These are now known as **orbitals** rather than orbits because they are complex three-dimensional shapes, as shown in Figure 15.2.10, rather than the simple circular paths once imagined by Rutherford and Bohr.

$$i\hbar\frac{\partial}{\partial t}\Psi = \hat{H}\Psi$$

FIGURE 15.2.9 Schrödinger's wave equation. While outside the scope of this course, it is interesting to see what Schrödinger's equation looks like. The Greek symbol Ψ (psi) represents the wave function of the electron.

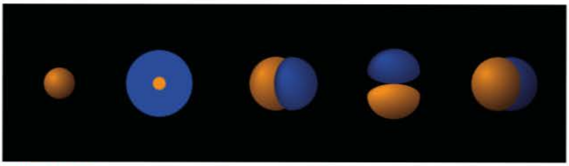


FIGURE 15.2.10 The shapes of the first five electron orbitals of a hydrogen atom.

PHYSICSFILE CCT

Schrödinger's cat

Schrödinger described the strange, counter-intuitive nature of quantum mechanical systems using a now-famous analogy known as Schrödinger's cat.

This is a thought experiment (i.e. Schrödinger did not actually perform the experiment) in which a cat is placed in a closed box with a flask of poison. A quantum mechanical system is set up such that there is a 50% chance of the flask being broken and the cat killed.

Schrödinger argued that until the box is opened to reveal the outcome of the experiment, the cat is considered as simultaneously alive *and* dead.

In a manner similar to the dual nature of light, the outcome (for the cat being alive or dead; for light being a wave or a particle) does not exist until an observation or measurement is made.

LIMITS TO MODELS AT VERY SMALL SCALES

It was becoming clear to scientists that the nature of the universe at the very smallest scale is fundamentally different to the way the universe is perceived at the macroscopic scale.

In everyday life, each object has a clearly definable position and motion. The classical laws of physics, developed by scientists from Newton through to Maxwell, are all based on this assumption, which is so fundamental to human experience that it is hard to imagine a universe where this is not the case.

However, this is exactly what is needed in order to explore the quantum universe. There is no particular reason why tiny particles such as electrons and photons should be similar to larger objects like balls or planets; scientists initially just extrapolated from their experience until the evidence showed that their assumptions were wrong.

Whenever a measurement is taken, a degree of error or *uncertainty* is involved. The certainty of the measurement is limited by measuring the resolution of the device used to make the measurement. For example, a ruler with markings one millimetre apart will have an uncertainty in any measurement of about half a millimetre. This is taken into account when commenting on the errors in a practical experiment or calculating the final uncertainty in a result. If more precision is needed in a final result, then a more precisely marked measuring device such as a micrometer or vernier caliper is needed to make the initial measurements. The smaller the divisions, the smaller the final uncertainty in the result.

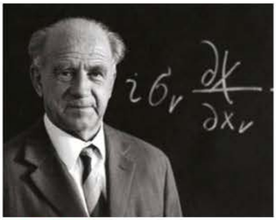


FIGURE 15.2.11 Werner Heisenberg won the Nobel Prize in 1932 for his work on the uncertainty principle. Heisenberg is regarded as a founder of quantum physics.

However, according to quantum mechanics there is a physical limit to the absolute accuracy of particular measurements. This limit is inherent in nature and is a result of both wave–particle duality and the interactions between the object being observed and the effect of the observation on that object (as Schrödinger tried to explain). The first scientist to clearly identify this limit was the German physicist Werner Heisenberg (Figure 15.2.11). The **Heisenberg uncertainty principle** describes a limit to which some quantities can be measured.

Imagine trying to find a ball in a pitch-black room. The only way to do so, assuming there is no light, is to feel around. The more you search and feel around, the more confident you can be that the object remains in the area of the room yet to be searched—until you actually touch it. Then there is every chance that it will roll away, the ball having been given momentum by your touch. You'll no longer know its future position. The very act of determining the position of the ball made knowing its future position less certain.

According to Heisenberg's uncertainty principle, the more exactly the position of a subatomic particle is known, the less is known about its momentum. Similarly, the more precisely the momentum of a particle is measured, the less certain is its exact position.

Similarly, at the quantum level, to measure the exact location of an electron it would be necessary to hit the electron with another particle such as a photon of light. However, as soon as the photon strikes the electron, it would cause the electron to move as the photon transfers some or all of its energy to the electron. The act of measurement causes a change in the value being measured. This is a general problem when trying to measure the location or motion of all subatomic particles.

Heisenberg's uncertainty principle applies particularly at the quantum level precisely because electrons aren't particles. The whole dual nature of matter idea means that electrons, and in fact all matter, behave both as a wave and as a particle. The uncertainty principle applies a limit to the simplified idea of electrons as particles; i.e. that the position and the velocity of an electron cannot both be known at exactly the same time, and that the amount of energy at a particular time, t, is also uncertain. For the normal-sized world around us, the inclusion of Planck's constant, h, in the measure of uncertainty means that the level of uncertainty in determining the position of everyday objects is extremely small—in fact, virtually insignificant.

However, at an atomic scale, this level of uncertainty is substantial. And since everyday objects are made up of atoms containing subatomic particles such as electrons, the basic understanding of matter comes down to this fundamental property of all quantum mechanical systems.

+ ADDITIONAL

The uncertainty principle

Heisenberg went on to explain his uncertainty principle with a formula that states that the product of the uncertainties in the position (Δx) and momentum (Δp) of a particle must always be greater than a certain minimum value related to Planck's constant. That is:

$$\Delta x \Delta p \geq \frac{h}{4\pi}$$

or alternatively in terms of energy (since the photon may transfer some or all of its energy to the electron):

$$\Delta E \Delta t \geq \frac{h}{4\pi}$$

where ΔE is the uncertainty in the energy of an object and Δt is the uncertainty in the time taken for the energy transfer to occur.

You may see this equation stated differently in some references. The difference is due to the way in which the variables are defined. The important part of the relationship is the fact that Δx and Δp are inversely proportional. When either Δx or Δp increases, the other must decrease.

This equation is sometimes called the 'indeterminacy principle'. It states that it is not possible to know both the position and the momentum of an object at exactly the same time (Table 15.2.1). The more accurately the position is measured, the greater the uncertainly in the momentum and vice versa. Note that this doesn't infer that an absolute measure of position cannot be made. Just that, in doing so, its momentum at that same time wouldn't be known and hence there is no way of knowing what the position would be a second later.

TABLE 15.2.1 It is not possible to accurately know both the position and the momentum of an object at exactly the same time.

	Position	Momentum
scenario 1	known	unknown
scenario 2	unknown	known

PHYSICS IN ACTION CCT

Viewing an electron

Imagine trying to view an electron with an optical microscope like that shown in Figure 15.2.12. For the electron to be seen, a photon would have to strike the electron and be reflected back to the observer. As noted earlier in this chapter when looking at X-ray diffraction, objects can be seen at their best when the wavelength of the electromagnetic radiation used is at least as small as the object. Since a short wavelength corresponds to a high frequency and high energy (i.e. $E = hf = \frac{hc}{\lambda}$), the photons needed to observe the electron would have high energy and thus would impart more momentum to the object being observed. The higher the energy, the shorter the wavelength and the better the potential resolution, but the more likely it would be that the photon would knock the electron off course and hence the object's position would be subject to greater uncertainty. Just attempting to observe the electron introduces significant uncertainty in either the position or the momentum of the electron.

How then can an electron be viewed?

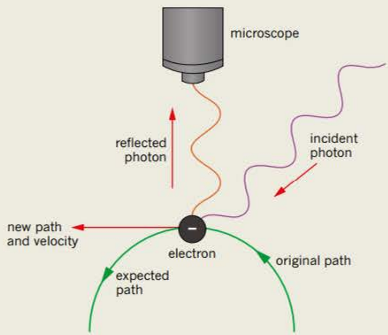


FIGURE 15.2.12 A thought experiment considering how an electron could be observed. The reflection of the photon needed to observe the electron introduces uncertainty in the position of the electron, making it unobservable.

A QUANTUM VIEW OF THE WORLD

Bohr's model of the atom was a mixture of classical and quantum theories, partially recognising the wave–particle duality of light and matter. However, it allowed for the development of new, more-radical theories to be developed by physicists such as Schrödinger and Heisenberg. And, intriguingly, quantum mechanics has confirmed certain aspects of the Bohr model, such as atoms existing only in discrete states of definite energy and the emission or absorption of photons of light when electrons make transitions from one energy state to another within an atom.

However, quantum mechanics goes much further—very much further than the brief introduction in this section. According to quantum mechanics, electrons don't exist in well-defined circular paths as we so regularly depict them in books. Because electrons are not particles, they don't follow particular paths in space and time at all. Rather, because of the wave nature of an electron, the paths can be thought of more as clouds, where the particular location of an electron at any point in time is based on probability since to measure a precise point, according to Heisenberg, only introduces uncertainty, preventing you from knowing where the electron would be at the next moment in time (Figure 15.2.13).

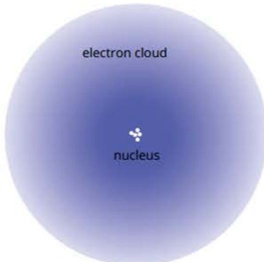


FIGURE 15.2.13 An atom as quantum mechanics views it. This conceptual image depicts the position of an electron around an atomic nucleus as an electron cloud or probability distribution of the location of the electron at any point in time.

The classical view of the world is a Newtonian one, where once the position and speed of an object are known, its future position can be predicted. This is termed a 'deterministic' model and works particularly well in predicting the positions of ordinary objects. This model suggests that the future of the universe, made up of particle-like objects, is completely knowable. Quantum mechanics proposes something very different, where the dual nature of particles—particularly fundamental particles such as the electron—prevents us from knowing the position and speed of an object at the same time. We can only calculate the probability that an electron will be observed at a particular place around an atom. In this view of the world there is some inherent unpredictability. In fact, it becomes meaningless to ask how an electron gets from one state to another when an atom emits or absorbs a photon of light—it just does.

15.2 Review

SUMMARY

- On the atomic level, energy and matter exhibit the characteristics of both waves and particles.
- All matter, like light, has a dual nature. Through everyday experience matter is particle-like, but under some situations it has a wave-like nature. This symmetry in nature—the dual nature of light and matter—is referred to as wave–particle duality.
- de Broglie viewed electrons as matter waves. His standing-wave model for electron orbits provided a physical explanation for electrons only being able to occupy particular energy levels in atoms. He suggested that the only way that the electron could maintain a steady energy level was if it established a standing wave.
- Quantum mechanics is the study of the wave properties of electrons. These wave properties are interpreted as describing the probability of finding an electron at a particular point in space.
- The wavelength of a particle is given by the de Broglie equation: $\lambda = \frac{h}{p}$ i.e. $\lambda = \frac{h}{mv}$.
- Heisenberg's uncertainty principle results from wave–particle duality and states that it is not possible to know the exact position and momentum of a particle simultaneously. The more precisely the position of a subatomic particle is known, the less is known about its momentum. Similarly, the more precisely the momentum of a particle is measured, the less certain is its exact position.
- The electron-diffraction experiment provides an example of the application of the uncertainty principle. The diffraction pattern is not produced by photon interactions but rather the probability of where a single photon may end up. Some positions are more likely than others, hence there is a larger central maximum intensity.
- In the quantum mechanical view of the atom, electrons are not particles and do not have clearly defined orbits. Their paths, due to their wave nature, can be explained as a probability distribution of their positions as particles or an electron cloud.

KEY QUESTIONS

To answer these questions, use $h = 6.63 \times 10^{-34} = 4.14 \times 10^{-15}$ eVs and $R = 1.097 \times 10^{7}\,m^{-1}$.

1 What is the de Broglie wavelength of an electron travelling at $1.0 \times 10^{6}\,m\,s^{-1}$?

2 Which of the following conclusions can be drawn from Louis de Broglie's investigation into the existence of matter waves?

A all particles exhibit wave behaviour
B only moving particles exhibit wave behaviour
C only charged particles exhibit wave behaviour
D only moving charged particles exhibit wave behaviour

3 In an experiment to determine the structure of a crystal, identical diffraction patterns were formed by a beam of electrons and a beam of X-rays with a frequency of 8.6×10^{18} Hz.

a Calculate the wavelength of the electrons.
b Calculate the speed of the electrons.

4 In a particular experiment, if the uncertainty about the position of a particle were decreased, then what will happen to the uncertainty about the speed of the particle?

5 Explain why Newtonian (classical) laws of physics are not appropriate when describing what occurs at the subatomic level.

6 Which of the following statements is *not* consistent with Heisenberg's uncertainty principle?

A The position and momentum of a particle can never be exactly known at the same time.
B The more exactly the position of particle is measured, the more uncertainty there is about its momentum.
C If the momentum of a particle is precisely measured, its position will be unknown.
D It is possible to precisely measure the momentum and position of a particle simultaneously.

Chapter review

15

KEY TERMS

de Broglie wavelength
energy levels
Heisenberg's uncertainty principle
ionise
laser
metal vapour lamp
orbit
orbital
quantum mechanics
standing wave

REVIEW QUESTIONS

1 Electrons in sodium gas are in the $n = 2$ excited state. As the electrons drop back to the ground state, what is the energy (in eV) of the emitted light? Use Figure 15.1.2 on page 393.

2 Electrons in sodium gas are in the $n = 4$ excited state. As the electrons drop back to ground state, what is the highest energy photon of light (in eV) that is produced? Use Figure 15.1.2 on page 393.

3 Can an atom emit photons (light) of any energy when an electron falls from a higher energy level to a lower energy level? Explain your answer.

4 Students are using spectroscopes to analyse the light emitted by a light globe and a sodium vapour lamp. Compare their observations.

5 By referring to the behaviour of electrons, explain how light is produced in an incandescent (filament) light globe.

6 Calculate the speed of an electron that has a de Broglie wavelength of 4.0×10^{-9} m.

7 Heisenberg's uncertainty principle was a contributing factor in showing Bohr's model of the atom to be inaccurate. How does the uncertainty principle contribute to the inaccuracy of Bohr's model?

8 How is the orange light that is produced by a sodium vapour street light created?

9 What is the de Broglie wavelength of a 40 g bullet travelling at $1.0 \times 10^3\,\text{m s}^{-1}$?

10 Would wave behaviours such as diffraction be noticeable for the bullet described in Question 9?

11 A particular atom has four energy levels. In this context, what does it mean to say that the levels are quantised?

12 Calculate the frequency of the photon produced when an electron in a hydrogen atom drops from the $n = 3$ energy level to its ground state, $n = 1$. Use Figure 15.1.7 on page 397.

13 Explain why the development of the Bohr model of the hydrogen atom was significant in the development of a comprehensive understanding of the nature of light.

14 Describe the relationship between the colours seen in the emission and absorption spectra of hydrogen.

15 According to Heisenberg's uncertainty principle, if the uncertainty in position is decreased, what will happen to the uncertainty in momentum?

16 If a photon of a very short wavelength were to collide with an electron, what would be the effect on the uncertainty of the position of the electron?

17 Bohr's quantised model of the atom was a significant development. However, it was limited in application. What was the Bohr model unable to explain?

18 When an electron drops from the $n = 5$ energy level of the hydrogen atom to the $n = 2$ energy level, a 434 nm photon is released. If the $n = 2$ orbit has an energy of −3.4 eV, what is the energy of the $n = 5$ orbit?

19 Explain why a cricket player does not have to consider the wave properties of a cricket ball while batting.

20 At what speed would a proton be travelling if it were to have the same wavelength as a gamma ray of energy 6.63×10^{-14} J? (Use the mass of a proton $m_p = 1.67 \times 10^{-27}$ kg.)

21 A charge q of mass m is accelerated from rest through a potential difference of V. Derive an expression that defines the de Broglie wavelength of the mass, λ, in terms of q, m and V.

22 A corollary of de Broglie's work on matter waves is that photons can be considered to have momentum. The momentum of photons, although small, has been measured under laboratory conditions. Use de Broglie's equation to find an equation for the momentum of a photon of wavelength λ.

23 Why can an electron microscope resolve images in finer detail than an optical microscope?

24 After completing the activity on page 392, reflect on the inquiry question: How is it known that classical physics cannot explain the properties of the atom?

CHAPTER

16 Properties of the nucleus

This chapter examines the nature and properties of alpha, beta and gamma radiation. It also explains the natural causes of nuclear instability, including the importance of the rate of decay, half-life of radioactive substances and penetrating power of the radiation. The half-lives of radioactive substances are also examined. In addition, nuclear fusion and nuclear fission reactions are discussed, including the forces that act within the nucleus and energy transfer and transformation phenomena in the production of nuclear energy.

Content

INQUIRY QUESTION

How can the energy of the atomic nucleus be harnessed?

By the end of this chapter you will be able to:

- analyse the spontaneous decay of unstable nuclei, and the properties of the alpha, beta and gamma radiation emitted (ACSPH028, ACSPH030) ICT
- examine the model of half-life in radioactive decay and make quantitative predictions about the activity or amount of a radioactive sample using the following relationships:
  - $N_t = N_0 e^{-\lambda t}$
  - $\lambda = \frac{\ln(2)}{t_{1/2}}$

  where N_t = number of particles at time t
  N_0 = number of particles present at $t = 0$
  λ = decay constant
  $t_{1/2}$ = time for half the radioactive amount to decay
- model and explain the process of nuclear fission, including the concepts of controlled and uncontrolled chain reactions, and account for the release of energy in the process (ACSPH033, ACSPH034) ICT
- analyse relationships that represent conservation of mass–energy in spontaneous and artificial nuclear transmutations, including alpha decay, beta decay, nuclear fission and nuclear fusion (ACSPH032) ICT N
- account for the release of energy in the process of nuclear fusion (ACSPH035, ACSPH036) ICT
- predict quantitatively the energy released in nuclear decays or transmutations, including nuclear fission and nuclear fusion, by applying: (ACSPH031, ACSPH035, ACSPH036) ICT N
  - the law of conservation of energy
  - mass defect
  - binding energy
  - Einstein's mass–energy equivalence relationship ($E = mc^2$).

16.1 Radioactive decay

PHYSICS INQUIRY N CCT

Sweet decay

How can the energy of the atomic nucleus be harnessed?

COLLECT THIS...

- 50 button-shaped lollies with a mark on one side, such as Skittles or M&M's. Coins could also be used.
- cup that fits all 50 lollies in it
- A3 piece of paper

DO THIS...

1 Fold the edges of the piece of paper up to create a shallow container to contain the lollies.

2 With 50 lollies in the cup, tip them into the paper container. Record the number of lollies that have the mark facing up. Remove the lollies without a mark facing up, and put them to the side. Do not eat them yet.

3 Place the remaining lollies back in the cup and tip them out again. Record the number of lollies with a mark facing up, and remove the ones without the mark facing up.

4 Repeat step 3 until no lollies are left.

RECORD THIS...

Draw a graph of the percentage of sweets left after each throw.

Describe the proportion of lollies that was removed after each throw and the resulting shape of the graph.

REFLECT ON THIS...

How can the energy of the atomic nucleus be harnessed?

What is the probability of a particular lolly being discarded after each throw?

Would changing the initial number of lollies affect the graph?

Can it be predicted when a particular lolly will be discarded?

Human senses cannot detect the radiation from radioactive atoms. High-energy radiation in higher than normal doses can be damaging to living tissue. Radiation and radioactive elements can also be used in a variety of applications that are beneficial. These radioactive atoms, or radioisotopes, will be discussed in this section.

ATOMS

GO TO ➤ Section 14.2 page 383

If an atom is **radioactive**, it will spontaneously emit **radiation** from its nucleus. Figure 16.1.1 shows this radiation emitted in the form of particles and electromagnetic energy (light). To understand radiation and radioactivity, it is necessary to know about the structure of the atom. Recall from Chapter 14 that the central part of an atom, the nucleus, consists of particles known as protons and neutrons. Collectively, these particles are called **nucleons** and are almost identical in mass and size.

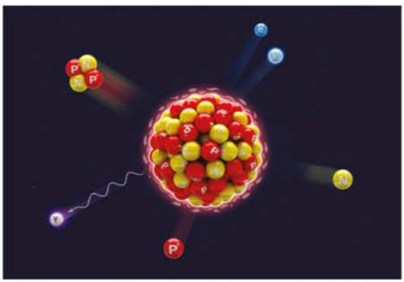


FIGURE 16.1.1 A radioactive nucleus is capable of emitting different types of radiation, although not all simultaneously.

The nucleons have very different electrical properties. Protons have a positive charge. Neutrons are electrically **neutral** so they have no charge. The nucleus contains nearly all of the atom's mass.

Most of the atom is empty space occupied only by negatively charged particles called electrons. These are much smaller and lighter than protons or neutrons. Figure 16.1.2 shows the structure of a typical atom.

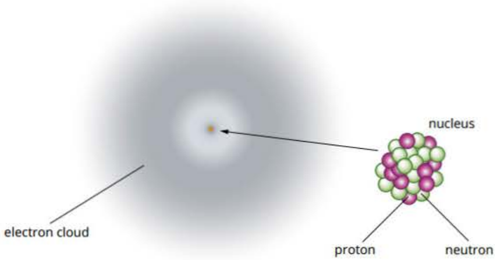


FIGURE 16.1.2 The nucleus of an atom occupies about 10^{-12} of the volume of the atom, yet it contains more than 99% of its mass. Atoms are mostly empty space. (Note that this atom is not drawn to scale.)

A particular atom can be identified by using atomic symbols that have the format shown in Figure 16.1.3.

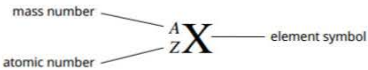


FIGURE 16.1.3 Atomic notation.

The **mass number** (A) is the total number of protons and neutrons in the nucleus.

The **atomic number** (Z) is the number of protons in the nucleus.

Atoms with the same number of protons belong to the same element. For example, if an atom has six protons in its nucleus (i.e. $Z = 6$), then the atom must be carbon. The number of neutrons does not affect which element the atom is, but it does affect the mass of the atom. Figure 16.1.4 shows how the size of the nucleus depends on the mass number. The more protons and neutrons there are in a nucleus, the heavier and larger it is.

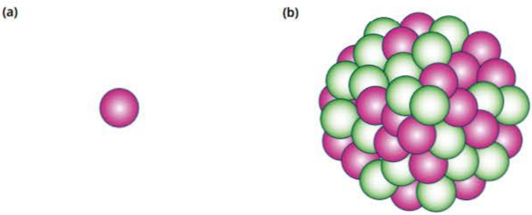


FIGURE 16.1.4 (a) and (b) are both nuclei; however, the hydrogen atom (a) has a very different size nucleus to a uranium atom (b). (Note: the nuclei are not drawn to scale.)

In an electrically neutral atom the number of electrons is equal to the number of protons. For example, any neutral atom of uranium ($Z = 92$) has 92 protons in the nucleus and 92 electrons in the electron cloud.

ISOTOPES

All atoms of a particular element will have the same number of protons, but may have a different number of neutrons. For example, lithium exists naturally in two different forms. One form has three protons and three neutrons. The other has three protons and four neutrons. These different forms of lithium are called **isotopes** of lithium. Isotopes have the same chemical properties but different physical properties such as density and volume. These isotopes are illustrated in Figure 16.1.5.

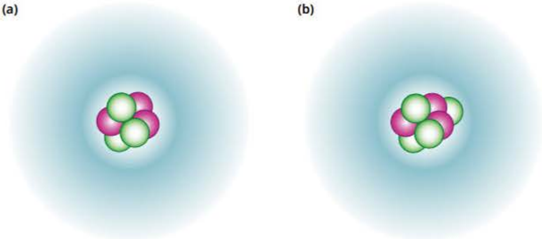


FIGURE 16.1.5 Two different isotopes of lithium: (a) $^{6}_{3}$Li and (b) $^{7}_{3}$Li.

The term **nuclide** is used when referring to a particular nucleus. For example, lithium-6 is a nuclide which has three protons and three neutrons.

There are three isotopes of hydrogen: the nuclide with one proton is called hydrogen; the nuclide with one proton and one neutron is called deuterium; and the nuclide with one proton and two neutrons is called tritium.

PHYSICSFILE L

Heavy water

A compound of oxygen and deuterium has identical chemical properties to ordinary water. However, the molecular mass of ordinary water is about 18 (16 + 1 + 1), while the molecular mass of water containing deuterium is 20 (16 + 2 + 2). Thus, water that contains deuterium has a higher density (by about 11%) and is commonly known as 'heavy water'. One use of heavy water is in nuclear reactors.

Worked example 16.1.1

WORKING WITH ISOTOPES

Consider the isotope of molybdenum, $^{95}_{42}$Mo. Work out the number of protons, nucleons and neutrons in this isotope.

Thinking	Working
The lower number is the atomic number.	atomic number = 42 This nuclide has 42 protons.
The upper number is the mass number. This indicates the number of particles in the nucleus, i.e. the number of nucleons.	mass number = 95 This nuclide has 95 nucleons.
Subtract the atomic number from the mass number to find the number of neutrons.	This isotope has 95 – 42 = 53 neutrons.

Worked example: Try yourself 16.1.1

WORKING WITH ISOTOPES

Consider the isotope of thorium, $^{230}_{90}$Th. Work out the number of protons, nucleons and neutrons in this isotope.

RADIOISOTOPES

Most of the atoms that make up the world around us are stable. Their nuclei have not altered in the billions of years since they were formed. These atoms will stay unchanged in the future. There are about 270 stable isotopes in nature. Tin ($Z = 50$) has ten stable isotopes, while aluminium ($Z = 13$) has just one.

There are also many naturally occurring isotopes that are unstable. An unstable nucleus may spontaneously become more stable by emitting a particle and as a result change into a different element or isotope. Unstable atoms are radioactive. An individual radioactive isotope is known as a **radioisotope**. Carbon has two stable isotopes: carbon-12 and carbon-13. Carbon also has one naturally occurring isotope that is unstable: carbon-14. The nucleus of a carbon-14 atom may spontaneously decay into a different substance, emitting high-energy particles that can be harmful. A known radioactive substance is identified by the radiation warning symbol shown in Figure 16.1.6.

FIGURE 16.1.6 This symbol is used to label and identify a radioactive source.

Figure 16.1.7 shows that every isotope of every element with an atomic mass greater than that of bismuth ($Z = 83$) is radioactive. The first 92 elements are naturally occurring.

Group

Period	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Period 1					1 H 1.008													2 He 4.003
2	3 Li 6.941	4 Be 9.012											5 B 10.81	6 C 12.01	7 N 14.01	8 O 16.00	9 F 19.00	10 Ne 20.18
3	11 Na 22.99	12 Mg 24.31											13 Al 26.98	14 Si 28.09	15 P 30.97	16 S 32.07	17 Cl 35.45	18 Ar 39.95
4	19 K 39.10	20 Ca 40.08	21 Sc 44.96	22 Ti 47.87	23 V 50.94	24 Cr 52.00	25 Mn 54.94	26 Fe 55.85	27 Co 58.93	28 Ni 58.69	29 Cu 63.55	30 Zn 65.38	31 Ga 69.72	32 Ge 72.64	33 As 74.92	34 Se 78.96	35 Br 79.90	36 Kr 83.80
5	37 Rb 85.47	38 Sr 87.61	39 Y 88.91	40 Zr 91.22	41 Nb 92.91	42 Mo 95.96	43 Tc (98)	44 Ru 101.1	45 Rh 102.9	46 Pd 106.4	47 Ag 107.9	48 Cd 112.4	49 In 114.8	50 Sn 118.7	51 Sb 121.8	52 Te 127.60	53 I 126.9	54 Xe 131.3
6	55 Cs 132.9	56 Ba 137.3	57 La 138.9	72 Hf 178.5	73 Ta 180.9	74 W 183.9	75 Re 186.2	76 Os 190.2	77 Ir 192.2	78 Pt 195.1	79 Au 197.0	80 Hg 200.6	81 Tl 204.4	82 Pb 207.2	83 Bi 209.0	84 Po (210)	85 At (210)	86 Rn (222)
7	87 Fr (223)	88 Ra (226)	89 Ac (227)	104 Rf (261)	105 Db (262)	106 Sg (263)	107 Bh (264)	108 Hs (267)	109 Mt (268)	110 Ds (271)	111 Rg (272)	112 Cn (285)	113 Nh (284)	114 Fl (289)	115 Mc (289)	116 Lv (292)	117 Ts (294)	118 Og (294)

Lanthanides

57 La 138.9	58 Ce 140.1	59 Pr 140.9	60 Nd 144.2	61 Pm (145)	62 Sm 150.4	63 Eu 152.0	64 Gd 157.3	65 Tb 158.9	66 Dy 162.5	67 Ho 164.9	68 Er 167.3	69 Tm 168.9	70 Yb 173.1	71 Lu 175.0

Every isotope of these elements is radioactive

Actinides

89 Ac	90 Th 232.0	91 Pa (231.0)	92 U 238.0	93 Np (237)	94 Pu (244)	95 Am (243)	96 Cm (247)	97 Bk (247)	98 Cf (251)	99 Es (252)	100 Fm (257)	101 Md (258)	102 No (259)	103 Lr (262)

FIGURE 16.1.7 The periodic table of the elements

Most of the elements found on Earth have naturally occurring radioisotopes; there are about 200 of these natural radioisotopes. During the twentieth century, an enormous number of radioisotopes were also artificially produced. Most of the radioisotopes used in industry, in medicine and for scientific research are artificially produced. Artificial radioisotopes are produced in nuclear reactors or particle accelerators.

FIGURE 16.1.8 Marie Curie conducted research of different radioactive substances.

RADIOACTIVITY

Around the turn of the twentieth century, scientists such as Marie Curie (Figure 16.1.8) were investigating the newly discovered radioactive substances polonium and radium. Ernest Rutherford and Paul Villard found that there were three different types of emission from these mysterious substances. They named them alpha, beta and gamma radiation.

Further experiments showed that the alpha and beta emissions were actually particles expelled from the nucleus. Gamma radiation was found to be high-energy electromagnetic radiation (light) also expelled from the nucleus. The term radioactive decay refers to the process that emits these particles and radiation from a nucleus.

The nature of these radiations will be discussed in this section.

Alpha (α) decay

When a heavy unstable nucleus undergoes radioactive decay, it may eject an **alpha particle**. This is a positively charged particle that consists of two protons and two neutrons. An alpha particle, symbol α, is identical to a helium nucleus and can also be written as ${}^{4}_{2}\text{He}$.

Uranium-238 is radioactive and may decay by emitting an alpha particle from its nucleus. Figure 16.1.9 shows the unstable nucleus of uranium-238 ejecting an alpha particle. This can be represented in a nuclear equation, which shows the changes occurring in the nuclei. Electrons are not considered in these equations, only the nucleons. The equation for the alpha decay of uranium-238 is:

$${}^{238}_{92}\text{U} \rightarrow {}^{234}_{90}\text{Th} + {}^{4}_{2}\text{He} + \text{energy}$$

The **parent nucleus** ${}^{238}_{92}\text{U}$ has spontaneously emitted an alpha particle (α) and has changed into a completely different element, ${}^{234}_{90}\text{Th}$. Thorium-234 is called the **daughter nucleus**. The energy released is mostly kinetic energy carried by the fast-moving alpha particle.

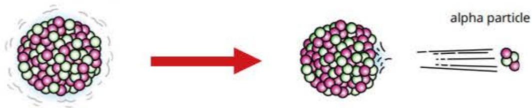


FIGURE 16.1.9 Alpha emission from uranium-238.

In any nuclear reaction, including radioactive decay, atomic and mass numbers are conserved. Energy is released during these decays.

When an atom changes into a different element, it is said to undergo a **nuclear transmutation**. In nuclear transmutations, electric charge is conserved. This results in the conservation of atomic number, i.e. the number of protons. The sums of atomic numbers on both sides of a nuclear equation must be equal. In the uranium decay equation, the atomic number is conserved: $92 = 90 + 2$. The mass number is also conserved: $238 = 234 + 4$.

PHYSICSFILE ICT WE

Radioactive lamps

The wicks or mantles (Figure 16.1.10) used in old-style camping lamps are slightly radioactive. They contain a radioisotope of thorium, an alpha-particle emitter. They have not been banned from sale so far because they contain only small amounts of the radioisotope and can be used safely by taking simple precautions such as washing hands and avoiding inhalation or ingestion.

However, a scientist from the Australian National University in Canberra has called for these mantles to be banned because they tend to crumble and turn to dust as they age. If this dust were inhaled, alpha particles could settle in someone's lung tissue, possibly causing cancers to form.

FIGURE 16.1.10 An old radioactive gas-light mantle.

Beta (β) decay

Many radioactive materials emit **beta particles**. There are two different types of beta particles: beta minus (β^-) and beta plus (β^+).

Beta minus (β^-)

This type of beta decay occurs when an electron is emitted from the nucleus of a radioactive atom, rather than from the electron cloud. This type of beta particle can be written as ${}^{0}_{-1}\beta$.

The atomic number of −1 indicates that the beta particle (the electron) has a single negative charge. The mass number of zero indicates that its mass is far less than that of a proton or a neutron.

Typically, beta-minus decay occurs if a nucleus has too many neutrons to be stable. A neutron spontaneously changes into a proton, a beta-minus particle (β^-, an electron), and an uncharged massless antimatter particle called an **antineutrino** ($\bar{\nu}$). This makes the nucleus more stable.

An example of an isotope that undergoes beta-minus decay is carbon-14. The other isotopes of carbon, i.e. carbon-12 and carbon-13, are both stable. Carbon-14 is unstable. It has too many neutrons and undergoes a beta-minus decay to become stable. In this process, one of the neutrons changes into a proton. Nitrogen-14 is then formed and energy is released. The beta-minus decay of carbon-14 is shown in Figure 16.1.11.

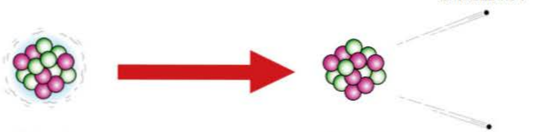


FIGURE 16.1.11 The beta-minus decay of carbon-14.

The nuclear equation for this decay is:

$${}^{14}_{6}\text{C} \rightarrow {}^{14}_{7}\text{N} + {}^{0}_{-1}\beta + \bar{\nu} + \text{energy}$$

The transformation taking place inside the nucleus is:

$${}^{1}_{0}\text{n} \rightarrow {}^{1}_{1}\text{p} + {}^{0}_{-1}\beta + \bar{\nu}$$

Notice that, in all these equations, the atomic and mass numbers are conserved. The antineutrino has no charge and has so little mass that both its atomic and mass numbers are zero.

Beta plus (β^+)

A different form of beta decay occurs when a nucleus has too many protons. In this case, a proton may spontaneously change into a neutron and emit a neutrino (ν) and a positively charged beta particle. This process is known as β^+ (beta-positive) decay. The positively charged beta particle is called a **positron**.

Positrons (${}^{0}_{+1}\beta$) have the same properties as electrons, but their electrical charge is positive rather than negative.

Gamma (γ) decay

After a radioisotope has emitted an alpha or beta particle, the daughter nucleus usually has excess energy. The protons and neutrons in the daughter nucleus then rearrange slightly and offload this excess energy by releasing a **gamma ray**, ${}^{0}_{0}\gamma$.

Gamma rays are high-energy electromagnetic radiation and so have no mass, are uncharged and travel at the speed of light ($3.00 \times 10^8\,m\,s^{-1}$).

A common example of a gamma-ray emitter is iodine-131. It decays by beta and gamma emission to form xenon-131, as shown in Figure 16.1.12.

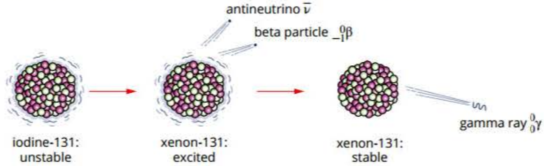


FIGURE 16.1.12 The gamma and beta decay of iodine-131.

The equations for this decay are:

$${}^{131}_{53}I \rightarrow {}^{131}_{54}Xe^* + {}^{0}_{-1}\beta$$

$${}^{131}_{54}Xe^* \rightarrow {}^{131}_{54}Xe + {}^{0}_{0}\gamma$$

The asterisk indicates that the nuclide is in an excited state.

Since gamma rays carry no charge and have no mass, they have no effect when balancing the atomic or mass numbers in a nuclear equation.

Worked example 16.1.2

RADIOACTIVE DECAY

Strontium-90 decays by radioactive emission to form yttrium-90. The equation is:

$${}^{90}_{38}Sr \rightarrow {}^{90}_{39}Y + X$$

Determine the atomic and mass numbers for X and identify the type of radiation being emitted.

Thinking	Working
The mass numbers of 90 are already balanced.	The mass number of X is zero.
Balance the atomic numbers.	$38 = 39 + a$ $a = 38 - 39 = -1$
X has an atomic number of −1 and a mass number of zero.	X is a beta-minus particle, ${}^{0}_{-1}\beta$.

Worked example: Try yourself 16.1.2

RADIOACTIVE DECAY

Polonium-218 decays by emitting an alpha particle and a gamma ray. The nuclear equation is:

$${}^{218}_{84}Po \rightarrow X + {}^{4}_{2}\alpha + \gamma$$

Determine the atomic number and mass number for X, then use the periodic table in Figure 16.1.7 on page 419 to identify the element.

WHY RADIOACTIVE NUCLEI ARE UNSTABLE

In Figure 16.1.13, the stable nuclides that exist in nature are indicated by purple squares. The radioisotopes that are alpha and beta emitters can be identified by the black plus and minus symbols. Most of these radioisotopes also emit gamma radiation.

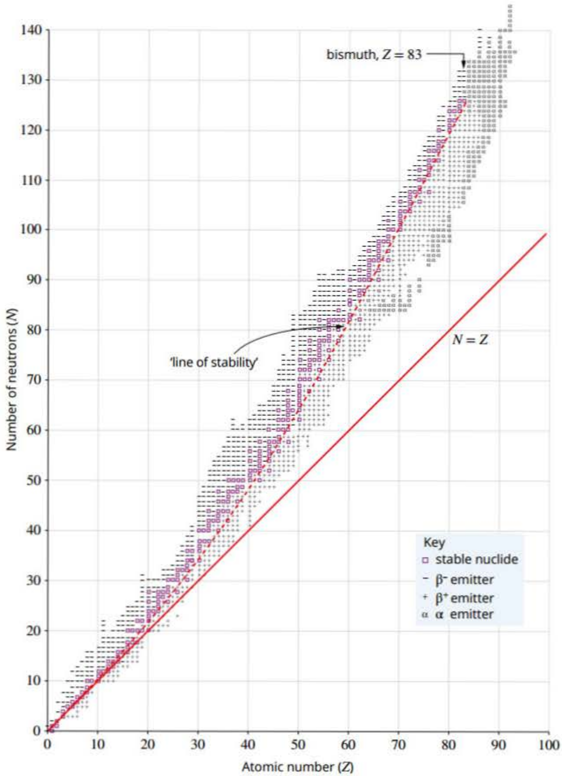


FIGURE 16.1.13 A chart showing stable and radioactive isotopes, plotted according to their number of protons (atomic number) and number of neutrons.

Within the nucleus, protons are in close proximity to other protons. This should seem odd since protons exert strong **electrostatic forces** of repulsion over each other. It might be thought that the nucleus should simply blow apart. As it doesn't, there must be other factors to consider. A force known as the **strong nuclear force** is also acting. This is a force of attraction that holds the nucleus together.

Electrostatic forces act between charged particles and can act over relatively large distances. In the nucleus, this means that each proton strongly repels every other proton so this force is trying to make the nucleus break apart. Neutrons are unaffected by electrostatic forces.

The strong nuclear force is a force of attraction that acts between every nucleon regardless of their charge. This force acts like 'nuclear glue'. Neutrons are attracted to nearby neutrons and protons. Protons are also attracted to nearby neutrons and protons. However, this force only acts over relatively short distances, so for nucleons on the opposite sides of a large nucleus, this force is not significant.

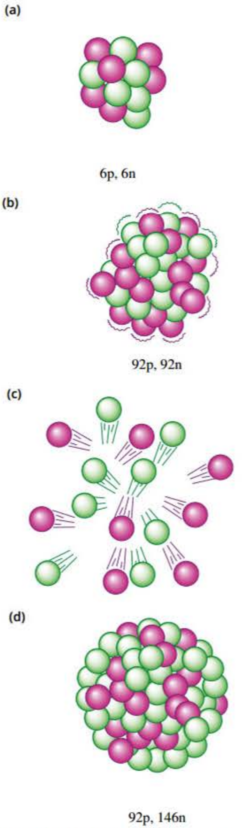


FIGURE 16.1.14 Stable and unstable nuclei. (a) A small nucleus such as carbon-12 is stable. This is because the electrostatic force of repulsion that acts between the protons is overcome by the strong nuclear force of attraction. (b) and (c) A large nucleus with equal numbers of protons and neutrons cannot exist. The electrostatic forces of repulsion between the protons would overcome the strong nuclear forces. (d) Additional neutrons increase the stability of large nuclei. The extra neutrons increase the influence of the strong nuclear force and act like a 'nuclear glue', holding the nucleus together.

In a stable nucleus, there is a delicate balance between the repulsive electric force and the attractive strong nuclear force. For example, bismuth-209, the heaviest stable isotope, has 83 protons and 126 neutrons. Here, the electrostatic repulsion of the protons is balanced by the strong attractive nuclear forces between the nucleons to make the nucleus stable. Compare this with bismuth-211. Its two extra neutrons upset the balance between forces. The nucleus of ^{211}Bi is unstable and it ejects an alpha particle in an attempt to become more stable.

From Figure 16.1.13 on page 423 it is evident that there is a 'line of stability' (indicated by the curved red dashed line on the graph) along which the stable nuclei tend to cluster. Nuclei away from this line are unstable.

For small nuclei with atomic numbers up to about 20, the ratio of neutrons to protons in stable nuclei is close to one. However, as the nuclei become bigger, this ratio increases for stable nuclei. Zirconium ($Z = 40$) has a neutron-to-proton ratio of about 1.25, while for mercury ($Z = 80$) the ratio is close to 1.66. This indicates that for higher numbers of protons, nuclei must have even more neutrons to remain stable. These neutrons act to dilute the repelling forces that exist between the extra protons.

Elements with more protons than bismuth ($Z = 83$) simply have too many repulsive charges in the nucleus. Additional neutrons are unable to stabilise these nuclei. All of these elements are unstable and radioactive. Figure 16.1.14 illustrates stable and unstable nuclei.

PHYSICS IN ACTION ICT WE

Detecting radiation

Our bodies cannot detect alpha, beta or gamma radiation. Therefore, a number of devices have been developed to detect and measure radiation.

A common detector is the **Geiger counter**. This can be used:

- by geologists searching for radioactive minerals such as uranium
- to monitor radiation levels in mines
- to measure the level of radiation after a nuclear accident, such as the accident at Fukushima, Japan, in 2011
- to check the safety of nuclear reactors
- to monitor radiation levels in hospitals and factories.

FIGURE 16.1.15 Thermoluminescent dosimeters are used by doctors, radiologists and scientists who work with radiation to monitor their exposure levels.

People who work in occupations that involve ongoing exposure to levels of ionising radiation usually pin a small radiation-monitoring device to their clothing. This is usually a thermoluminescent dosimeter (TLD), as pictured in Figure 16.1.15. These are used by personnel in nuclear power plants, radiotherapy departments at hospitals, airport security gates and uranium mines.

PHYSICS IN ACTION ICT

How technetium is produced

Technetium-99m is the most widely used radioisotope in nuclear medicine. It is used for diagnosing and treating cancer. However, this radioisotope decays relatively quickly and so usually needs to be produced close to where it will be used. Technetium-99m is produced in small nuclear generators that are located in hospitals around the country. In this process, the radioisotope molybdenum-99 is used as the parent nuclide. Molybdenum-99 decays by beta emission to form a relatively stable (or metastable) isotope of technetium, technetium-99m, as shown below:

$$^{99}_{42}\text{Mo} \rightarrow {}^{99m}_{43}\text{Tc} + {}^{0}_{-1}\beta + \nu$$

Technetium-99m is flushed from the generator using a saline solution. The radioisotope is then diluted and attached to an appropriate chemical compound before being administered to the patient as a tracer. Technetium-99m is purely a gamma emitter. This makes it very useful as a diagnostic tool for locating and treating cancer. Its decay equation is:

$$^{99m}_{43}\text{Tc} \rightarrow {}^{99}_{43}\text{Tc} + \gamma$$

PROPERTIES OF ALPHA, BETA AND GAMMA RADIATION

In the early experiments with radioactivity, emissions from a sample of radium were directed through a magnetic field as shown in Figure 16.1.16. The emissions followed three distinct paths, which suggested that there were three different forms of radiation being emitted. The emissions each had different charges, masses and speeds.

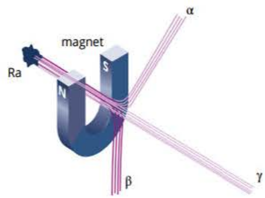


FIGURE 16.1.16 Applying a magnetic field shows that there are three different types of emissions from a radium source.

Alpha (α) particles

Alpha particles consist of two protons and two neutrons. This means that they are relatively heavy and slow moving. Alpha particles are emitted from the nucleus at speeds of up to 20 000 km s^{-1} (2.0×10^7 m s^{-1}), just less than 10% of the speed of light (Figure 16.1.17).

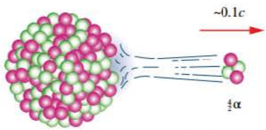


FIGURE 16.1.17 The speed and structure of an α particle.

Alpha particles have a double positive charge. This, combined with their relatively slow speed, makes them very easy to stop. They only travel a few centimetres in air before losing their energy, and will be completely absorbed by a thin card or a human hand (Figure 16.1.18). They have a poor **penetrating ability**.

An example of an isotope that emits α radiation (or undergoes α decay) is the isotope of americium $^{241}_{95}$Am. Americium can be found in ionisation smoke detectors.

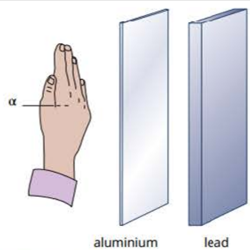


FIGURE 16.1.18 The penetrating ability of alpha radiation.

Beta (β) particles

Beta particles are fast-moving electrons (β^-) or positrons (β^+). Beta-minus particles are created when a neutron decays into a proton. Beta-plus particles are created when a proton decays into a neutron.

Beta particles are much lighter than alpha particles. As a result, they leave the nucleus with far higher speeds—up to 90% of the speed of light (c), as shown in Figure 16.1.19.

FIGURE 16.1.19 The speed and structure of a beta particle.

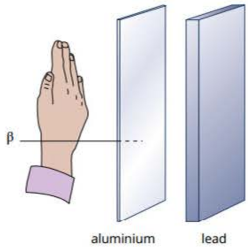


FIGURE 16.1.20 The penetrating ability of beta radiation.

Beta particles are more penetrating than alpha particles. They are faster and have a smaller charge than alpha particles. Beta particles will travel a few metres through air and through a human hand. Typically, a sheet of aluminium about 1 mm thick will stop them, as shown in Figure 16.1.20.

Gamma (γ) rays

Gamma rays are electromagnetic radiation with a very high frequency. Figure 16.1.21 shows where gamma rays lie along the electromagnetic spectrum. They have no rest mass and travel at the speed of light: $3.0 \times 10^8\,\mathrm{m\,s^{-1}}$ or $300\,000\,\mathrm{km\,s^{-1}}$ (Figure 16.1.22).

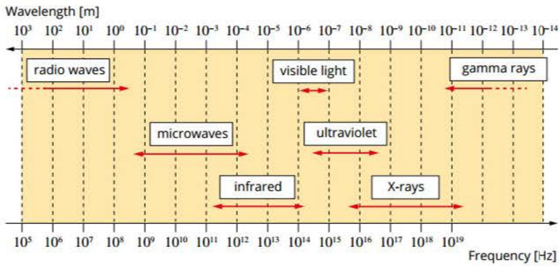


FIGURE 16.1.21 The electromagnetic spectrum contains many different types of radiation that differ in their wavelength and frequency. Gamma rays have very high frequencies and very short wavelengths, making them very energetic and highly penetrating.

FIGURE 16.1.22 The speed and nature of gamma radiation.

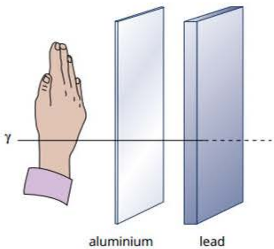


FIGURE 16.1.23 The penetrating ability of γ radiation.

PA 8.1

Gamma rays have no electric charge. Their high energy and uncharged nature make them a very penetrating form of radiation. Gamma rays can travel an almost unlimited distance through air and even through a human hand, an aluminium sheet and a few centimetres of lead (Figure 16.1.23). Even a metre of concrete would not completely absorb a beam of gamma rays.

The properties of alpha (α), beta (β) and gamma (γ) radiation are summarised in Table 16.1.1.

TABLE 16.1.1 A comparison of the properties of alpha, beta and gamma radiation.

Property	α particle	β particle	γ-ray
Mass	heavy	light	none
Speed	up to $20\,000\,\mathrm{km\,s^{-1}}$ or about 10% of the speed of light	about 90% of the speed of light	the speed of light
Charge	+2	−1 or +1	0
Range in air	a few centimetres	1 or 2 m	many metres
Penetration in matter	~10^{-2} mm	a few millimetres	high

16.1 Review

SUMMARY

- The nucleus of an atom consists of positively charged protons and neutral neutrons. Collectively, protons and neutrons are known as nucleons. Negatively charged electrons surround the nucleus.
- The nucleus of the atom is extremely small but contains most of the atom's mass.
- The atomic number, *Z*, is the number of protons in the nucleus. The mass number, *A*, is the number of nucleons in the nucleus, i.e. the combined number of protons and neutrons. Elements are represented as ${}^{A}_{Z}\mathrm{X}$.
- Isotopes of an element have the same number of protons but different numbers of neutrons. Isotopes of an element are chemically identical to each other, but have different physical properties.
- An unstable isotope—a radioisotope—may spontaneously decay by emitting alpha, beta or gamma radiation from its nuclei.
- An alpha particle (α) consists of two protons and two neutrons and is emitted from the nuclei of some radioisotopes. It is identical to a helium nucleus and can be written as ${}^{4}_{2}\mathrm{He}$.
- A beta-minus particle (β^- or ${}^{0}_{-1}\beta$) is an electron that has been emitted from the nucleus of a radioactive atom as a result of a neutron transmutating into a proton. An antineutrino is always emitted with a beta-minus particle.
- A beta-plus particle (β^+) or positron is a positively charged electron (${}^{0}_{+1}\beta$) that has been emitted from the nucleus of a radioactive atom as a result of a proton transmutating into a neutron. A neutrino is always emitted with a beta-plus particle.
- A gamma ray, γ, is high-energy electromagnetic radiation that is emitted from the nuclei of radioactive atoms. Gamma rays always accompany alpha or beta emission.
- In any nuclear reaction, both the atomic and mass numbers are conserved.
- Nuclear stability is the result of the strong nuclear force between nucleons over a very short distance, which opposes electrostatic repulsion between protons.
- Alpha (α) particles are ejected from the nucleus at around 10% of the speed of light. They have a double positive charge and are relatively heavy. Alpha particles have poor penetrating power.
- Beta (β) particles emanate from the nucleus at up to 90% of the speed of light. They are much lighter than α particles. Beta-minus (β^-) particles have a single negative charge. Beta-plus (β^+) particles have a single positive charge. Beta radiation has a moderate penetrating ability.
- Gamma (γ) rays are high-energy electromagnetic radiation and so travel at the speed of light and have no charge. They have high penetrating power.

KEY QUESTIONS

1 What is the collective term used for protons and neutrons?

2 How many protons and how many neutrons are in the ${}^{197}_{79}\mathrm{Au}$ nuclide?

3 How many nucleons are there in the ${}^{235}_{92}\mathrm{U}$ nuclide?

4 How is the number of electrons in a neutral atom determined?

5 Determine the nature of the unknown, X, for the following transmutation:

$${}^{218}_{86}\mathrm{Rn} \rightarrow {}^{214}_{84}\mathrm{Po} + \mathrm{X}$$

6 What type of beta decay occurs in this nuclear reaction?

$${}^{214}_{82}\mathrm{Pb} \rightarrow {}^{214}_{83}\mathrm{Bi} + \mathrm{Y}$$

7 What type of decay occurs when a nucleus has too many protons?

8 Which type of radiation (alpha, beta or gamma):

a can easily penetrate aluminium foil

b is ejected when a neutron decays into a proton

c travels relatively slowly at typically around 10% of the speed of light

d travels at speeds of up to 90% of the speed of light

e has no charge?

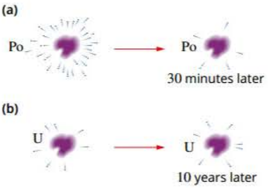


FIGURE 16.2.1 The activity of (a) polonium-218 and (b) uranium-235.

> The half-life ($t_{1/2}$) of a radioisotope is the time that it takes for half of the nuclei of the sample radioisotope to decay.

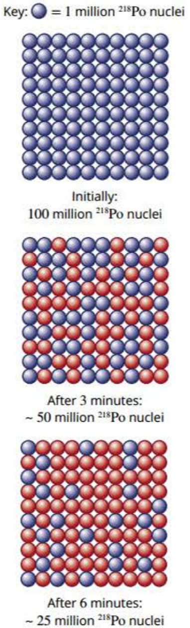


FIGURE 16.2.2 The decay of polonium-218 over two half-lives, showing how only one-quarter (25%) of the original radioisotope remains after two half-lives.

16.2 Half-life

Different radioisotopes will emit radiation and will decay at very different rates. A Geiger counter held close to a small sample of polonium-218 will initially detect a very high level of radiation. Half an hour later, this count rate will have dropped to almost zero.

Compare this with a similar sample of uranium-235. A Geiger counter held close to the uranium will show a low count rate. However, as time passes, the count rate does not seem to change. If you came back decades later, the count level would still be the same. Figure 16.2.1 illustrates this.

The **half-life** of a radioisotope describes how long it takes for half of the atoms in a given mass to decay. The count rate is the **activity** of the sample. These ideas will be studied in this section.

HALF-LIFE

All radioisotopes are unstable but some are more unstable than others. In the previous example, polonium-218 is more unstable than uranium-235. One way of determining this instability is by measuring the half-life ($t_{1/2}$) of the radioisotope.

The half-life of polonium-218 is 3 minutes. Consider a sample of polonium that initially contains 100 million undecayed polonium-218 nuclei, as shown in Figure 16.2.2. Over the first 3 minutes about half of these will have decayed, leaving around 50 million polonium-218 nuclei. Over the next 3 minutes half of these 50 million polonium-218 nuclei will decay, leaving approximately 25 million of the original radioactive nuclei. The process continues as time passes.

> The number of particles remaining at a point in time can be found using the following formulae:
>
> $$N_t = N_0 e^{-\lambda t}$$
>
> $$\lambda = \frac{\ln(2)}{t_{1/2}}$$
>
> where
>
> N_t is the number of radioactive nuclei remaining at time t
>
> N_0 is the initial number of particles present (i.e. $t = 0$)
>
> λ is the decay constant
>
> t is the period of time that the radioactive nuclei has decayed
>
> $t_{1/2}$ is the half-life, i.e. the time for half the radioactive sample to decay

> The number of half-lives in a period of time can be found using:
>
> $$n = \frac{t}{t_{1/2}}$$
>
> where
>
> n is the number of half-lives elapsed
>
> t is the period of time that the radioactive nuclei has decayed
>
> $t_{1/2}$ is the half-life

By combining some of these equations, it is possible to produce an equation for the number of particles remaining in terms of the number of half-lives elapsed:

$$N_t = N_0 e^{-\lambda t} = N_0 e^{-\ln(2) \cdot n} = N_0 \left(\frac{1}{2}\right)^n$$

$$N_t = N_0\left(\frac{1}{2}\right)^n$$

where N_t is the number of radioactive nuclei remaining
N_0 is the initial number of radioactive nuclei
n is the number of half-lives elapsed

As time passes, a smaller and smaller proportion of the original radioisotope remains in the sample. The graph in Figure 16.2.3, known as a decay curve, shows this.

Even a very small radioactive sample will contain billions of atoms. It is important to know that although the behaviour of such a large sample of nuclei can be mathematically predicted, the behaviour of one particular nucleus cannot. It has a 50% chance of decaying in each half-life. The half-life of a radioisotope is constant and cannot be changed by chemical reactions, heat and so on. Half-life is solely determined by the instability of the nuclei of the radioisotope.

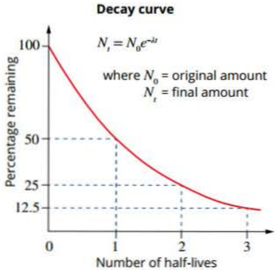
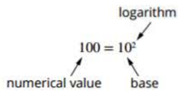


FIGURE 16.2.3 A decay curve for a radioisotope.

SKILLBUILDER N

Natural logs

A logarithm is a way of describing the power or index of a function. For example, in the figure to the right, the 2 is the logarithm with a base of 10. This can be shown with the formula: $\log_{10}(100) = 2$.



The base of the natural logarithm, ln, is the constant e. For example, $\log_e(2) = \ln(2) \approx 0.69$ (to two significant figures).

Scientific calculators may have a button titled ln which can be used to calculate the natural log function.

A few logarithm rules are outlined below:

- $e^{\ln x} = x$
- $\ln(1) = 0$
- $\ln(e) = 1$

Worked example 16.2.1

HALF-LIFE

A sample of the radioisotope thorium-234 contains 8.0×10^{12} nuclei. The half-life of thorium-234 is 24 days.

a Calculate the decay constant for thorium-234.	
Thinking	**Working**
Recall the formula for the decay constant.	$\lambda = \frac{\ln(2)}{t_{1/2}}$
Substitute $t_{1/2}$ into the equation and solve for λ. As the half-life given for thorium is in days, your answer will be in day^{-1}.	$\lambda = \frac{\ln(2)}{t_{1/2}}$ $= \frac{\ln(2)}{24}$ $= 0.029\ \text{day}^{-1}$

b How many thorium-234 nuclei will remain in the sample after 120 days?	
Thinking	**Working**
Recall the formula for the number of nuclei remaining after time t.	$N_t = N_0 e^{-\lambda t}$
Substitute $N_0 = 8.0 \times 10^{12}$ and the decay constant from part (a) into the equation. Calculate the number of nuclei remaining.	$N_t = N_0 e^{-\lambda t}$ $= 8.0 \times 10^{12} \times e^{-0.029 \times 120}$ $= 2.5 \times 10^{11}$ nuclei

WS 8.7

Worked example: Try yourself 16.2.1

HALF-LIFE

A sample of the radioisotope sodium-24 contains 4.0×10^{10} nuclei. The half-life of sodium-24 is 15 hours.

a Calculate the decay constant for sodium-24.

b How many sodium-24 atoms will remain in the sample after 150 hours?

Activity is measured in becquerels, Bq.

1 Bq = 1 disintegration per second.

PHYSICSFILE N WE ICT

Measuring radioactivity

A Geiger counter can be used to record the number of radioactive decays occurring in a sample each second. It measures the activity of the sample.

Over time, the activity of a sample of a radioisotope will decrease. More and more of the radioactive nuclei will have decayed, and at some point it will no longer emit radiation.

If a sample of polonium-218 ($t_{1/2}$ = 3 minutes) has an initial activity of 2000 Bq, then after one half-life its activity will be 1000 Bq. After a further 3 minutes, the activity of the sample will have reduced to 500 Bq, and after a further 3 minutes it will be 250 Bq, and so on.

Uranium-235 has a half-life of 700 000 years. Its activity will remain virtually constant for decades and will certainly not change over 3 minutes.

The same half-life equation can be used to calculate the final activity of a radioactive sample after a number of half-lives.

The activity of the nuclei remaining after a number of half-lives can be found mathematically using:

$$A = A_0\left(\frac{1}{2}\right)^n$$

where

A is the activity of the radioactive nuclei remaining (Bq)

A_0 is the initial activity of the radioactive nuclei (Bq)

n is the number of half-lives elapsed.

PHYSICSFILE ICT

Common radioisotopes and their applications

The half-lives of some common radioisotopes are shown in Table 16.2.1. The half-life of a radioisotope will determine what it is used for. For example, the most commonly used medical tracer, technecium-99, has a short half-life of just 6 hours. The short half-life means that radioactivity does not remain in the body any longer than necessary. On the other hand, the radioisotope used in a smoke detector, americium-241, is chosen because of its long half-life, 461 years. The smoke detector can continue to function for a very long time, as long as the battery is replaced each year.

TABLE 16.2.1 Some common radioisotopes and their half-lives.

Isotope	Emission	Half-life	Application
Natural			
polonium-214	α	0.000 16 s	nothing at this time
strontium-90	β	28.8 years	cancer therapy
radium-226	α	1630 years	once used in luminous paints
carbon-14	β	5730 years	carbon dating of fossils
uranium-235	α	700 000 years	nuclear fuel, rock dating
uranium-238	α	4.5 billion years	nuclear fuel, rock dating
thorium-232	α	14 billion years	fossil dating, nuclear fuel
Artificial			
technetium-99m	β	6 h	medical tracer
sodium-24	β	15 h	medical tracer
iodine-131	β	8 days	medical tracer
phosphorus-32	β	14.3 days	medical tracer
cobalt-60	β	5.3 years	radiation therapy
americium-241	α	460 years	smoke detectors
plutonium-239	α	24 000 years	nuclear fuel, rock dating

DECAY SERIES

PA 8.2 PA 8.3

Generally, when a radionuclide decays, its daughter nucleus is not completely stable and is itself radioactive. This daughter nucleus will then undergo further decay. Eventually a stable isotope is reached and the sequence ends. This is known as a **decay series**. An example of a decay series is shown in Figure 16.2.4. This particular series shows the decay of uranium-238 (shown at the top of the chart) into lead-206 (shown at the bottom of the chart).

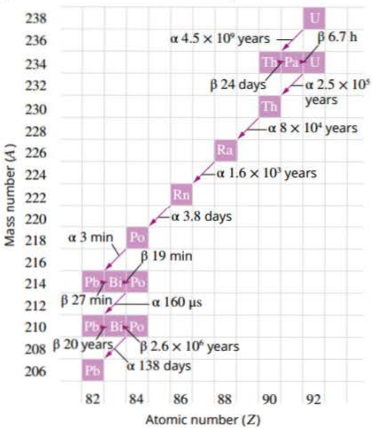


FIGURE 16.2.4 The uranium decay series. The half-life and emissions are indicated on each of the decays as radioactive uranium-238 is gradually transformed into stable lead-206. Mining companies find significant quantities of lead at uranium mines.

The Earth is 4.5 billion years old—old enough to have only four naturally occurring decay series that remain active. These are:

- the uranium series in which uranium-238 eventually becomes lead-206
- the actinium series in which actinium-235 eventually becomes lead-207
- the thorium series in which thorium-232 eventually becomes lead-208
- the neptunium series in which neptunium-237 eventually becomes bismuth-209. (Since neptunium-237 has a relatively short half-life, it is no longer present in the crust of the Earth, but the rest of its decay series is still continuing.)

Geologists analyse the proportions of the radioactive elements in a sample of rock to gain a reasonable estimate of the rock's age. This technique is known as rock dating.

PHYSICS IN ACTION ICT

Radiocarbon dating

Carbon dating is a technique used by archaeologists to determine the ages of fossils and ancient objects that were made from plant matter. This method involves measuring and comparing the proportion of two isotopes of carbon, carbon-12 and carbon-14 in the specimen.

Carbon-12 is a stable isotope but carbon-14 is radioactive. Carbon-14 only exists in trace amounts in nature. Carbon-12 atoms are about 1 000 000 000 000 (10^{12}) times more common than carbon-14 atoms.

Carbon-14 has a half-life of 5730 years and decays by beta-minus emission to nitrogen-14. Its decay equation is:

$$^{14}_{6}\text{C} \rightarrow {}^{14}_{7}\text{N} + \beta^{-} + \overline{\nu}$$

Both carbon-12 and carbon-14 can combine with other atoms in the environment. For example, they both combine with oxygen to form carbon dioxide. Plants and animals take in carbon-based molecules from the air and food. This means that all living organisms contain the same percentage of carbon-14. In the environment, the production of carbon-14 is matched by its decay and so the proportion of carbon-14 atoms to carbon-12 remains constant.

After an organism dies, the amount of carbon-14 it contains will decrease as these atoms decay to form nitrogen-14 and are not replaced from the environment. The number of atoms of carbon-12 does not change as carbon-12 is a stable atom. So, over time, the proportion of carbon-14 to carbon-12 atoms decreases.

The proportion of carbon-14 to carbon-12 in a dead organism can be compared with that found in living organisms and the approximate age of the specimen can be determined from the half-life of carbon-14.

Consider this example. The count rate from a 1 g sample of carbon that has been extracted from an ancient wooden spear is 10 Bq. A 1 g sample of carbon from a living piece of wood has a count rate of 40 Bq. We can assume that this was also the initial count rate of the spear. For its count rate to have reduced from 40 to 10 Bq, the spear must be (40 → 20 → 10) two half-lives of carbon-14 old, i.e. about 11 500 years old.

In 1988, scientists used carbon-dating techniques to show that the Shroud of Turin (Figure 16.4.5) was probably a medieval forgery. It had been claimed that the Shroud of Turin was the piece of cloth that was the burial shroud of Jesus Christ. Carbon-dating tests on small samples of the cloth established that there was a high probability that it was made between 1260 and 1390, not around the time of Christ's death.

FIGURE 16.2.5 The Shroud of Turin.

16.2 Review

SUMMARY

- The rate of decay of a radioisotope is measured by its half-life, $t_{1/2}$. This is the time that it takes for half of the radioisotope to decay.
- The number of nuclei, N_t, remaining after some time, t, is given by the equation:

$$N_t = N_0 e^{-\lambda t}$$

- The decay constant, λ, can be calculated with the equation:

$$\lambda = \frac{\ln(2)}{t_{1/2}}$$

- The activity of a sample indicates the number of emissions per second. Activity is measured in becquerels (Bq), where 1 Bq = 1 emission per second.
- The number of atoms of a radioisotope will decrease over time. Over one half-life, the number of atoms of a radioisotope will halve.
- When a radionuclide decays, its daughter nucleus is usually itself radioactive. This daughter will then decay to a granddaughter nucleus, which may also be radioactive, and so on. This is called a decay series.

KEY QUESTIONS

1 What is meant by the activity of a radioisotope?

2 Technetium-99m has a half-life of 6.0 hours. A sample of the radioisotope originally contains 8.0×10^{10} atoms. How many technetium-99m nuclei remain after 6.0 hours?

3 Iodine-131 has a half-life of 8 days. A sample of the radioisotope initially contains 2.4×10^{12} iodine-131 nuclei. How many iodine-131 nuclei remain after 24 days?

4 Radioactive materials are considered to be relatively safe when their activity has fallen below 0.1% of the initial value.

a How many half-lives does this take?

b Plutonium-239 is a by-product of nuclear reactors. Its half-life is 24 000 years. How long does the plutonium-239 have to be stored as nuclear waste before it is considered safe to handle?

5 If a particular atom in a radioactive sample has not decayed during the previous half-life, what is the percentage chance that it will decay in the next half-life?

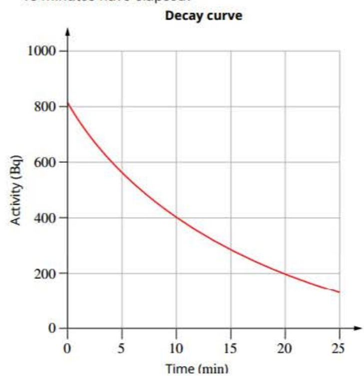
6 A hospital in Alice Springs needs 12 µg of the radioisotope technetium-99m. The specimen has to be ordered from Sydney. The half-life of technetium-99m is 6 hours and the delivery takes 24 hours. How much must be produced in Sydney to satisfy the Alice Springs order?

7 The activity of a radioisotope changes from 6000 Bq to 375 Bq over a period of 60 minutes. Calculate the half-life of this radioisotope.

8 A Geiger counter is used to measure the radioactive emissions from a certain radioisotope. The activity of the sample is shown in the graph.

a What is the half-life of the radioisotope according to the graph?

b What would the activity of the sample be after 40 minutes have elapsed?



9 According to Figure 16.2.4 on page 431, what type of decay does lead-210 undergo and what is its half-life?

10 In the uranium decay series shown in Figure 16.2.4, $^{234}_{92}U$ decays to eventually produce stable $^{206}_{82}Pb$. How many alpha and beta-minus decays have occurred?

DS 8.1

16.3 Nuclear fission and fusion

In 1905, Albert Einstein theorised that mass, m, and energy, E, are equivalent through the equation $E = mc^2$. This led to the realisation that vast amounts of energy lie unharnessed within the nuclei of atoms. The ramifications of Einstein's work and the discovery of nuclear fission were realised in 1945 with the explosion of the first atomic bomb in the desert near Alamogordo in New Mexico, USA. In this section, nuclear fission and the energy that it can unleash (Figure 16.3.1) will be explored.

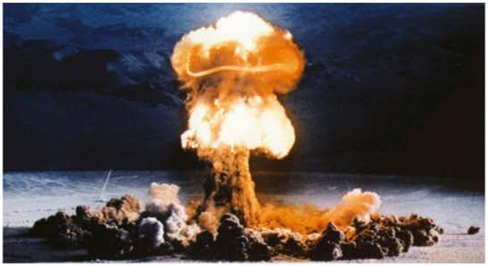


FIGURE 16.3.1 An atomic bomb explosion and its associated mushroom cloud. Nuclear fission can release huge amounts of energy.

INSIDE THE NUCLEUS

The current understanding of the basic properties and structure of the nucleus is the result of intense scientific investigation in the early part of the twentieth century. Physicists such as Becquerel, Rutherford, Chadwick, Geiger, Marsden and Harkins were instrumental in the development of the model of the nucleus that exists today. These renowned scientists are shown in Figure 16.3.2.

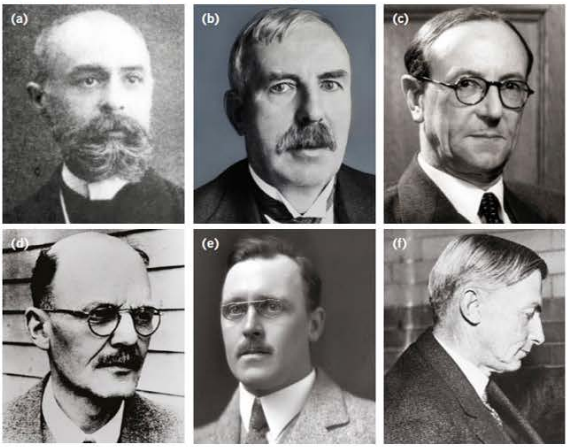


FIGURE 16.3.2 (a) Henri Becquerel, (b) Ernest Rutherford, (c) James Chadwick, (d) Hans Geiger, (e) Ernest Marsden and (f) William Harkins.

Recall that within the nucleus, there are protons in close proximity to other protons, explained by the strong nuclear force (see Section 16.1). An example of how the electrostatic and strong nuclear forces act in a nucleus is shown in Figure 16.3.3.

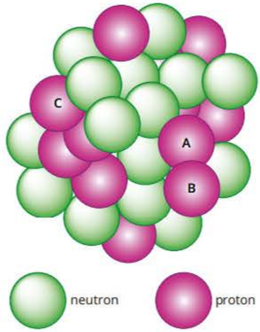


FIGURE 16.3.3 Proton A both attracts and repels proton B, but, at short distances, the attraction due to the strong nuclear force is much greater than the repulsion due to the electrostatic force. Proton A also both attracts and repels proton C, but because of the greater distance between them, the force of repulsion is larger. However, proton A and proton C do not fly apart due to the strong attractive forces exerted on them by adjacent neutrons.

NUCLEAR FISSION

The discovery of the neutron by James Chadwick in 1932 enabled scientists to explore the behaviour of larger atomic nuclei. Up until then, physicists such as Enrico Fermi had been firing alpha particles at target nuclei and analysing the results. Chadwick found that with larger target nuclei, the positive alpha particles were too strongly repelled from the positively charged nuclei and collisions did not occur.

The advantage of a neutron is that it is neutral and so is not repelled by any target nucleus. The bombarding neutrons can be absorbed into the nucleus of the target atom, as shown in Figure 16.3.4. This makes neutrons very useful as a form of radiation. They are used in many experiments to artificially transmutate (change the form of) different isotopes.

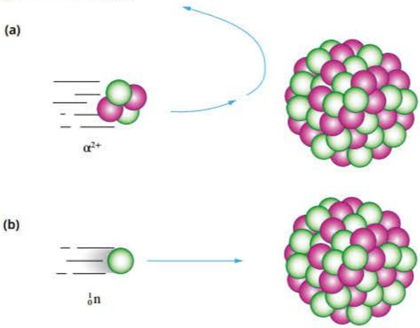


FIGURE 16.3.4 (a) Charged alpha particles are repelled by a nucleus. (b) Uncharged neutrons are able to smash into a nucleus.

Nuclear **fission** occurs when an atomic nucleus splits into two or more pieces. This is usually triggered or induced by the absorption of a neutron, as shown in Figure 16.3.5. Nuclides that are capable of undergoing nuclear fission after absorbing a neutron are said to be fissile. Only a handful of fissile nuclides exist in nature.

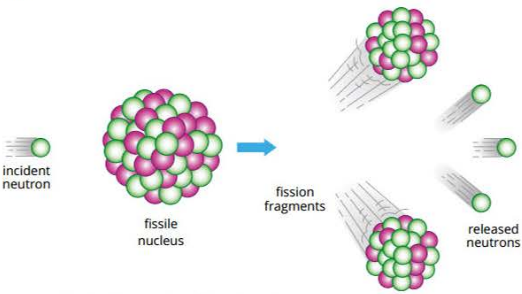


FIGURE 16.3.5 Nuclear fission is the splitting of a nucleus.

PHYSICSFILE ICT

Strong nuclear force

The existence of the strong nuclear force was first proposed by Japanese theoretical physicist Hideki Yukawa in 1935. However, the properties of this force are so complex that it took until 1975 for physicists to develop a mathematical model that could successfully describe it.

Uranium-235 and plutonium-239 are fissile and can be made to split when bombarded by a slow-moving neutron. Uranium-238 and thorium-232 require a very high-energy neutron to induce fission, so they are regarded as fissionable, but non-fissile.

RELEASE OF NEUTRONS DURING FISSION

Uranium-235 and plutonium-239 are the fissile nuclides most commonly used in nuclear reactors and nuclear weapons. They are more fissile than uranium-238 and thorium-232.

When a uranium-235 or plutonium-239 nucleus absorbs either a slow- or fast-moving neutron, it becomes unstable and spontaneously undergoes fission. However, fission is more likely to be induced by a slow-moving neutron because it is more easily captured.

A uranium-235 nucleus may split in many different ways. When a sample of uranium-235 undergoes fission, a wide variety of fission products are produced. Figure 16.3.6 shows one outcome, but many others are possible. Usually either two or three neutrons are released. For uranium-235, an average of 2.47 neutrons per fission has been determined.

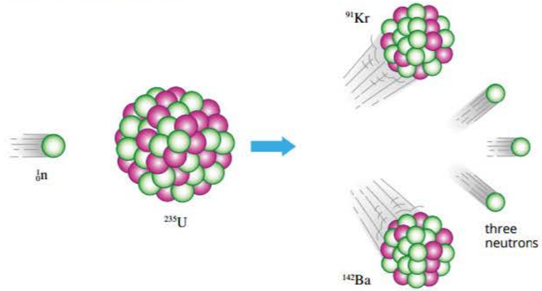


FIGURE 16.3.6 One possible outcome for the neutron-induced fission of uranium-235. This example shows three neutrons released along with krypton-91 and barium-142 daughter nuclides.

A typical fission reaction for uranium-235 is:

$$^{1}_{0}n + ^{235}_{92}U \rightarrow ^{236}_{92}U \rightarrow ^{91}_{36}Kr + ^{142}_{56}Ba + 3^{1}_{0}n + \text{energy}$$

Krypton-91 and barium-142 are known as the daughter nuclei or **fission fragments**. Three neutrons are freed from this uranium nucleus when it splits. Note that, in the same way as for radioactive decay, both the atomic number, Z, and mass number, A, are conserved in these nuclear reactions. For the reaction equation shown, the atomic numbers on either side of the arrows add up to 92 and the mass numbers add up to 236.

In the end, the decay products of the nuclear fission process form a lethal cocktail of radioactive isotopes. It is these radioactive fission fragments that comprise the bulk of the high-level waste produced by nuclear reactors.

Plutonium-239 will also undergo fission in a variety of ways. It releases an average of 2.89 neutrons per fission, slightly more than uranium-235.

There are two types of nuclear fission reactions: uncontrolled and controlled. During a nuclear fission reaction, neutrons are ejected: these neutrons can create further nuclear fission reactions. During an uncontrolled nuclear fission reaction, the subsequent nuclear fission reactions are allowed to proceed without any moderation (such as removing neutrons released during the reaction). During a controlled nuclear fission reaction, the subsequent nuclear fission reactions are moderated and controlled by removing ejected neutrons.

PHYSICS IN ACTION ICT

Enrico Fermi, Lise Meitner and nuclear fission

Enrico Fermi, pictured in Figure 16.3.7, was born in Italy in 1901. He completed his doctorate and post-doctorate work in physics at the University of Pisa and in Germany. Fermi had immigrated to the USA by the time the nuclear age dawned in the 1930s. The neutron had just been discovered in 1932, which enabled scientists to fire neutral particles at atomic nuclei for the first time. Fermi was at the forefront of this research.

Fermi bombarded uranium-238 atoms with neutrons and found that uranium-238 nuclei absorbed the neutrons and formed a radioactive isotope of uranium. This isotope then decayed by emitting a beta-minus particle to become neptunium, which then emitted another beta-minus particle to become plutonium, two completely undiscovered elements. Fermi had successfully produced the world's first artificial and transuranic (i.e. after uranium) elements. The nuclear reactions for this process are:

$$^{1}_{0}n + ^{238}_{92}U \rightarrow ^{239}_{92}U$$

$$^{239}_{92}U \rightarrow ^{239}_{93}Np + ^{0}_{-1}\beta$$

$$^{239}_{93}Np \rightarrow ^{239}_{94}Pu + ^{0}_{-1}\beta$$

FIGURE 16.3.7 Enrico Fermi.

In 1938, following on from Fermi's work, two German scientists, Otto Hahn and Fritz Strassmann, were also bombarding uranium ($Z = 92$) in an attempt to produce some transuranic elements ($Z > 92$). They found that, rather than producing larger elements, they were getting isotopes of barium ($Z = 56$). Hahn wrote to his colleague Lise Meitner, pictured in Figure 16.3.8, about this unexpected result. She then discussed this with her nephew Otto Frisch, a nuclear physicist, and realised that the bombarding neutrons were causing the uranium nuclei to split. If barium ($Z = 56$) was one of the products, then krypton ($Z = 36$) must be another. This was found to be the case. It was Frisch who coined the term 'fission' and Meitner who proposed that energy would be released during this process.

After the start of World War II, Enrico Fermi was commissioned by President Roosevelt to design and build a device that would sustain the fission process in the form of a chain reaction. In 1942, Fermi succeeded in this task. A squash court at the University of Chicago was used as the site for the world's first nuclear reactor. It produced less than 1 W of power—not even enough to power a small light globe! This sounds like a bit of a failure, but, in fact, achieving fission for the first time was a very important breakthrough. The reactor was later modified to produce about 200 W. Fermi died of cancer in 1954. One year after his death, the element with atomic number 100 was artificially produced and named fermium, Fm, in his honour.

FIGURE 16.3.8 Lise Meitner.

NUCLEAR FUSION

Nuclear fusion is a process that has been occurring inside the Sun and other stars for billions of years. Fusion involves the combining of small nuclei such as hydrogen and helium to form a larger nucleus. The example of nuclei fusing to form a helium atom is shown in Figure 16.3.9a on page 438. The amount of energy released per nucleon with fusion is greater than with fission, and there is no radioactive waste produced. Scientists are working on experimental fusion reactors such as ITER (International Thermonuclear Experimental Reactor) in France and the National Ignition Facility in the USA.

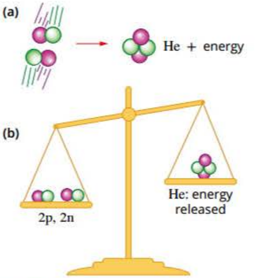


FIGURE 16.3.9 (a) When two isotopes of hydrogen fuse to form a helium nucleus, energy is released. (b) The mass of the reactants is slightly greater than the mass of the products when the nuclei combine during fusion.

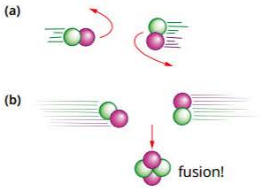


FIGURE 16.3.10 (a) Slow-moving nuclei do not have enough energy to fuse together. The electrostatic forces cause them to be repelled from each other. (b) If the nuclei have sufficient kinetic energy, then they will overcome the repulsive forces and move close enough together for the strong nuclear force to come into effect. At this point, fusion will occur and energy will be released.

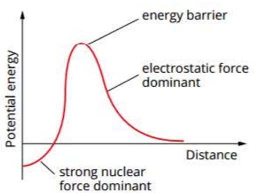


FIGURE 16.3.11 If two hydrogen-2 (deuterium) nuclei are to get close enough for the strong nuclear force to act, they must overcome the energy barrier presented by the electrostatic force.

The fusion process created so far has only lasted for a fraction of a second, and it is not expected that a fusion reactor will successfully operate for many years. Researchers at Lockheed Martin in the USA are working on a compact fusion reactor. In 2014 they claimed a prototype will be running by 2019. This claim has been met with scepticism by some in the scientific community.

As in the cases of radioactive decay and nuclear fission, the mass of the reactants is slightly greater than the mass of the products when the nuclei combine during fusion. This mass difference is represented by the unbalanced scales shown in Figure 16.3.9b.

Nuclear fusion is a very difficult process to recreate. The main problem is that nuclei are positively charged. They exert an electrostatic force of repulsion on each other; that is, they push each other away. As such, it is not easy to force the nuclei together. Remember that the electrostatic force is a long-range force and the strong nuclear force of attraction only acts at much shorter distances.

As two nuclei approach each other, the electrostatic force will cause them to be repelled. Slow-moving nuclei with relatively small amounts of kinetic energy will not be able to get close enough for the strong nuclear force to come into effect. Fusion will not happen, as can be seen in Figure 16.3.10a.

If the nuclei are travelling towards each other at higher speeds, as shown in Figure 16.3.10b, they may have enough kinetic energy to overcome the repulsive force. The nuclei can now get close enough for the strong nuclear force to start acting. If this happens, fusion will occur.

The graph in Figure 16.3.11 shows the effect of the electrostatic force and the strong nuclear force on the potential energy of a pair of deuterium ($^{2}_{1}H$) nuclei. At large separation distances, the electrostatic force dominates and the nuclei repel each other (shown to the right of the energy barrier in the graph). At small separation distances, the strong nuclear force dominates and the nuclei can fuse together. However, to get the nuclei to this point, they need an enormous amount of energy. Temperatures in the hundreds of millions of degrees are required. This enormous amount of energy enables the nuclei to overcome the energy barrier shown in the graph and fuse together.

As in fission, in any fusion reaction the atomic numbers and mass numbers on either side of the equation are conserved. The fusion of two hydrogen-2 (deuterium) nuclei is shown below:

$$^{2}_{1}H + ^{2}_{1}H \rightarrow ^{3}_{1}H + ^{1}_{1}H$$

The atomic numbers add up to two on both sides and the mass numbers add up to four on both sides. However, the total mass of the reactants will be greater than the total mass of the products.

PHYSICSFILE EU S ICT

Hydrogen bomb

In 1952, a fusion reaction was used to power the world's first hydrogen bomb. It had five times the destructive power of all the conventional bombs that were dropped during the whole of World War II. The high temperature achieved by a fissile fuel explosion was used to initiate the fusion reaction. In other words, an atomic bomb was used as the fuse for the hydrogen fusion bomb. A hydrogen bomb dropped at Bikini Atoll in 1956 is shown in Figure 16.3.12.

FIGURE 16.3.12 The hydrogen bomb dropped at Bikini Atoll in 1956.

16.3 Review

SUMMARY

- Within a nucleus, forces of attraction and repulsion are acting. The long-distance electrostatic force of repulsion acts between the protons. A short-distance strong nuclear force of attraction acts between every nucleon.
- Nuclear fission occurs when a nucleus is made to split into two or more fragments and release a number of neutrons. This can be induced by striking a fissile nucleus with a neutron. A relatively large amount of energy is released during this process.
- Nuclear fusion is the combining of light nuclei to form heavier nuclei. Extremely high temperatures are required for fusion to occur.
- A small amount of mass is lost during a fusion reaction.
- Nuclei are positively charged and so repel each other due to electrostatic forces. In a fusion reaction, approaching nuclei must have enough speed to overcome the electrostatic forces and get close enough for the strong nuclear force to take effect. The energy that is required to achieve this is called the energy barrier.
- The amount of energy released per nucleon is greater for fusion than for fission.

KEY QUESTIONS

1 What is the strong nuclear force?

2 Why is dangerous waste produced by nuclear reactors?

3 Consider one particular neutron in the nucleus of a gold atom. Describe the forces that the neutron experiences from other nucleons.

4 Which of these nuclides below are fissile and which are non-fissile?
cobalt-60, uranium-235, uranium-238, plutonium-239

5 Determine the number of neutrons (a) released in this fission reaction:

$$^{1}_{0}n + ^{235}_{92}U \rightarrow ^{148}_{57}La + ^{85}_{35}Br + a^{1}_{0}n$$

6 How does fusion differ from fission?

7 Two slow-moving protons are travelling directly towards each other. Explain whether the protons will collide and fuse together.

8 Two fast-moving protons are travelling directly towards each other. The protons collide and fuse together. Explain why this happens.

9 What happens to the number of nucleons during the fusion reaction below?

$$^{2}_{1}H + ^{3}_{1}H \rightarrow ^{4}_{2}He + ^{1}_{0}p$$

16.4 Energy from nuclear reactions

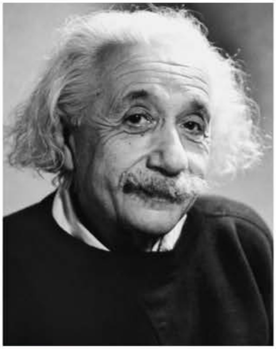


FIGURE 16.4.1 Albert Einstein.

GO TO ➤ Section 12.3 page 325

It is well established that the mass of any nucleus is always less than the mass of its individual nucleons. Two separate protons and two separate neutrons will have slightly more total mass than a helium nucleus.

Albert Einstein, pictured in Figure 16.4.1, provided the explanation of the origins of this missing mass. He showed that mass and energy were not completely independent quantities. Indeed, mass can be converted into energy and energy can be converted into mass.

ENERGY RELEASED DURING NUCLEAR REACTIONS

The energy released as a result of a mass defect (mass decrease) is given by Einstein's famous equation, introduced in Chapter 12:

> $E = mc^2$
>
> where
>
> E is energy (J)
>
> m is the mass defect (the decrease in mass; kg)
>
> c is the speed of light ($3.00 \times 10^8\,\mathrm{m\,s^{-1}}$)

The chemical reactions that you have probably performed at school typically release only a few electronvolts of energy. Compared with this, an enormous amount of energy is released during nuclear reactions. This has made nuclear energy the focus of scientific research over the past century.

During radioactive decay, millions of electronvolts of energy can be released. Alpha-particle decay usually involves the release of 5–10 MeV (5–10 million electronvolts) of energy. Nuclear fission involves much more energy again—typically around 200 MeV. This energy is mainly in the form of the kinetic energy of the fission fragments and neutrons, as well as the emission of energy as gamma radiation.

During any fission reaction, the combined mass of the incident neutron and the target nucleus is always slightly greater than the combined mass of the fission fragments and the released neutrons. For example, in Figure 16.4.2, the mass of the incident neutron and the uranium-235 nucleus is greater than the combined masses of the fission products—barium-142, krypton-91 and the three neutrons.

> An electronvolt is the energy that an electron would gain if it were accelerated by a voltage of 1 volt and is equal to 1.602×10^{-19} J.
>
> To convert from eV to joules: multiply by 1.602×10^{-19} J
>
> To convert from joules to eV: divide by 1.602×10^{-19} J

Only a very small proportion of the original mass of the nuclei is available as usable energy—typically around 0.1%. If you had a 1.000 kg block of pure uranium-235 that underwent fission completely, at the end you would have a block of radioactive fission fragments with a mass of around 0.999 kg.

> The energy of moving objects such as cars and tennis balls is measured in joules. However, nuclei, subatomic particles and radioactive emissions have such small amounts of energy that the joule is inappropriate.
>
> The energy of subatomic particles and radiation is usually given in electronvolts (eV). One electronvolt is an extremely small amount of energy.

> To calculate the percentage of mass that is converted into energy during fission or fusion, use the equation:
>
> $\%\ \text{mass decrease} = \frac{\text{mass defect}}{\text{initial mass}} \times \frac{100}{1}$

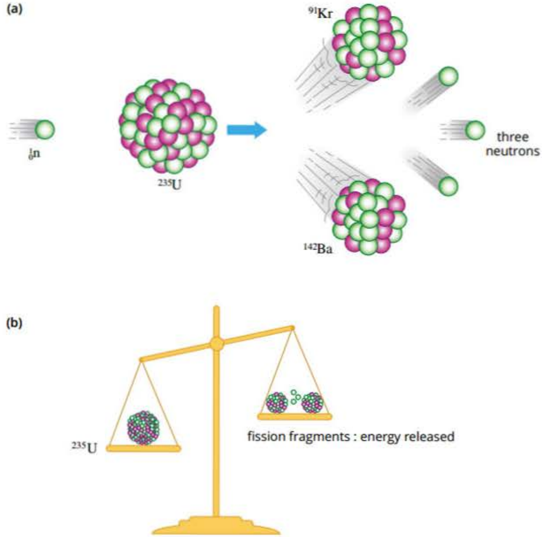


FIGURE 16.4.2 (a) Each fission of a uranium-235 nucleus releases about 200 MeV of energy. (b) The fission fragments have a lower combined mass than the uranium nucleus. The missing mass is converted into the 200 MeV of energy according to the equation $E = mc^2$.

Worked example 16.4.1

FISSION

Plutonium-239 is a fissile material. When a plutonium-239 nucleus is struck by and absorbs a neutron, it can split in many different ways. Consider the example of a nucleus that splits into barium-145 and strontium-93 and releases some neutrons.

The nuclear equation for this is:

$$^1_0n + ^{239}_{94}Pu \rightarrow ^{145}_{56}Ba + ^{93}_{38}Sr + a^1_0n + \text{energy}$$

a How many neutrons are released during this fission process, i.e. what is the value of *a*?	
Thinking	**Working**
Analyse the mass numbers (Z).	$1 + 239 = 145 + 93 + (a \times 1)$ $a = (1 + 239) - (145 + 93)$ $= 2$ Two neutrons are released during fission.

b During this single fission reaction, there was a loss of mass (a mass defect) of 3.07×10^{-28} kg. Calculate the amount of energy that was released during the fission of a single plutonium-239 nucleus. Give your answer in both MeV and joules.	
Thinking	**Working**
The energy released during the fission of this plutonium nucleus can be found by using $E = mc^2$.	$E = mc^2$ $= (3.07 \times 10^{-28}) \times (3.00 \times 10^8)^2$ $= 2.76 \times 10^{-11}$ J
To convert J into eV, divide by 1.602×10^{-19}. Remember that 1 MeV = 10^6 eV.	$E = \frac{2.76 \times 10^{-11}}{1.602 \times 10^{-1}}$ $= 1.73 \times 10^8$ eV $= 173$ MeV

c The combined mass of the plutonium nucleus and bombarding neutron was 3.99×10^{-25} kg. What percentage of this initial mass was converted into the energy produced during the fission process?	
Thinking	**Working**
Use the relationship $\%\text{ mass decrease} = \frac{\text{mass defect}}{\text{initial mass}} \times \frac{100}{1}$.	$\%\text{ mass decrease} = \frac{\text{mass defect}}{\text{initial mass}} \times \frac{100}{1}$ $= \frac{3.07 \times 10^{-28}}{3.99 \times 10^{-25}} \times \frac{100}{1}$ $= 0.0769\%$ This is a very small percentage loss in mass.

Worked example: Try yourself 16.4.1

FISSION

Plutonium-239 is a fissile material. When a plutonium-239 nucleus is struck by and absorbs a neutron, it can split in many different ways. Consider the example of a nucleus that splits into lanthanum-143 and rubidium-94 and releases some neutrons.

The nuclear equation for this is:

$$^{1}_{0}\text{n} + ^{239}_{94}\text{Pu} \rightarrow ^{143}_{57}\text{La} + ^{94}_{37}\text{Rb} + a^{1}_{0}\text{n}$$

a How many neutrons are released during this fission process, i.e. what is the value of *a*?

b During this single fission reaction, there was a loss of mass (a mass defect) of 4.58×10^{-28} kg. Calculate the amount of energy that was released during fission of a single plutonium-239 nucleus. Give your answer in both MeV and joules, to three significant figures.

c The combined mass of the plutonium nucleus and bombarding neutron was 2.86×10^{-25} kg. What percentage of this initial mass was converted into the energy produced during the fission process?

WS 8.8 WS 8.9

BINDING ENERGY

The total mass of a stable nucleus is slightly less than the combined mass of the individual protons and neutrons. Einstein realised that this mass defect is linked to energy using $E = mc^2$.

- During nuclear fusion when a nucleus forms, mass is lost and converted into energy.
- During nuclear fission when an unstable nucleus splits into fragments, mass is lost and converted into energy.
- The energy released per nucleon is greater during fusion than during fission.

The mass defect in nuclei can be converted into energy using $E = mc^2$. This is known as the **binding energy** of the nucleus. The binding energy of the nucleus indicates how much energy is needed to separate the nucleus into individual protons and neutrons.

Each nucleus has its own binding-energy value, with a higher value indicating a more stable nucleus. A binding-energy-per-nucleon graph, as shown in Figure 16.4.3, allows a comparison of nuclear stabilities.

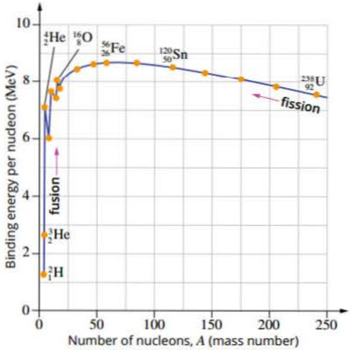


FIGURE 16.4.3 The graph of binding energy per nucleon shows that elements with mass numbers between 40 and 80 are the most stable.

Figure 16.4.3 can be analysed to better understand fission and fusion.

- Small nuclei have lower binding energy per nucleon values, indicating that they are easier to break apart compared to larger nuclei. Helium-4 has a relatively high value, indicating that it is stable.
- The binding energy per nucleon increases dramatically for the very small nuclei. As they fuse together, the binding energy per nucleon increases. This is the energy released during fusion.
- Elements with mass numbers between 40 and 80 have nuclei that are tightly bound. It takes more energy to break these nuclei apart. These are the most stable nuclei. These elements have the highest binding-energy-per-nucleon values on the graph.
- Larger nuclei have lower binding energy per nucleon values, indicating relatively lower stabilities.
- If a large nucleus such as uranium splits into two fragments, the binding energy per nucleon of the fragments increases. This is the energy released during fission.
- Iron (Fe), with a mass number of 56, has the most stable nucleus.
- Nuclei smaller than iron undergo fusion and release energy. Nuclei larger than iron undergo fission and release energy.

During nuclear reactions (e.g. nuclear fusion and nuclear fission), the principle of conservation of energy applies. When atoms lose some of their original mass, due to the conservation of energy that lost mass appears as generated energy, as noted by Einstein's equation $E = mc^2$.

16.4 Review

SUMMARY

- When fission occurs, the mass of the fission fragments is always less than the mass of the original particles. This decrease in mass is proportional to energy, as given by $E = mc^2$.
- A small amount of mass is lost during a fusion reaction. This mass is related to the energy produced according to $E = mc^2$.
- The energy of subatomic particles is measured in electronvolts (eV).
- Binding energy indicates how much energy is needed to separate the nucleus into individual protons and neutrons.
- The binding energy per nucleon increases dramatically for the very small nuclei. As nucleons fuse together, the binding energy per nucleon increases. This is the energy released during fusion.
- If a large nucleus such as uranium splits into two fragments, the binding energy per nucleon increases. This is the energy released during fission.

KEY QUESTIONS

1 The mass defect during the following process is 2.12×10^{-28} kg. Give your answers correct to three significant figures.

$$^{1}_{0}n + ^{235}_{92}U \rightarrow ^{148}_{57}La + ^{85}_{35}Br + a^{1}_{0}n$$

a Calculate the energy (in joules) released in this fission process.

b Express this energy in electronvolts.

2 The fusion reaction that is most promising for use in nuclear fusion reactors is:

$$^{2}_{1}H + ^{3}_{1}H \rightarrow ^{4}_{2}He + ^{1}_{0}n$$

Why is energy released during this process?

3 Qualitatively describe how the amount of energy released per nucleon during a single nuclear fission reaction compares to the amount of energy released per nucleon for a single fusion reaction.

4 During the process of nuclear fusion, mass is lost and this appears as energy. What is the approximate percentage of mass lost during a typical fusion reaction?

5 Consider the fusion reaction shown below.

$$^{2}_{1}H + ^{3}_{1}H \rightarrow ^{a}_{b}X + ^{1}_{0}n$$

a Determine the values of *a* and *b* and hence the symbol of the unknown element X.

b 33 MeV of energy is released. What is the mass defect in kg? Give your answer to two significant figures and in scientific notation.

6 The following fusion reaction is taking place in the Sun.

$$^{3}_{2}He + X \rightarrow ^{4}_{2}He + ^{1}_{1}p + ^{1}_{1}p$$

During each fusion reaction, 23 MeV of energy is released.

a What are the atomic and mass numbers of particle X and what is its symbol?

b Convert 23 MeV into joules. Give your answer to two significant figures and in scientific notation.

c Determine the mass defect for this fusion process. Give your answer to two significant figures and in scientific notation.

7 What happens to the binding energy and the stability of two hydrogen-2 nuclei when they are fused together to form helium-4?

Chapter review

KEY TERMS

activity
alpha particle
antineutrino
atomic number
beta particle
binding energy
daughter nucleus
decay series
electrostatic force
fission
fission fragment
gamma ray
Geiger counter
half life
isotope
mass number
neutral
nuclear transmutation
nucleon
nuclide
parent nucleus
penetrating ability
positron
radiation
radioactive
radioisotope
strong nuclear force

REVIEW QUESTIONS

1 How many protons and neutrons are in the ${}^{45}_{20}Ca$ nuclide?

2 Give the meaning of the term 'fissile'.

3 Use the periodic table in Figure 16.1.7 on page 419 to determine the number of protons, neutrons and nucleons in cobalt-60.

4 Determine the nature of the unknown, X, for the following transmutation:

$$ {}^{60m}_{27}Co \rightarrow {}^{60}_{27}Co + X $$

(60m means the nuclide is metastable and has a higher level of stability than very short-lived isotopes. The mass number is still 60.)

5 What type of radiation does potassium-48 (atomic number 19) emit? Use Figure 16.1.13 on page 423 to answer this question.

6 Are all atoms fissile? Give examples to support your answer.

7 Consider one particular proton in the nucleus of a zinc atom. Describe the forces that the proton experiences from other nucleons.

8 Identify each of these radiation types:

a ${}^{0}_{-1}A$

b ${}^{1}_{1}B$

c ${}^{4}_{2}C$

d ${}^{1}_{0}D$

e ${}^{0}_{0}E$

f ${}^{0}_{1}F$

9 Some nuclei can be made unstable by firing neutrons into them. The neutron is captured and the nucleus becomes unstable. The nuclear equation when the stable isotope boron-10 transmutates by neutron capture into a different element, X, by emitting alpha particles is:

$$ {}^{10}_{5}B + {}^{1}_{0}n \rightarrow X + {}^{4}_{2}He $$

Identify the unknown element, X, and its mass and atomic numbers.

10 Identify each of the unknown particles X and Y in the following nuclear transmutations.

a ${}^{14}_{7}N + \alpha \rightarrow {}^{17}_{8}O + X$

b ${}^{27}_{13}Al + Y \rightarrow {}^{27}_{12}Mg + {}^{1}_{1}H$

11 Both neutrons and alpha particles can be used to trigger nuclear fission. Explain why neutrons are better than alpha particles for inducing fission.

12 A typical fusion reaction is ${}^{2}_{1}H + {}^{2}_{1}H \rightarrow {}^{3}_{1}H + {}^{1}_{1}H$. Why are high temperatures such as 100 million degrees needed for this reaction to occur?

13 Find the values of *x* and *y* in each of these radioactive decay equations.

a ${}^{208}_{81}Ti \rightarrow {}^{x}_{y}Pb + \beta^{-}$

b ${}^{180}_{80}Hg \rightarrow {}^{x}_{y}Pt + \alpha$

14 Fluorine-18 is a radioisotope that is used for detecting tumours. It is formed when radioactive neon-18 decays by positron emission. Fluorine-18 in turn also decays by positron emission. The equations are as follows:

$$ {}^{18}_{10}Ne \rightarrow {}^{a}_{b}X + {}^{0}_{+1}\beta $$

$$ {}^{a}_{b}X \rightarrow {}^{c}_{d}Y + {}^{0}_{+1}\beta $$

Determine the values of *a*, *b*, *c*, *d* and identify X and Y, which are the daughter nuclei that result from this process.

15 The radioisotope nitrogen-12 decays by emitting a positron and a neutrino. The decay equation for nitrogen-12 is:

$$ {}^{12}_{7}N \rightarrow X + {}^{0}_{+1}\beta + {}^{0}_{0}\nu $$

Identify particle X.

16 A stable isotope of neon has 10 protons and 10 neutrons in each nucleus. Every proton is repelling all the other protons. Why is the nucleus stable?

17 Which type of radiation out of alpha, beta and gamma:

a is the fastest

b has the greatest penetrating power?

18 Health workers who deal with radiation to treat cancer often have to wear a lead vest to protect their vital organs from exposure to radiation. Which type(s) of radiation is the lead apron shielding them from?

19 A nuclear physicist was bombarding a sample of beryllium-7 with a beam of electrons in an effort to smash the electrons into the beryllium nuclei. Why would it be quite difficult for a collision between the electrons and the nuclei to occur?

20 A radioactive isotope X has a half-life of 20 minutes. A sample starts with 6.0×10^{14} atoms of the isotope. What amount of the original isotope will remain after 20 minutes?

21 Protactinium-234 is a radioactive element with a half-life of 70 s. If a sample of this radioisotope contains 6.0×10^{10} nuclei, how many nuclei of this element will remain after 140 s?

22 Radiotherapy treatment of brain tumours involves irradiating the target area with radiation from an external source. Why is cobalt-60 (a gamma emitter with a half-life of 5.3 years) generally used as the radiation source for this treatment?

23 Consider this fission reaction of uranium-235:

$$ {}^{1}_{0}n + {}^{235}_{92}U \rightarrow {}^{141}_{55}Cs + {}^{93}_{37}Rb + 2{}^{1}_{0}n $$

During this fission reaction there is a mass defect of 4.99×10^{-28} kg. How much energy in joules is produced per reaction?

24 Consider this fusion reaction:

$$ {}^{1}_{1}H + {}^{3}_{2}He \rightarrow {}^{4}_{2}He + {}^{0}_{+1}e + \nu $$

Hydrogen and helium-3 are being fused together and a helium-4 nucleus is being created along with a positron and a neutrino. 21 MeV of energy is released.

a How does the combined mass of the hydrogen and helium-3 nucleus compare with the combined mass of the helium-4 nucleus, positron and neutrino?

b Where has the energy that was released come from?

c Convert the energy into joules.

d What is the mass defect of this fusion reaction?

25 Compare the waste and the energy per nucleon produced by fusion reactors with those of fission reactors.

26 What happens to the binding energy per nucleon and the stability of the nucleus when a uranium-238 nucleus splits apart to form two smaller nuclei?

27 The binding energy per nucleon for iron (mass number 56) is higher than for other elements. What does this mean for the stability of iron nuclei?

28 X-ray light produced by a synchrotron has energy of 5.0×10^{4} eV. What is the energy of this light in joules?

29 After completing the activity on page 416, reflect on the inquiry question: How can the energy of the atomic nucleus be harnessed?

CHAPTER 17

Deep inside the atom

Physics is the story of how we know about the physical universe around us. It is a fascinating story that started with people like Aristotle over two thousand years ago, and is far from over. This chapter brings together the huge and the tiny. The largest machine ever built by man, the Large Hadron Collider (LHC), is designed to detect the smallest particles in the universe. It does this by giving two beams of protons travelling in opposite directions huge amounts of energy and then colliding them. The energy density of these collisions is so high it enables physicists to re-create the conditions that existed right back to around a millionth of a millionth (10^{-12}) of a second after the big bang and hence to see what sort of particles made up the very early universe—and how they evolved to become our current universe.

Content

INQUIRY QUESTION

How is it known that human understanding of matter is still incomplete?

By the end of this chapter you will be able to:

- analyse the evidence that suggests:
  - that protons and neutrons are not fundamental particles
  - the existence of subatomic particles other than protons, neutrons and electrons
- investigate the Standard Model of matter, including:
  - quarks, and the quark composition hadrons
  - leptons
  - fundamental forces (ACSPH141, ACSPH142) ICT
- investigate the operation and role of particle accelerators in obtaining evidence that tests and/or validates aspects of theories, including the Standard Model of matter (ACSPH120, ACSPH121, ACSPH122, ACSPH146). ICT

17.1 The Standard Model

PHYSICS INQUIRY CCT

What's in the box?

How is it known that human understanding of matter is still incomplete?

COLLECT THIS...

- five identical opaque boxes with lids that seal or can be taped shut
- 10 different objects, chosen by someone else, that fit within the boxes. Ensure they are varied in size, shape and material.
- four known test objects of different shape and material
- detecting apparatus such as magnets, scales and a metal detector

DO THIS...

1 Get someone else to place the objects in four of the boxes and seal them so you cannot see what is in them. Some of the boxes should have one item in them, and some more than one.

2 Make initial observations about the contents of one of the boxes. Record your observations.

3 Plan a test that could help you identify how many and what kind of objects are in the sealed boxes using the one empty box, the test objects and detecting apparatus.

4 Conduct the test on the four sealed boxes. Record the evidence obtained.

5 Make sure your total number of predicted objects totals 10. Decide if one or more of the objects are from the group of known objects.

RECORD THIS...

Describe the process used to infer properties of hidden objects.

Present a table of your results.

REFLECT ON THIS...

How is it known that human understanding of matter is still incomplete?

What property is the easiest to measure and most reliable to interpret?

What type of evidence is the most difficult to interpret?

How would you convince your peers of your theory about the box contents?

As a spin-off from the investigations into nuclear physics that were conducted towards the end of World War II, particle physicists began to predict and discover new subatomic particles. Subatomic particles are those smaller than an atom. Physicists recognised that as the energy of bombarding particles was increased, new particles were being formed.

At first, physicists built particle detectors and waited for high-energy cosmic rays from space to smash into their targets in order to see what nuclear fragments (pieces) they could identify. However, physicists recognised that in order to probe more effectively into the nucleus, they would need to build more powerful machines such as particle accelerators. Particle accelerators accelerate protons and electrons to energies high enough to form the particles that the physicists were predicting would exist. Figure 17.1.1 on page 449 shows the motion of particles detected in a particle accelerator.

These new discoveries led to the development of the **Standard Model** of particle physics. The Standard Model is the best description so far of the fundamental building blocks of matter and the forces that govern them. This section will explore the particles of the Standard Model.

THE STANDARD MODEL OF PARTICLE PHYSICS

Currently, the Standard Model of particle physics is the most successful theory for predicting the behaviour and properties of the particles that exist in nature. The Standard Model was developed by experiment and theory together. As technology improved, new particles were found, and as the model developed, more particles were predicted and were then found. The Standard Model is a mathematical description of all known particles and three of the forces acting between them.

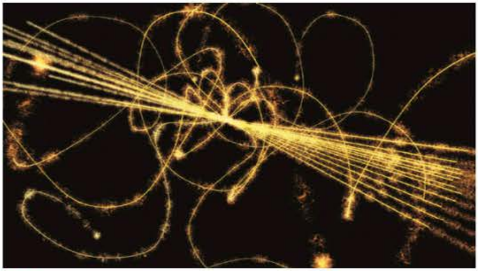


FIGURE 17.1.1 This image shows the paths of many particles in a particle accelerator. Energy is converted into mass as pairs of particles and their antiparticles are created from a gamma ray. The big bang theory and the Standard Model of particle physics tell us that this is how the matter inside us was created in the first seconds of the universe billions of years ago.

The essential components of the Standard Model are that forces between particles are mediated (caused) by other particles and that all matter is made of 12 fundamental particles.

The Standard Model consists of two types of particles: **fermions**, which make up all matter, and **gauge bosons**, which mediate the various interactions between other particles.

PHYSICS IN ACTION CCT ICT

Quarks and other subatomic particles

Our understanding of the atom has changed greatly in the past 100 years. It was once thought that atoms were like miniature billiard balls: solid and indivisible. That idea was changed forever when the first subatomic particles—the electron, the proton and then the neutron—were discovered in the period from 1897 to 1932.

Since World War II, further research has uncovered about 300 other subatomic particles! Examples of these include pi-mesons, mu-mesons, kaons, tau leptons and neutrinos. For many years, physicists found it difficult to make sense of this array of subatomic particles.

Then in 1964 Murray Gell-Mann put forward a simple theory. He suggested that most subatomic particles were themselves composed of a number of more fundamental particles called **quarks**. Currently, it is accepted that there are six different quarks, each with different properties (and strange names!): up, down, charmed, strange, top and bottom. The latest quark to be identified was the top quark, whose existence was confirmed in 1995. The proton consists of two up quarks and one down quark, while neutrons consist of one up quark and two down quarks. Subatomic particles that consist of quarks are known as **hadrons**. The family of particles called the leptons has six members:

FIGURE 17.1.2 The Large Hadron Collider particle accelerator is at CERN, the European centre for high-energy physics. It accelerates protons from rest to 99.99995% of the speed of light in under 20 s.

electron, electron-neutrino, muon, muon neutrino, tau and tau neutrino. **Leptons** are indivisible point particles; they are not composed of quarks.

A significant amount of effort and money has been directed to testing Gell-Mann's theory—both theoretically and experimentally. This has involved the construction of larger and larger particle accelerators such as the ones at Fermilab in Chicago and CERN in Geneva (Figure 17.1.2).

While the current theory suggests that quarks and leptons are the ultimate fundamental particles that cannot be further divided, the nature of scientific theories and models is such that they can change as new experimental data are obtained. Are quarks and leptons made of smaller particles again? Time will tell!

PHYSICSFILE

The graviton

Theoretical physics predicts that there should be a boson for gravity too, called a graviton. However, this has not yet been found. Luckily the effects of gravity—although familiar to us on a human scale—are negligible at the subatomic scale, meaning the Standard Model works accurately without this missing piece.

GAUGE BOSONS

There are four forces that can act between particles and govern their behaviour:

- strong nuclear
- electromagnetic
- weak nuclear
- gravitational.

Each has a different strength and acts over different distances. For example, the **strong nuclear force** acts over very short distances (that is, on the subatomic scale), but is very large. Gravity acts over an infinite range, but is relatively weak.

The fundamental assumption in the Standard Model is that three of these forces (strong nuclear, electromagnetic and weak nuclear) arise through the exchange of particles called gauge bosons (or just bosons). Bosons are often called force-carrying, force-mediating or exchange particles.

Each force has its own boson. In this context, consider these forces as their particle equivalents, according to the notion of wave–particle duality. In other words, consider these forces as particles.

Forces through exchange of particles

Previously, forces were thought of as being exerted on particles by fields. For example, there would be a region around a charged particle where another charged particle experiences a force. This may seem quite puzzling, as the force is applied without any direct interaction by the two particles. The same effect is felt when two magnets are brought together. The magnets do not need to touch in order to feel the force between them. In the Standard Model this is resolved and forces are thought to be exerted through the exchange of other particles.

To use an analogy to see how a force can be exerted on two particles through the exchange of another particle, consider Figure 17.1.3. Two inline skaters stand stationary and then begin to pass footballs back and forth to each other. As they do this, they will begin to move away from each other. This is due to the conservation of momentum each time they throw and catch the ball. This situation could be likened to two particles experiencing a repulsive (pushing away) force.

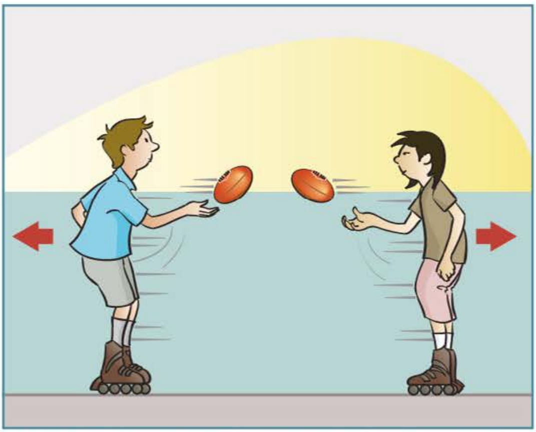


FIGURE 17.1.3 Inline skaters exchanging footballs act as an analogy for particles experiencing repulsive forces due to the exchange of gauge bosons.

A force of attraction can also be illustrated using the same analogy as shown in Figure 17.1.4. If the two skaters now exchange the footballs by trying to grab them out of each other's hands, they will exert a force of attraction on each other. This would cause them to move together and can be likened to two particles experiencing an attractive force.

Going back to a particle level, Figure 17.1.5 represents an electron approaching another electron. Each electron emits a photon that is absorbed by the other. This causes each electron to experience a force of repulsion. The photon has been responsible for the electromagnetic force acting on the two electrons. Table 17.1.1 gives a summary of the nature of these particles, their strength and the range over which they can exert a force.

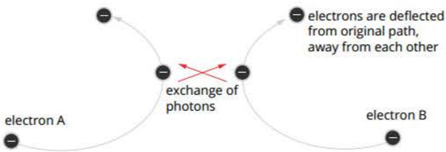


FIGURE 17.1.5 Two electrons approach each other, are repelled and then move away from each other. The two electrons exchange a photon that is the carrier of the electromagnetic force.

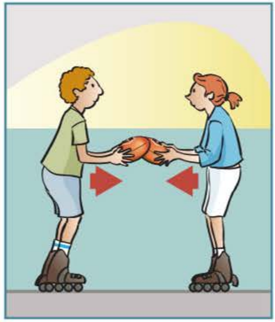


FIGURE 17.1.4 Inline skaters trying to grab footballs act as an analogy for particles experiencing attractive forces due to the exchange of gauge bosons.

TABLE 17.1.1 Features of the various gauge bosons and the forces they are responsible for.

Force	Nature	Relative strength	Range (m)	Force carrier (gauge boson)
strong nuclear	bonds nucleons together; acts between quarks	1	10^{-15} (~size of nucleus)	gluon
electromagnetic	responsible for both electric and magnetic fields exerting forces of attraction or repulsion	$\frac{1}{137}$	infinite	photon
weak nuclear	causes radioactive decay	10^{-6}	10^{-18} (less than the width of a proton)	W^+, W^- and Z
gravity	a force of attraction that acts between any two objects with mass	6×10^{-39}	infinite	graviton (theoretical)

FERMIONS

The Standard Model states that all particles of matter are made of one or more of the 12 fundamental or elementary particles and their **antimatter** opposites. Being a fundamental particle means that, to the best of scientific knowledge, it is not composed of other smaller particles. These fundamental particles are called fermions, which are defined by their **quantum numbers**.

Fermions are divided into two groups of particles called quarks and leptons. Six of the 12 particles are quarks and combine in groups of two or three to form the hundreds of known particles.

Quarks

All matter is made of particles comprising protons, neutrons and electrons. However, protons and neutrons are actually made of even smaller particles called quarks. Quarks are the only things that interact using the strong nuclear force.

High-energy particle physicists seem to be a creative bunch: not only do they choose unusual names for their new particles but they also choose unusual ways to describe them and their properties. The six different types of quarks are referred to as the six 'flavours' of quarks.

FIGURE 17.1.6 This is an artist's representation of the three quarks that make up a proton (uud).

PHYSICSFILE L

Naming quarks

The name 'quark' was taken by Murray Gell-Mann from the book *Finnegans Wake*, by James Joyce: 'three quarks for Muster Mark'.

The six flavours of quarks are:

- up (u)
- down (d)
- strange (s)
- charmed (c)
- bottom (b)
- top (t).

The last four quark names also apply to new quantum numbers and their conservation laws, called strangeness (s), charm (c), bottomness (b) and topness (t). Quarks have their antimatter opposites called antiquarks that have the opposite sign for all of their quantum numbers. Quarks have the properties of **baryon number**, spin and charge, which must add together to give the total baryon number, spin and charge of the particles that they combine to make (Figure 17.1.6).

Table 17.1.2 shows the quantum numbers for quarks. (In this table, e is the fundamental or elementary charge. Its magnitude is equal to the charge on a proton, i.e. $e = q_p = -q_{e^-}$.)

All quarks experience the strong nuclear force, and this separates them from the leptons. Leptons do not experience the strong nuclear force. Quarks also have non-integer electromagnetic charges, meaning their electromagnetic charges are all fractions. Leptons have electromagnetic charges of −1 or 0.

When quarks bond together they form hadrons. Note that quarks have only been found bound into hadrons; they have never been found singly.

Hadrons

The common feature of hadrons is that they all interact by exchanging **gluons**, which are the particles that mediate the strong nuclear force. Hadrons that carry an electromagnetic charge can also interact by exchanging photons, but the effect of the gluons far outweighs any other force-mediating particle.

This group of particles is further categorised into two groups, the **baryons** and the **mesons**, based on a property called baryon number (B), described below.

TABLE 17.1.2 The quantum numbers for quarks.

Particle/ antiparticle name	Symbol	Electromagnetic charge (Q)	Baryon number (B)	Strangeness (S)	Charm (c)	Bottomness (b)	Topness (t)
up	u	$+\frac{2}{3}e$	$\frac{1}{3}$	0	0	0	0
down	d	$-\frac{1}{3}e$	$\frac{1}{3}$	0	0	0	0
strange	s	$-\frac{1}{3}e$	$\frac{1}{3}$	−1	0	0	0
charmed	c	$+\frac{2}{3}e$	$\frac{1}{3}$	0	+1	0	0
bottom	b	$-\frac{1}{3}e$	$\frac{1}{3}$	0	0	−1	0
top	t	$+\frac{2}{3}e$	$\frac{1}{3}$	0	0	0	+1
anti-up	$\bar{u}$	$-\frac{2}{3}e$	$-\frac{1}{3}$	0	0	0	0
anti-down	$\bar{d}$	$+\frac{1}{3}e$	$-\frac{1}{3}$	0	0	0	0
anti-strange	$\bar{s}$	$+\frac{1}{3}e$	$-\frac{1}{3}$	+1	0	0	0
anti-charmed	$\bar{c}$	$-\frac{2}{3}e$	$-\frac{1}{3}$	0	−1	0	0
anti-bottom	$\bar{b}$	$+\frac{1}{3}e$	$-\frac{1}{3}$	0	0	+1	0
anti-top	$\bar{t}$	$-\frac{2}{3}e$	$-\frac{1}{3}$	0	0	0	−1

Mesons are particles that have a baryon number of zero. In this group are many particles and their antimatter particles; for example, the pion (π^+), antipion (π^-) and pi-zero (π^0), the kaon (K^+) and antikaon (K^-), and the eta (η^0). Note that the antimatter particle of the eta is considered to be itself.

Mesons consist of a quark and an antiquark pair. For example the pion-plus (π^+) consists of an up quark and an anti-down quark ($u\bar{d}$). Therefore the electromagnetic charge of the pion-plus is $\left(+\frac{2}{3}\right)+\left(+\frac{1}{3}\right)=\left(+\frac{3}{3}\right)=+1$, but its baryon number is $\left(+\frac{1}{3}\right)+\left(-\frac{1}{3}\right)=\left(\frac{0}{3}\right)=0$, as it should be for all particles that are not baryons. Table 17.1.3 shows some of the many mesons and their constituent quarks.

TABLE 17.1.3 Various mesons and their constituent quarks.

Meson name	Symbol	Baryon number (*B*)	Strangeness (*S*)	Charm (*c*)	Bottomness (*b*)	Topness (*t*)	Quarks
pion-plus	π^+	0	0	0	0	0	$u\bar{d}$
pion-minus	π^-	0	0	0	0	0	$\bar{u}d$
kaon-plus	K^+	0	+1	0	0	0	$u\bar{s}$
kaon-minus	K^-	0	−1	0	0	0	$\bar{u}d$
rho-plus	ρ^+	0	+1	0	0	0	$u\bar{d}$
rho-minus	ρ^-	0	−1	0	0	0	$\bar{u}d$
phi	φ	0	0	0	0	0	$\bar{s}s$
D-plus	D^+	0	0	+1	0	0	$c\bar{d}$
D-zero	D^0	0	0	+1	0	0	$c\bar{u}$
D-plus-s	D_s^+	0	+1	+1	0	0	$c\bar{s}$
B-minus	B^-	0	0	0	−1	0	$b\bar{u}$
upsilon	Y	0	0	0	0	0	$b\bar{b}$

Baryons are particles that have a baryon number of 1 for normal matter or −1 for antimatter. In this group are the familiar proton (p^+) and antiproton ($\bar{p}$), neutron (n) and antineutron ($\bar{n}$), along with hundreds of other particles and their antiparticles; for example, the lambda-zero (Λ^0) and antilambda-zero ($\bar{\Lambda}^0$), sigma-plus (Σ^+), sigma-zero (Σ^0) and sigma-minus (Σ^-), and the xi-zero (Ξ^0) and omega-minus (Ω^-).

Figure 17.1.7 shows the different groups of elementary particles covered so far and how the groups overlap.

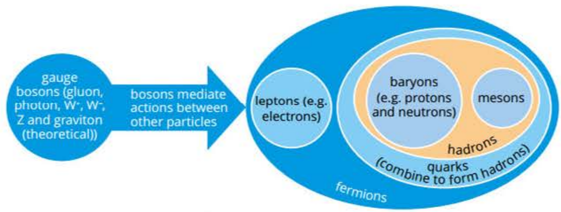


FIGURE 17.1.7 The different groups of elementary particles

Baryons, including the proton and neutron, consist of three quarks. For example, the proton is made up of two up quarks and a down quark (uud), and the neutron is made up of one up quark and two down quarks (ddu), as shown in Figure 17.1.8.

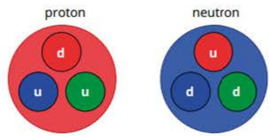


FIGURE 17.1.8 The proton and neutron are both baryons which are made of three quarks.

Each of these quarks must have a different **colour charge** of red, green or blue. They must always bond together to give a total colour charge of 'white', which means for baryons:

- red + blue + green (as with the proton and neutron shown above)
- antired + antiblue + antigreen

and for mesons:

- red + antired
- blue + antiblue
- green + antigreen.

This is why there are never any 'free' individual quarks by themselves.

They must also combine to have the total electromagnetic charge of the proton, which is one 'fundamental unit' (e or $+1.602 \times 10^{-19}$ C). Therefore quarks must have an electromagnetic charge that is less than the fundamental unit. In fact, quarks have an electromagnetic charge of either $+\frac{2}{3}e$ (up, charmed and top) or $-\frac{1}{3}e$ (down, strange and bottom), where e is the former fundamental unit of electromagnetic charge.

The proton's electromagnetic charge is therefore made up of:

$$\left(+\tfrac{2}{3}\right)+\left(+\tfrac{2}{3}\right)+\left(-\tfrac{1}{3}\right)=\left(+\tfrac{3}{3}\right)=+1$$

A neutron is made up of two down quarks and an up quark (ddu) of different colours, which equates to an electromagnetic charge of:

$$\left(-\tfrac{1}{3}\right)+\left(-\tfrac{1}{3}\right)+\left(+\tfrac{2}{3}\right)=\left(+\tfrac{0}{3}\right)=0$$

Note that the baryon number of each quark is $+\frac{1}{3}$ so three quarks add to give a baryon number of +1. Antiquarks have a baryon number of $-\frac{1}{3}$, so three antiquarks add to give a baryon number of −1. Adding the masses of the separate quarks in a proton is less than the mass of the proton. This is because most of the mass of the proton is stored in the gluons that hold the quarks together. Table 17.1.4 shows some of the many baryons and their constituent quarks.

TABLE 17.1.4 Some of the many baryons and their constituent quarks.

Baryon name	Symbol	Baryon number (*B*)	Strangeness (*S*)	Charm (*c*)	Bottomness (*b*)	Topness (*t*)	Quarks
proton	p	+1	0	0	0	0	uud
antiproton	$\overline{p}$	−1	0	0	0	0	$\overline{u}\overline{u}\overline{d}$
neutron	n	+1	0	0	0	0	udd
antineutron	$\overline{n}$	−1	0	0	0	0	$\overline{u}\overline{d}\overline{d}$
Lambda-plus	Λ^+	+1	0	+1	0	0	udc
Lambda-zero	Λ^0	+1	−1	0	0	0	uds
Sigma-plus	Σ^+	+1	−1	0	0	0	uus
Sigma-zero	Σ^0	+1	−1	0	0	0	uds
Sigma-minus	Σ^-	+1	−1	0	0	0	dds
Xi-zero	Ξ^0	+1	−2	0	0	0	uss
Xi-plus	Ξ^+	+1	−2	0	0	0	dss
Omega-minus	Ω^-	+1	−3	0	0	0	sss

MATTER AND ANTIMATTER

As mentioned with the particles discussed so far, most particles have what is called an antiparticle. These have the same properties of mass, spin, charge and life span as the particle. However, their electric charge and their quantum numbers are the same in magnitude but have the opposite sign. For example, a positron has the same mass as an electron, but has a positive charge.

There are two conventions used to indicate an antiparticle.

- The antiparticle of an uncharged particle is indicated by placing a bar above the symbol for the normal matter particle. An example is an electron **neutrino**. This particle has the symbol ν_e and its antiparticle, the anti-electron neutrino, has the symbol $\bar{\nu}_e$.
- The antiparticle of charged particles is given the symbol of the particle but with the opposite sign; for example, the antiparticle of the muon (μ^-) is the antimuon (μ^+).

Leptons

The particles in this group are the ones that interact by exchanging W and Z bosons, which mediate the **weak nuclear force**. Leptons that carry a charge can also interact by exchanging photons, which mediate the electromagnetic force. Leptons do not interact via the strong nuclear force carriers.

Electrons are a type of lepton. These particles also include the muon and tau particles, as well as their corresponding neutrinos and the antimatter opposites of these six particles. In your studies of nuclear physics, you will have seen that an electron, ejected from the nucleus in a beta-minus decay, will always be emitted along with an antineutrino. In a beta-plus decay, an antimatter positron is emitted with a normal neutrino.

The LEP (Large Electron–Positron collider) at CERN was the most powerful collider of leptons ever built. The ALEPH (Apparatus for LEP Physics) experiment determined the mass of the W and Z bosons and found that the number of particles with light neutrinos was three (Figure 17.1.9).

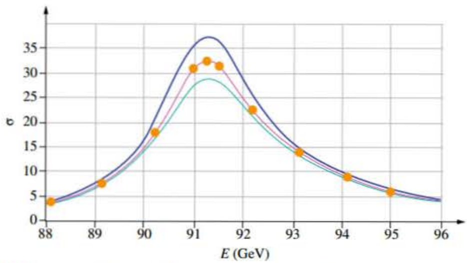


FIGURE 17.1.9 These three lines model the existence of two (green), three (red) or four (blue) neutrinos. The data from millions of Z particles, observed by the ALEPH experiment at CERN, produced the orange dots. There is a strong correlation between the data and the three-neutrino model.

When nuclei undergo beta-minus decay (β^-), an electron (e^-) is emitted from the nucleus, along with its corresponding electron antineutrino ($\bar{\nu}_e$). In a beta-plus decay (β^+), a positron (e^+) is emitted along with the electron neutrino (ν_e). Neutrinos are essentially massless particles that interact so weakly with matter that they can go through the entire Earth like photons (light) go through a pane of glass. Since neutrinos have zero charge they do not experience the electromagnetic force.

Another lepton particle called the muon (μ^-) can also be emitted from the nucleus along with the muon antineutrino ($\bar{\nu}_\mu$) in a similar way to the beta particle. Muons are very similar to electrons, but are 207 times larger. There is also an antimuon (μ^+) along with its associated muon neutrino (ν_μ). These muon neutrinos are found to be different to the electron neutrinos.

The last leptons are the tau (τ^-) with its corresponding antitau neutrino ($\bar{\nu}_\tau$), and the antitau (τ^+) with its tau neutrino (ν_τ). The tau is extremely massive in comparison to the electron, at 3477 times its mass. Table 17.1.5 gives a summary of leptons and their properties. Table 17.1.6 shows the six antileptons with their symbols, charges and conservation numbers.

TABLE 17.1.5 The properties of leptons.

Lepton name	Symbol	Charge	Electron lepton number	Muon lepton number	Tau lepton number
electron	e^-	−1	−1	0	0
electron neutrino	ν_e	0	+1	0	0
muon	μ^-	−1	0	+1	0
muon neutrino	ν_μ	0	0	+1	0
tau	τ^-	−1	0	0	+1
tau neutrino	ν_τ	0	0	0	+1

TABLE 17.1.6 The properties of antileptons.

Antiparticle name	Symbol	Charge	Electron lepton number	Muon lepton number	Tau lepton number
positron	e^+	+1	−1	0	0
electron antineutrino	$\bar{\nu}_e$	0	−1	0	0
antimuon	μ^+	+1	0	−1	0
muon antineutrino	$\bar{\nu}_\mu$	0	0	−1	0
antitau	τ^+	+1	0	0	−1
tau antineutrino	$\bar{\nu}_\tau$	0	0	0	−1

FIGURE 17.1.10 This image records a proton and antiproton annihilation. This event was recorded in 1955 in a photographic emulsion at the Bevatron accelerator at the Lawrence Berkeley Laboratory, California.

Matter vs antimatter: Annihilation

When physicist Paul Dirac first proposed antimatter, he did more than just predict its existence. He also predicted that matter and antimatter will **annihilate** when they collide. This is observed and exploited on a daily basis in experiments and applications involving antimatter.

Figure 17.1.10 shows an antiproton entering along the track marked L (top), before colliding with a proton. When the proton and antiproton mutually annihilate, the huge amount of energy released produces several new particles whose tracks form the 'star' pattern seen in this image.

Another example is the annihilation of an electron and positron. This produces two photons that carry away the energy initially contained in the electron and positron. These events are a stunning verification of Einstein's famous equation $E = mc^2$, where mass (m) and energy (E) are equivalent and can be converted from one form to the other.

The opposite of annihilation is also observed. This is called particle–antiparticle pair production. Energy in the form of a photon can create a particle–antiparticle pair if the photon has energy greater than or equal to the mass of the particle–antiparticle pair. Pair production also illustrates the relationship between mass and energy in Einstein's equation.

17.1 Review

SUMMARY

- Our understanding of the universe progresses by both theory and experimentation. New particles are sometimes predicted before any experimental evidence is available, while sometimes observations in experimental events lead to the discovery of a new particle.
- The Standard Model of particle physics explains three of the four fundamental forces in the universe (electromagnetism, the strong nuclear force and the weak nuclear force) in terms of an exchange of particles called gauge bosons.
- The gauge bosons for these three forces are the photon for electromagnetism, the gluon for the strong nuclear force, and the Z, W^- and W^+ for the weak nuclear force.
- All matter in the universe is made of fundamental particles called quarks and leptons.
- There are six quarks which all experience the strong nuclear force mediated by gluons. These combine to form hadrons and cannot exist alone. The hadrons include baryons made of three quarks and mesons made of two quarks (a quark–antiquark pair).
- Leptons are a group of particles that interact by exchanging W and Z bosons, which mediate the weak nuclear force. Electrons are in this group of particles.
- Antimatter particles have similar properties, like mass, spin and lifetime, as their corresponding particles. Their electric charge and other characteristics called quantum numbers are the same in magnitude but have an opposite sign.
- When a matter particle and its antimatter particle collide, they completely annihilate to produce a photon with the equivalent energy of the two particles, calculated by Einstein's mass–energy equation $E = mc^2$.

KEY QUESTIONS

1 Between which particles does the strong nuclear force act?

2 Which gauge bosons are responsible for the weak nuclear, strong nuclear and electromagnetic forces?

3 The particles of the Standard Model have been classified into three main groups. Name the three groups and give the main characteristic of each.

4 What are the fundamental differences between the two groups of particles within the fermions (quarks and leptons)?

5 A common analogy used to explain how forces are mediated by the exchange of particles involves two skaters passing balls to each other. What does the ball represent in this analogy?

6 Name the two groups of hadrons and explain the differences between the types of particle in each group.

7 Classify each of the following as a gauge boson, a lepton or a hadron: proton, gluon, electron, photon, muon, neutron, neutrino.

17.2 Evidence for the Standard Model

The Australian Nuclear Science and Technology Organisation—ANSTO—is the Australian Government's national nuclear organisation, headquartered in Lucas Heights (in the southern suburbs of Sydney). One of ANSTO's missions is to operate nuclear facilities to benefit the Australian community, industry and scientific research. These facilities include the Australian Centre for Neutron Scattering (the OPAL reactor) located in New South Wales, and the Australian Synchrotron, located in Melbourne. The Australian Synchrotron is the most powerful synchrotron in the Southern Hemisphere.

A **synchrotron** is a type of particle accelerator. Electrons with energies as high as 3 billion electronvolts are accelerated around a huge evacuated ring to almost the speed of light. These charges are forced to follow a curved path by a magnetic field. As they accelerate around curves, the electrons give off bursts of radiation. This synchrotron radiation is channelled down tubes called beamlines and utilised by researchers in a range of experimental stations.

Synchrotrons can be used as supermicroscopes to reveal the hidden structure of fibres, chemical proteins and enzymes by using powerful techniques such as X-ray diffraction. Synchrotrons can also improve medical imaging techniques and can distinguish features of cells up to 1000 times smaller than previously possible. X-ray lithography can be used to etch microscopic patterns on materials and construct micromachines.

This section discusses the discovery of new particles, and how particle accelerators play a key role in this.

PREDICTING THE EXISTENCE OF NEW PARTICLES

One of the most exciting aspects of physics is the discovery of something new. It could be a new theory or a more elegant solution to a problem. For particle physicists the ultimate achievement is to discover a new particle. These discoveries could unfold in one of two ways:

- the particle is predicted to exist through logical reasoning by theoretical physicists
- the particle is observed in a collision event by experimental physicists.

Whether by theory or observation, finding new particles generates great excitement and activity in the world of high-energy particle physics. Theoretical and experimental methods have been used over the years to discover, for example, the neutron, the neutrino, the positron and the Higgs boson. The discovery of the neutron and positron was discussed in greater detail in Chapter 14.

GO TO ➤ Section 14.2 page 385

The discovery of the neutrino

Physicists studying the energy of the beta particles ejected from radioisotopes discovered that they seemed to vary by a small amount. They were missing a quantity of energy, which contravened (contradicted) the conservation laws of energy, spin and momentum. Wolfgang Pauli recognised that as the laws must not be violated there must be another very small particle produced in beta decay that carried the missing momentum and energy. The name given to this particle was the neutrino.

Neutrinos are particles that have been found to possess the lowest mass in nature, and they exist all throughout the universe. Neutrinos have no charge and their mass has only recently been discovered to be about one-billionth that of a proton (about 10^{-36} kg). While you have been reading these sentences, billions of neutrinos have passed right through your body, and continued on to pass right through the Earth. Fortunately neutrinos interact with matter very rarely and so are completely harmless. It has been estimated that if neutrinos passed through a piece of lead 8 light-years thick, they would still have only a 50% chance of being absorbed.

The discovery of the Higgs boson

Until the 1960s, physicists did not know how particles got their mass. British physicist Peter Higgs, and others, proposed an answer to this question in 1964. Using the Standard Model, they predicted that there was a yet-undiscovered particle that would interact with other particles and give them what is measured as mass. This particle came to be known as the **Higgs boson**.

The search for this, and for other particles, led over 10 000 physicists and engineers to collaborate to build the **Large Hadron Collider** (LHC) near Geneva, Switzerland. This is the largest and most complex experimental facility ever built. One of its primary goals was to test the prediction of the existence of the Higgs boson.

The search continued until 2012, when evidence for the discovery of a candidate for the Higgs particle was first announced (Figure 17.2.1). In 2013 two of the original researchers, Peter Higgs and François Englert, were awarded the Nobel Prize in Physics for their prediction and their work to prove the prediction.

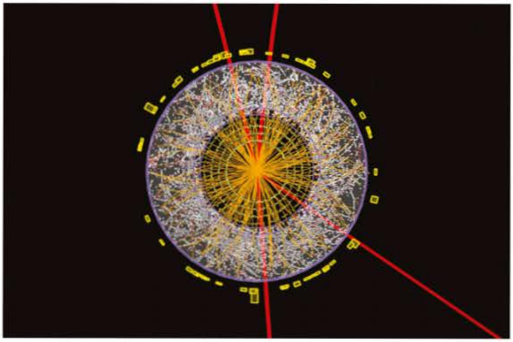


FIGURE 17.2.1 The image shows tracks of particles and measurements of their energies in the ATLAS detector at the LHC. The nature and energies of the particles produced are consistent with predictions of the formation of a Higgs boson.

SYNCHROTRONS

Synchrotron light was first discovered in the 1940s, when it was observed being produced in particle accelerators used for theoretical physics. When first discovered, this radiation was seen as an unwanted by-product of the acceleration process, as its release robbed accelerating particles of energy. It was only later that the useful benefits of such radiation became apparent. Since their origins in the 1940s, synchrotrons have undergone progressive evolution.

Originally, large particle accelerators were designed to investigate the nature of matter by examining the structure of atoms and molecules via collisions. Strong magnets directed the particles to collide with a target or with another moving particle. Scientists used the data gained from these collisions to learn more about the make-up of the subatomic particles fired from the machine or the target samples that were hit.

In contrast to these types of particle accelerators built for collisions, a synchrotron light source is designed to use electrons to generate beams of infrared, UV, visible and X-ray radiation. Synchrotron light, or synchrotron radiation, is the term given to a range of electromagnetic radiation of wavelengths from approximately 10^{-3} to 10^{-10} m. This electromagnetic radiation is produced by charged particles such as electrons or protons as they travel in a curved path at speeds close to that of light. The beam of the synchrotron light produced falls in the shape of a cone ahead of the travelling charged particles.

The Australian Synchrotron at Monash University, shown in Figure 17.2.2, is an example of this type of particle accelerator.

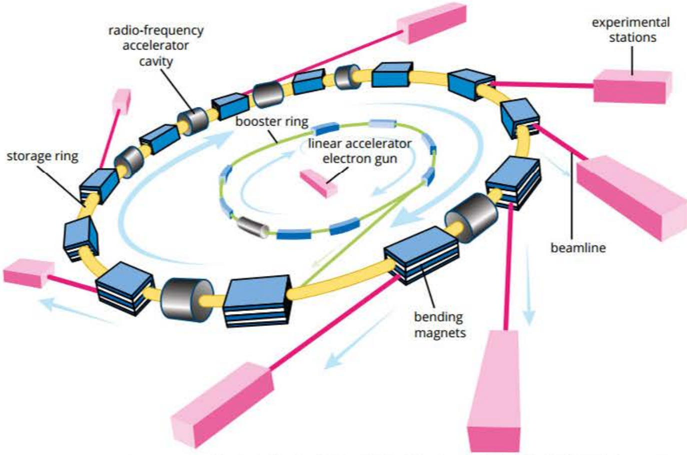


FIGURE 17.2.2 The Australian Synchrotron is about the size of a football field. This scale is necessary to contain the electrons, which are travelling at almost the speed of light as they zoom around the storage ring. This diagram shows the main features and experimental stations arranged around the storage ring.

The electron linac

GO TO ➤ Section 14.1 page 376

Cathode ray tubes (introduced in Chapter 14) are useful particle accelerators but are limited to using voltages over a few tens of kilovolts. A linear accelerator, or **linac**, accelerates particles in straight lines. The first linac was built in 1928 by the Norwegian engineer Rolf Widerøe. It consisted of three hollow metal tubes inside an evacuated cylinder. These are called drift tubes and were used in Widerøe's machine to accelerate potassium ions to an energy of 50 000 eV (50 keV). This type of accelerator is referred to as a standing-wave linac.

The Australian Synchrotron and other electron linacs make use of travelling waves rather than standing waves in order to accelerate particles. The travelling-wave linac consists of:

- an electron gun
- a vacuum system
- focusing elements
- RF (radio-frequency) cavities.

GO TO ➤ Section 12.2 page 309

As shown in Figure 17.2.3, electrons escape from the electron gun as they boil off the heated filament of the assembly. From here, they accelerate across a potential difference of about 100 kV and exit the electron gun at a velocity of approximately half the speed of light. At such velocities, the effects of relativity (introduced in Chapter 12) come into play.

The electron beam travels through an ultra-high vacuum within the linac, to prevent energy loss through interaction with air particles. As the electrons travel, focusing elements act on the beam to constrict it to a narrow beam in the centre of the vacuum tube.

Electrons are accelerated to close to the speed of light by the end of the linac. Such acceleration is critical to the production of synchrotron light in the storage ring.

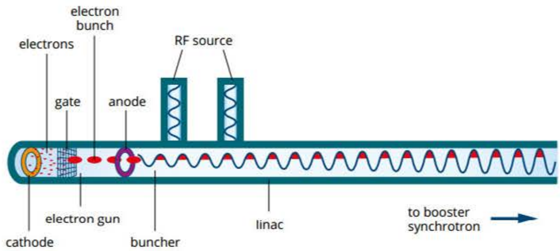


FIGURE 17.2.3 A diagram of a travelling-wave linac.

This huge acceleration is achieved by cylindrical RF (radio-frequency) cavities that surround the electron beam. These cavities produce intense electromagnetic radiation at several hundred megahertz (MHz). The RF radiation propagates through the linac as a travelling wave. When timed correctly, electrons can, in effect, 'ride the crest' of this RF wave, resulting in their acceleration to enormous speeds.

In the Australian Synchrotron, electrons are released from the gate of the electron gun in pulses every 2×10^{-9} s to travel towards the anode. These electrons accelerate as they pass through the RF cavity with the crest of the RF radiation and are slowed down when they pass through with the trough of the RF radiation. This effect causes the electrons to become bunched into groups as they travel through the linac itself. The frequency of RF radiation is timed to accelerate the arrival of each electron bunch. The linac used in the Australian Synchrotron gives the electrons a kinetic energy of 100 MeV.

The booster ring

Within the circular **booster ring** of the Australian Synchrotron, bending magnets provide a force at right angles to the motion of the electrons in order to bend them into a circular path. In this ring, the energy of the electrons is boosted from 100 to 3000 MeV, or 3 GeV (gigaelectronvolts or 10^9). The energy boost is supplied by an RF chamber through which the electrons travel on each orbit of the ring.

The storage ring

The booster ring channels the electrons into the **storage ring**, a doughnut-shaped tube (Figure 17.2.4). In the Australian Synchrotron this ring has a radius of 34.3 m and a circumference of 216 m. Around this ring are 14 bending magnets, each 1.7 m long. These keep the electrons in the circular path. They are separated by 14 straight sections in which focusing magnets keep the electrons confined to a flat beam less than half a millimetre wide and only two-hundredths of a millimetre high.


FIGURE 17.2.4 The booster to storage ring transfer line in the Australian Synchrotron.

In the storage ring of the synchrotron, electrons orbit for hours at a time at speeds near that of light. A series of magnets makes them bend in arcs as they travel through the ring. As the electrons change direction (i.e. accelerate), they emit synchrotron radiation.

Several different types of magnets are used to direct the beam of charged particles. Bending magnets, called dipole magnets, guide the particles through small arcs that combine to produce 360° of bending around the ring. Other magnets called quadrupole and sextupole magnets refocus the beam to prevent it diverging and keep the particles in stable orbits. Steering and corrector magnets correct and fine-tune the orbit to one-millionth of a metre. The specific arrangement and strength of the magnets, called the lattice, dictates the characteristics of the synchrotron light produced, including:

- brightness
- polarisation
- energy distribution
- coherence.

As the electron beam circulates, it radiates synchrotron light and loses energy. To counter this, the electrons pass through RF cavities, like those in the booster ring, to replenish the energy lost. The RF cavities have electromagnetic fields oscillating at radio frequencies produced by amplifiers located next to the storage ring. These fields oscillate in polarity extremely quickly, up to 500 million times per second (500 MHz). The process ensures that the electrons stay at a constant energy and remain stored in the ring.

Despite the RF cavities, the beam is still not perfectly stable. All synchrotron beams will gradually reduce in intensity with time. Some electrons are lost to collisions between electrons and gas molecules in the near vacuum of the ring. To minimise these losses, the vacuum chamber must be kept at a pressure of about one-thousandth of one-billionth of normal atmospheric pressure, or less than 10^{-7} Pa. Under these conditions, the beam typically loses half of its intensity over a 5–50 hour period. New electrons are injected into the beam at 4–24 hour intervals to replace those lost through collisions and energy losses.

The unused high-energy X-rays given off by the storage ring are continually absorbed by radiation shielding. The shield wall surrounding the storage ring is usually made of lead and concrete, forming a tunnel that completely encloses the storage ring, except for the beamlines through which radiation is guided. This design feature is critical for employee safety during synchrotron operation.

Beamlines

A **beamline** is the path that synchrotron light travels from the storage ring, where it is produced, to its target experimental work (Figure 17.2.5). The point at which the beamline meets the storage ring is called the front end. A beamline is typically a stainless steel tube, 15–35 m in length and around 4 cm in diameter. The dimensions depend greatly on the technique being performed on the beamline and the application of that technique. A typical beamline consists of an optics room, an experiment room and a control room.

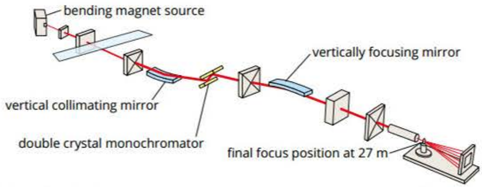


FIGURE 17.2.5 The arrangement of mirrors and crystal monochromator in the beamline for studying proteins and microcrystals and small molecule X-ray diffraction at ANSTO.

Inside the optics room, synchrotron light is modified according to the needs of its experimental use. Sometimes scientists will wish to use only a specific range of wavelengths of synchrotron light for their experiments, rather than all of the light produced. A device called a monochromator, either a crystal or a grating, is used as a wavelength selector. As a beam hits this device, particular wavelengths are diffracted at different angles. By rotating the monochromator, a specific light frequency can be selected from the broad band of frequencies available in the incident beam. As it is prepared for its role in an experiment, synchrotron light may also be:

- aligned using slits
- refocused using mirrors
- lessened in intensity using attenuators within the optics room.

Thin beams of synchrotron light are directed onto a specific target or sample being examined within an experiment room. Scientists control their experiments from an external control room in which they are protected from the intense electromagnetic radiation being used in the experiments. All synchrotrons have a number of beamlines, each directing the synchrotron light to an experimental station. The Australian Synchrotron started out with 13 beamlines but has been designed to allow for additional beamlines as demand for experimental time increases.

COLLIDERS AND PARTICLE PHYSICS

Since the announcement of the proof of the existence of the Higgs boson, colliders have been the superstars of particle accelerators. The best known of the colliders is the Large Hadron Collider (LHC). Figure 17.2.6 shows an image of particles that were created in the LHC. The LHC is only one of a group of accelerators housed at CERN, Geneva, near the French–Swiss border. Since 1934, colliders have been at the forefront of experimentation in particle physics. This research has led not only to a better understanding of the fundamental nature of matter but also to a range of discoveries that are being applied in areas as diverse as medicine and environmental science.

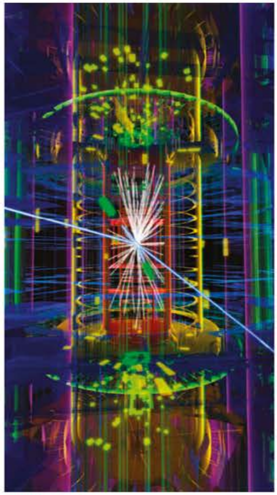


FIGURE 17.2.6 This image shows the shower of particles that were created following a particle collision in the ATLAS detector. The particles were accelerated in the LHC at CERN.

The evolution of colliders

Cyclotrons

The cathode ray tube and the Van de Graaff generator were arguably the first particle accelerators. The cathode ray tube was developed by German physicist Ferdinand Braun in 1897. The Braun tube, as it was originally known, directed electrons from an unheated cathode in an evacuated glass tube, through a magnetic field onto a phosphor-covered screen. The magnetic field was used to change the direction of the electrons and accelerate them towards a particular portion of the screen. Varying the strength of the magnetic field changed the acceleration of the electrons. The basic principles of the Braun tube became the basis of cathode ray tube televisions and computer screens.

Braun's development of the cathode ray tube followed J. J. Thompson's discovery of electrons in the same year. To physicists, the discovery of the electron suggested that atoms could be broken down into smaller and smaller particles and the science of particle physics was born. During the 1920s, physicists believed that electrons and the earlier-discovered protons were the fundamental particles within an atom. However, discoveries in the 1930s and 1940s quickly made physicists rethink this idea, and the study of elementary particle physics began. Into the field of particle physics came the **cyclotron** (Figure 17.2.7), a device specifically developed to allow further investigation of fundamental particles.

FIGURE 17.2.7 In the 1930s, physicist Ernest Lawrence built the first successful cyclotron, at Berkeley, California. It was just 13 cm in diameter and it accelerated protons to an energy of 80 keV.

An American physicist, Ernest Lawrence, was the first to produce a working cyclotron in 1932. This first cyclotron was just 13 cm in diameter and accelerated protons to 80 keV ($1\,\text{keV} = 1.6 \times 10^{-16}\,\text{J}$).

Unlike Van de Graaff generators that only accelerate electrons once via the voltage on the dome of the generator, the cyclotron accelerates the particles over and over, leading to particle energies many times the original. In Lawrence's cyclotron, protons were injected into the centre of a cylindrical space between the poles of a large electromagnet. The magnetic field caused the particles to move in a circular path, hence the name. An electric field accelerated the particles further, causing them to move in an outward spiralling path towards the outside edge of the cyclotron, with ever-increasing energy. At the rim, the particles were directed out of the cyclotron to hit a target. The high-energy particles colliding with the target generated secondary particles, which could then be studied.

It is hard to imagine that the start of developments in the study of particle physics in the Large Hadron Collider (LHC) came from an instrument just 13 cm across. Lawrence continued to develop successive models of the cyclotron, which increased the size and the energies produced. By 1945 he'd built a 4.7 m 'synchrocyclotron', which accelerated particles to energies of 730 MeV.

As the study of particle physics accelerated in the 1950s, so did the energies developed by successive generations of particle accelerators. To date, the largest cyclotron ever built, an 18 m unit at the University of British Columbia in Canada, remains in operation and produces protons with energies of 500 MeV. Until the 1950s, when it was superseded by the synchrotron, the cyclotron was the most powerful particle accelerator available to particle physics.

PHYSICS IN ACTION CCT

Cyclotrons

To perform experiments at very high energies, linear accelerators would need to be extremely long. For this reason, in the 1930s the American physicist Ernest O. Lawrence designed the first circular accelerator. This is called a cyclotron, and it won Lawrence a Nobel Prize in 1939. In some respects, the cyclotron operates as a spiral-shaped linac. Protons are often used as the accelerating particles in this machine.

Here, the many drift tubes are replaced by two semicircular, D-shaped, hollow copper chambers, called 'dees'. These are the positive and negative electrodes of the cyclotron between which exists a strong electric field. The dees sit back to back, giving the cyclotron its circular shape, and lie between the poles of a powerful electromagnet. The inside of the metallic dee is shielded from the electric field. The magnetic field acts on the particles, producing a circular path. When a particle emerges from the dee, the sign of the accelerating potential is reversed, causing the particle to speed up towards the other dee. This occurs so that a proton will accelerate towards a negatively charged dee as it exits the positively charged dee. Each time the particles cross the gap between the dees, their speed increases and they travel in a semicircle of larger radius. They gain energy with each revolution until they have sufficient energy to exit the accelerator (Figure 17.2.8).

A key to the operation of the cyclotron is that the frequency of the radio-frequency (RF) generator that produces the alternating field must match the frequency of the circulating charged particles. The charged particles travel in a path of radius:

$$r = \frac{mv}{qB}$$

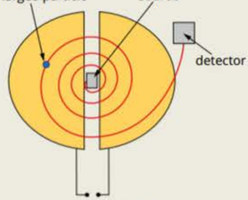


FIGURE 17.2.8 The cyclotron operates in some ways like a spiral linac—particles are accelerated from a source, through semicircular chambers called dees, until they gain sufficient energy to exit.

Their speed is then $v = \frac{rqB}{m}$ and the time taken for one orbit of the cyclotron is $t = \frac{d}{v}$, where d is the path distance of one revolution:

$$d = \frac{2\pi r}{v}$$

$$= \frac{2\pi mv}{qBv} \text{ (substituting the above expression for path radius, } r\text{)}$$

$$= \frac{2\pi m}{qB}$$

Strangely enough, the time taken for one revolution of the cyclotron does not depend upon the velocity of circulating charges. This is because as the speed increases, the radius of path travelled also increases and the time taken for each orbit remains the same.

In 1943, the Adelaide-born physicist Marcus Oliphant, while working in Britain, suggested modifying the cyclotron design to produce a synchrotron.

CERN

The term CERN is derived from the French 'Conseil Européen pour la Recherche Nucléaire', or European Council for Nuclear Research. CERN was founded in 1952 to establish a world-class facility for particle-physics research in Europe. Initially concentrating on the inside of the atom, hence the word 'nuclear', CERN's main area of research is the study of the fundamental particles that make up matter and the forces acting between them. It is referred to as the European laboratory for particle physics because of this research focus. CERN started up its first synchrotron in 1957. The development of the CERN facility since then is indicative of the development of collider technology in general. Figure 17.2.9 shows how the technology at CERN has developed over time.

The first accelerator commissioned at CERN was a 600 MeV synchrocyclotron, a cyclotron in which higher energies were reached compared with a standard cyclotron. This was achieved by synchronising the accelerating voltage with the particle velocity. While in terms of energy this wasn't a significant advance over Lawrence's original cyclotron, the design of this and other synchrocyclotrons did allow for potentially much higher energy levels. The 1950s Cold War–era became something of a technological race to develop the highest energy accelerator.

After the launch of the Russian 'Synchrophasotron', a 10 GeV accelerator, CERN developed the Proton Synchrotron (PS), first accelerating protons in November 1959. The Proton Synchrotron achieved a beam energy of 28 GeV on its first day of operation. It was the world's highest energy particle accelerator at the time. Since then, the PS has had its proton beam intensity increased to 1000 times its original intensity and it has accelerated many different kinds of particles besides protons. It remains in active use today, feeding accelerated particles to experiments and to other accelerators.

Particle accelerators have continued to develop in design, in the range of particles being accelerated and in the energy levels attained. In 1976, CERN switched on the 'Super Proton Synchrotron', with a beam of protons circulating a distance of 7 km. It achieved 400 GeV for the first time and is now running at up to 450 GeV. In 2008, the 27-kilometre-diameter Large Hadron Collider started up. It is the largest particle accelerator ever developed, and initial intensities of up to 8 TeV (8×10^{12} eV) were achieved, which were later increased to 13 TeV being possible. This represents an increase in energy levels of almost 30 000 times greater than the levels achieved by the cyclotrons of the early 1950s.

Not all particle accelerators accelerate particles. The antiproton decelerator at CERN has been designed to slow down antiprotons, reducing their energy, to allow the study of antimatter. Table 17.2.1 compares the advances in the design, energy level and range of particles that are accelerated as particle accelerator designs have advanced.

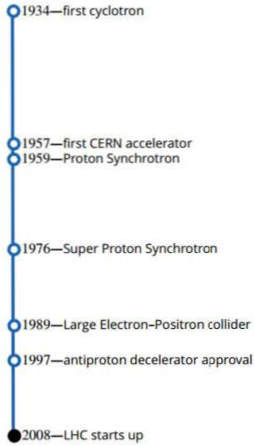


FIGURE 17.2.9 Timeline of the introduction of accelerators at CERN. The first unit deployed at CERN was in 1957, 23 years after Ernest Lawrence built the first cyclotron.

TABLE 17.2.1 A comparison of particle accelerators.

Commissioned	Particle accelerator	Energy level	Particles accelerated
1934	first Lawrence cyclotron	80 eV	protons
1945	Lawrence Synchrocyclotron	730 eV	protons
1957	first CERN Synchrocyclotron	600 MeV	protons
1957	Russian Dubna Synchrophasotron	10 GeV	protons
1959	CERN Proton Synchrotron	28 GeV	protons, range of other particles
1976	CERN Super Proton Synchrotron	400–450 GeV	protons, range of other particles
1989	Large Electron–Positron collider	100–209 GeV	electrons, positrons, range of other particles
2008	Large Hadron Collider	14 TeV	protons, lead ions, hydrogen ions, positrons, antiprotons

PHYSICS IN ACTION ICT

Application of particle accelerators

Particle accelerators aren't just used as a research tool for particle physicists. The Large Hadron Collider got so much media attention with the proof of the existence of the Higgs boson in 2012 that you'd think that was its sole purpose. These large particle colliders will continue to be developed into the future as scientists try to produce experimental evidence for new theories and ideas. Large-scale developments of the LHC will see it continue to play a leading role well into the future.

Particle accelerators take many forms—from synchrotrons like the Australian Synchrotron at Monash University, to relatively small-scale particle accelerators used for implanting ions in silicon wafers, and even the humble cathode ray tube found in old-style televisions. This range of particle accelerators can be used for a wide range of applications and for the development of diverse technologies.

Other areas where particle accelerator research is leading to groundbreaking discoveries include (but are not limited to):

- cancer treatment
- carbon dating
- art restoration
- the production of artificial body parts.

Believe it or not, X-ray microscopy has even led to better nappies!

PARTICLE PHYSICS AND COLLIDERS

FIGURE 17.2.10 Japanese physicist Hideki Yukawa predicted the existence of mesons.

WS 8.10 WS 8.11

By the mid 1930s it was understood that atoms were made up of smaller particles and that those particles were not limited to the proton, neutron and electron. The positron, neutrino and photon, or gamma (γ) particle, had also been discovered. In 1935, Japanese physicist Hideki Yukawa (Figure 17.2.10) predicted the existence of an additional particle in order to balance the strong nuclear force. (The strong nuclear force holds particles together in the atomic nucleus and has since been found to hold together elementary particles.) Yukawa predicted that this particle would have a mass somewhere between that of an electron and a proton, so it was called a meson, meaning 'in the middle'.

The meson that Yukawa predicted was finally discovered in cosmic rays in 1947. The search for it also led to the discovery of other particles in the interim and the development of elementary particle physics and the scientific instruments with which the interactions of particles making up an atom could be observed.

The Standard Model of particle physics describes the workings of the atom as much more complicated than the simple models proposed at the beginning of the twentieth century. Research following the discovery of the meson led to not just a few more subnuclear particles, as originally expected, but many hundreds more. Colliders, and the ever-increasing particle energies that successive improvements to particle accelerators allowed, have been fundamental to these discoveries.

Recall that the hadrons are a group of particles that includes protons and neutrons. Leptons are the group of particles that includes electrons. Hadrons interact via the strong nuclear force; leptons don't and are far more numerous. While leptons are considered elementary particles, collider experiments indicated that hadrons have an internal structure. It was discovered that hadrons are made up of more fundamental elementary point particles termed quarks. Scientists proposed that there are six quarks:

- up
- down
- charmed
- strange
- top (also called truth)
- bottom (also called beauty).

So far there is no direct evidence of up and down quarks; they may not even exist singly. Strong evidence of the truth or top quark wasn't observed until 1995. Its comparatively high rest mass required a proton–antiproton collision of almost 2 TeV in order to detect the decay products from this very-short-lived (10^{-23} s) particle. The LHCb (Large Hadron Collider beauty) detector has been specifically built to detect the bottom or beauty quark using the upgraded LHC.

As mentioned in the previous section, all particles also have antiparticles—subatomic particles that have the same mass as the particle but with the opposite electric or magnetic properties. Many subnuclear particles are unstable and decay. Unstable particles influenced by the strong nuclear force decay more quickly than those caused by the weak nuclear force. Recording collision events in particle accelerators is essential to observing these extremely short-lived particles via observations of the particles to which they decay.

Table 17.2.2 summarises many of the subatomic particles along with their charge, rest mass and life span until they decay into other particles or forms. The extremely short life spans of many of the particles makes observing them via a single collision highly problematic. This is why there are millions of collision events in a typical run at the LHC.

TABLE 17.2.2 Characteristics of subatomic particles and their antiparticles.

Category	Name	Symbol and charge	Antiparticle symbol	Rest mass (MeV c^{-2})*	Lifetime (s)
Gauge boson	photon	γ	γ (self)	0	stable
	W	W^+	W^-	80.33×10^3	3×10^{-25}
	Z	Z^0	Z^0 (self)	90.19×10^3	3×10^{-25}
Leptons	electron	e^-	e^+	0.511	stable
	electron neutrino	ν_e	$\bar{\nu}_e$	0	stable
	muon	μ^-	μ^+	105.7	2.2×10^{-6}
	muon neutrino	ν_μ	$\bar{\nu}_\mu$	0	stable
	tau	τ^-	τ^+	1777	2.91×10^{-13}
	tau neutrino	ν_τ	$\bar{\nu}_\tau$	0	stable
Hadrons: mesons	pion	π^+	π^-	139.6	2.60×10^{-8}
	kaon (also other variations)	K^+	K^-	493.7	1.24×10^{-8}
	eta	η	η (self)	547.5	5×10^{-19}
	and others				
Hadrons: baryons	proton	p	$\bar{p}$	938.3	stable
	neutron	n	$\bar{n}$	939.6	887
	lambda	Λ^0	$\bar{\Lambda}^0$	1115.7	2.63×10^{-10}
	sigma (also other variations)	Σ^0	$\bar{\Sigma}^0$	1189.4	0.80×10^{-10}
	xi (also other variations)	Ξ^-	Ξ^+	1321.3	1.64×10^{-10}
	omega	Ω^-	Ω^+	1672.5	0.82×10^{-10}

*1 MeV $c^{-2} = 1.783 \times 10^{-30}$ kg

The design of particle accelerators and the associated detectors must take into account the predicted life span of the particle before it decays, the charge on the particle and the relative rest mass (a larger mass means that more energy is required) in order to be sure that the particle being detected meets theoretical predictions within reasonable experimental constraints.

The Large Hadron Collider

The Large Hadron Collider (LHC) at CERN is one of the latest developments in particle accelerator technology. Due to its enormous size and high energies, the particle accelerator is probably the best known outside of particle physics research. However, the LHC is only the most recent and highest energy accelerator in a chain of accelerators and experiments that work together to feed particle research activities at CERN. Together, the particle detectors working from the LHC allow specialised investigations at the leading edge of particle-physics research.

Facts and figures about the LHC:

- It gets its name from: its size (Large); its ability to accelerate both protons and ions, both of which are hadrons (Hadron); and because it is a collider in which two particle beams travelling in opposite directions collide at points where the two rings of the LHC intersect (Collider).
- It is the largest scientific instrument ever built: the tunnel has a circumference of 27 km with an average depth of 100 m below ground level (Figure 17.2.11). Being located underground has the extra benefit of shielding the accelerator from external radiation, although cost was the main factor for building underground.
- The LHC was initially considered in the 1980s with the intent of producing an accelerator capable of reaching the (predicted) energies needed for research on particles that would have existed just 10^{-12} s after the big bang.
- It was built at a cost of €3 billion (equivalent to around AU$4.5 billion), with financial contributions from non-member countries (i.e. Japan, India and the USA) helping to reduce the construction time.
- The project was built for an initial four experiments (ATLAS, ALICE, CMS and LHCb). Two additional experiments have since been added.
- Its first collisions in October 2008 produced particles with an (then record) energy of 4 TeV. After a two-year upgrade, it is now capable of energies potentially reaching 13 TeV or 6.5 TeV per proton beam in a head-on collision. Lead-ion beams now have a collision potential of 1150 TeV.
- Protons complete over 11000 revolutions of the 27 km circumference each second, travelling at a speed of 0.999999991 times that of light. There are up to 600 million collisions per second.

FIGURE 17.2.11 A section of the LHC's 27 km tunnel.

The main steps in the acceleration of particles in the LHC are outlined below.

- Hydrogen atoms are pumped from storage and are stripped of their electrons to create protons.
- The protons travel through a linac reaching an energy of 50 MeV and are injected into the Proton Synchrotron booster.
- The Proton Synchrotron booster accelerates the protons to 1.4 GeV, after which they are fed into the Proton Synchrotron and accelerated to 25 GeV.
- Protons are then sent to the Super Proton Synchrotron and are accelerated to 450 GeV.
- Finally, the protons are injected into the LHC in both clockwise and anticlockwise directions to create the two beams that will collide.
- The beams are accelerated to 6.5 TeV each over a period of 20 minutes, after which they can be kept in the storage rings for a number of hours.

Goals of the LHC

The Standard Model summarises the current understanding of particle physics. It has been tested and has been successful in predicting previously unknown particles. However, it is unable to answer questions on dark matter and why there is comparatively very little antimatter in the universe, among others. The main goals of the LHC aim to add to the understanding of particle physics beyond the Standard Model. Some examples are given below.

- Bosons are also referred to as force-carrying particles and are used to explain forces being exerted on particles by fields. Scientists are seeking a unified theory for the four natural forces. The Standard Model links the weak nuclear force, the strong nuclear force and the electromagnetic force, but is unable to construct a similar theory for gravity. Supersymmetry—the existence of more massive particles than are currently known—could lead to a unified theory.
- The reasons why objects and particles have mass cannot be predicted from the Standard Model. A separate model is needed. The Higgs field is theorised to interact with particles such that they acquire their mass. The Higgs field has a Higgs boson particle associated with it. The successful detection of the Higgs boson and subsequent determination of its mass was a major achievement of the LHC.
- Astronomical observations suggest that only 4% of all of the matter in the universe is visible. The LHC is searching for evidence of dark matter and dark energy, which are theorised to account for 23% and 73% of the remaining matter respectively.
- Matter and antimatter existed together in equal quantities at the time of the big bang but today, as far as we know, there is comparatively very little antimatter. Experiments at the LHC will attempt to determine why.
- Heavy ions such as lead colliding at high energies form hot, dense matter. The LHC will be used to investigate the state of matter called the quark–gluon plasma, which is theorised to have existed in the early universe.

The initial four experiment stations developed as part of the LHC project are described below.

LHCb

LHCb stands for 'Large Hadron Collider beauty', an experiment designed to investigate the slight differences between matter and antimatter. It is shown in Figure 17.2.12 on page 470. It was specifically built to detect the bottom (or beauty) quark, also called the b quark, which is a frequent decay product for the Higgs boson. By studying the b quark, scientists hope to better understand what happened immediately after the big bang.

According to the theory, approximately 14 billion years ago, the big bang produced both matter and antimatter. As the matter cooled and expanded, its composition changed. Within one second, it is believed that antimatter had largely ceased to exist, leaving matter to create the universe as you know it. The b quark, an antiparticle, is believed to be key in understanding that process.

Every collision event at the LHC forms a broad range of quarks. The ATLAS and CMS detectors enclose the entire collision point within a detector. By contrast, the LHCb is able to detect b quarks by having a series of sub-detectors at points largely forwards of the collision point in order to focus on particles thrown forwards of the collision. The first sub-detector is close to and forwards of the collision event; the rest of the chain of sub-detectors are arranged one behind the next, over a total length of 20 metres.

Since the quarks created by the collision event decay quickly into other subatomic particles, the LHCb features moveable tracking detectors consisting of a forward spectrometer and planar detectors.

Weighing 5600 tonnes, the LHCb is sited 100 metres below ground. In May 2015, when the LHC was reopened after an upgrade to energies of 13 TeV, the LHCb detected events from proton–proton collisions. These initial collisions were only in the 450 GeV range and hence not suitable for full physics studies, but were able to confirm that the detector was recording suitable events. Follow-up studies in June 2015 were among the first to use the increased energy of the LHC.

FIGURE 17.2.12 The underground cavern holding the LHC beauty, or LHCb, experiment of the Large Hadron Collider at CERN. Much of the work on this experiment is directed to the study of antiparticles.

17.2 Review

SUMMARY

- All particle accelerators require a source of charged particles. In a cathode ray tube these are provided by a device called an electron gun.
- The force acting on the charged particles from electric and magnetic fields will change the direction of the beam. As the charged particles are accelerated by these forces, work is done and the charges gain kinetic energy.
- A synchrotron is a doughnut-shaped particle accelerator designed to circulate electrons around a closed path at speeds very close to that of light. The accelerating electrons generate beams of infrared, UV, visible and X-ray radiation.
  - Electrons are emitted from an electron gun in pulses. They are accelerated through the linac by powerful bursts of radio-frequency (RF) radiation.
  - Electrons then travel around a booster ring, and are accelerated further as they pass through RF cavities until reaching an energy of 3 GeV. The magnetic field of the bending magnets is increased as the velocity of electrons increases within the booster ring.
  - Electrons are channelled into the storage ring. Synchrotron radiation (light) is produced as the particles travel through the strong magnetic fields of the dipole magnets.
  - Electrons replace energy lost due to the production of synchrotron light as they pass through radio-frequency cavities in the storage ring.
  - Synchrotron light leaves the storage ring, passing down beamlines to a number of independent experimental stations.
- Colliders are used to collide two beams of particles head-on. They were developed for the study of elementary particle physics.
- Ernest Lawrence was the first to produce a working cyclotron in 1932. This first cyclotron was just 13 cm in diameter and accelerated protons to 80 keV (1 keV = 1.6×10^{-16} J).
- As the study of particle physics rapidly progressed in the 1950s, so too did the energies developed by successive generations of particle accelerators. The LHC can now achieve energy levels of 13 TeV, almost 30 000 times that achieved in the early 1950s.
- Particle accelerators have also developed to be able to accelerate a range of particles, including protons, electrons, antiprotons, positrons and ions, to allow specific experiments in particle physics.
- Not all particle accelerators accelerate particles. The antiproton decelerator at CERN has been designed to slow down antiprotons to allow the study of antimatter.
- The design of particle accelerators and the associated detectors must take into account the charge on the particle being investigated, the relative rest mass (a larger mass means that more energy is required) and the predicted lifetime of the particle before it decays, in order to be sure that the particle being detected meets theoretical predictions within reasonable experimental constraints.
- The main goals of the LHC aim to add to the current understanding of particle physics beyond the Standard Model.

KEY QUESTIONS

1 What property do particles need to have in order to be accelerated by particle accelerators?

2 In a cathode ray tube, how are particles accelerated?

3 List the four main features of the Australian Synchrotron in order from the initial emission of electrons at the cathode through to the experiment and control rooms.

4 In the booster ring of a synchrotron, what kinds of speeds can the electrons be accelerated to?

5 How did cyclotrons get their name?

6 How is the synchrotron an advance on the earlier cyclotron designs?

7 The LHCb detector has been designed to investigate differences between matter and antimatter. What design feature is incorporated into the LHCb to enable this form of investigation?

Chapter review

17

KEY TERMS

annihilation
antimatter
baryon
baryon number
beamline
booster ring
colour charge
cyclotron
fermion
gauge boson
gluon
hadron
Higgs boson
linac
Large Hadron Collider
lepton
meson
neutrino
quantum number
quark
Standard Model
storage ring
strong nuclear force
synchrotron
weak nuclear force

REVIEW QUESTIONS

1 What happens in the process called annihilation?

2 Which quarks make up protons and neutrons?

3 Which forces are explained by the Standard Model of particle physics?

4 Explain how the proton gains a +1 charge from its constituent quarks and how a neutron has a neutral charge from its constituent quarks.

5 A neutron in an evacuated container decays to produce a proton and two other particles. The proton is then attracted to an electron and becomes the nucleus of a hydrogen atom. The atom then slowly drifts to the base of the container.
Order the four fundamental forces that have been involved in the sequence of events described above from first to last. Place any forces that were not involved last in your list.

6 In your own words, explain how the Standard Model explains three of the four fundamental forces. As part of your answer, identify the three forces and the associated particle(s).

7 Is an electron a composite or a fundamental particle and is it a lepton or a hadron?

8 Where in a cathode ray tube are electrons released from?

9 What is linac the shortened term for?

10 What is a beamline?

11 Synchrotron radiation is described as having a high degree of collimation and very low beam divergence. Which particular experimental technique is this characteristic of synchrotron radiation ideal for?

12 In which section of the Australian Synchrotron do electrons move from a heated filament?

13 High-energy electrons travelling in a curved path under the influence of a magnetic field will emit light. Within what frequency range is this light? (Give a qualitative answer only.)

14 Strong experimental evidence for the truth or top quark wasn't observed until 1995. What physical constraint of particle accelerator design would have been the main limiting factor that meant the top quark couldn't be observed earlier?

15 The Standard Model of particle physics seeks to unify theories about the weak nuclear force, electromagnetic force and the strong nuclear force. It has been particularly successful in predicting particular particles. Name two areas where the Standard Model is deficient.

16 Compared with X-rays from conventional sources, X-rays from synchrotron radiation have a considerably higher intensity. What benefit does this offer researchers?

17 What specific attributes of synchrotron radiation make synchrotron sources more useful than laser sources for particular investigations of elements that absorb particular wavelengths?

18 Copy the table and list the properties of X-rays emitted by a conventional X-ray tube and a synchrotron.

X-ray tube	Synchrotron

19 Describe the role of the storage ring in a synchrotron.

20 In which section of the Australian Synchrotron is the strength of the magnetic field increased as the velocity of the electrons increases?

21 In which section of the Australian Synchrotron are electrons accelerated by RF radiation?

22 The Large Hadron Collider, or LHC, accelerates particles from a proton or hadron source. It is the largest scientific instrument ever built. Place the following components of the LHC in order from source to detector:

storage ring
electrons are stripped away
hydrogen atom source
linac
detectors
Proton Synchrotron Booster
Super Proton Synchrotron

23 After completing the activity on page 448, reflect on the inquiry question: How is it known that human understanding of matter is still incomplete?

MODULE 8 • REVIEW

REVIEW QUESTIONS

From the universe to the atom

WS 8.12 WS 8.13 MR 8

Multiple choice

1 Which of the following atomic models did Rutherford propose?

A solid ball
B plum pudding
C nuclear
D planetary

2 In general, which of the following is true for most stars?

A luminosity increases with the surface temperature
B luminosity decreases with the surface temperature
C luminosity is not related to the surface temperature
D luminosity is directly proportional to the surface temperature

3 Main sequence stars are known to obey a mass–luminosity relationship. If a star is found to have a luminosity eight times that of the Sun, what is its mass likely to be (in terms of the mass of the Sun, $M_\odot$)?

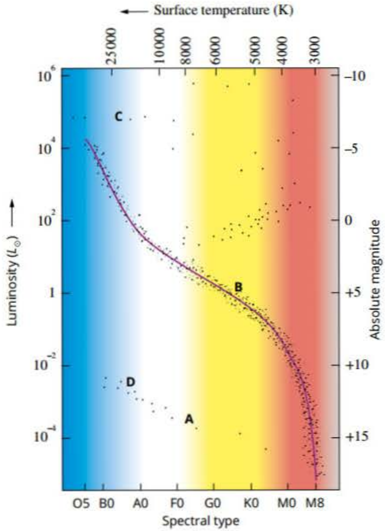
A 8 $M_\odot$
B 4 $M_\odot$
C 2 $M_\odot$
D 1 $M_\odot$

4 Which of the following best describes what is plotted on the Hertzsprung–Russell diagram?

A absolute magnitude against the luminosity of the stars
B apparent magnitude against the temperature of the stars
C temperature against spectral type of the stars
D luminosity against spectral type of the stars

The following information relates to questions 5 to 7.

Consider the Hertzsprung–Russell diagram below.



5 Which letter corresponds to a Sun-like star?

A A
B B
C C
D D

6 Which letter corresponds to a blue supergiant?

A A
B B
C C
D D

7 Which letter corresponds to an old star that once was a Sun-like, main-sequence star?

A A
B B
C C
D D

MODULE 8 • REVIEW

8 Which of the following best outlines the life cycle for a massive star, a star with a mass much greater than that of the Sun?

A main-sequence star → planetary nebula → red supergiant → supernova → black hole
B main-sequence star → planetary nebula → supernova → red supergiant → black hole
C planetary nebula → main-sequence star → supernova → red supergiant → black hole
D planetary nebula → main-sequence star → red supergiant → supernova → black hole

9 The energy output from the Sun involves the fusion of four protons to form one helium nucleus, two antineutrinos, two gamma rays and 24.7 MeV of energy. What is the energy equivalent of 24.7 MeV in joules?

A 3.95×10^{-18} J
B 3.95×10^{-15} J
C 3.95×10^{-12} J
D 3.95×10^{-9} J

10 The first fission reaction ever to be observed is shown in the equation below:

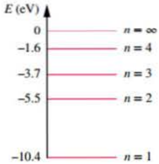
$${}^{235}_{92}\text{U} + {}^{1}_{0}\text{n} \rightarrow {}^{92}_{36}\text{Kr} + {}^{141}_{56}\text{Ba} + x\,{}^{1}_{0}\text{n} + \text{energy}$$

How many neutrons (x) are produced in this fission reaction?

A 1 **B** 2 **C** 3 **D** 4

The following information relates to questions 11 and 12.

An electron beam of energy 7.0 eV passes through some mercury vapour in the ground state and excites the electrons to the $n = 3$ energy level.

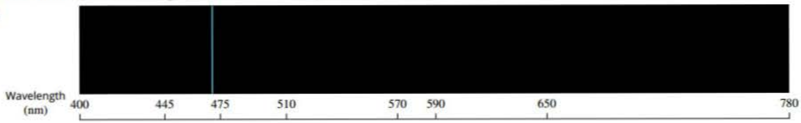


11 A photon collides with a mercury atom in the ground state. As a result, a 30.4 eV electron is ejected from the atom. What was the wavelength of the incident photon?

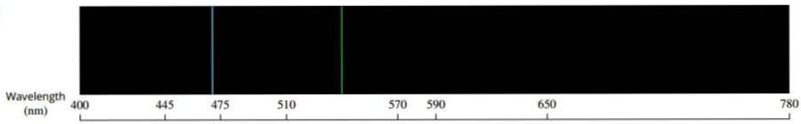
A 3.04×10^{-8} m
B 3.5×10^{-8} m
C 3.6×10^{-8} m
D 4.09×10^{-8} m

12 Calculate the wavelengths of the emitted photons present during the transition from the $n = 3$ energy level to the ground state. Which of the following spectra describes any of these photons?

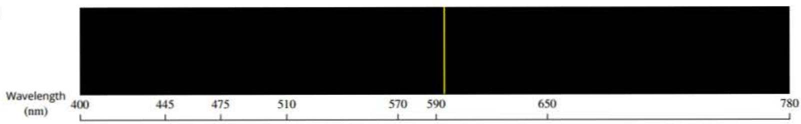
A



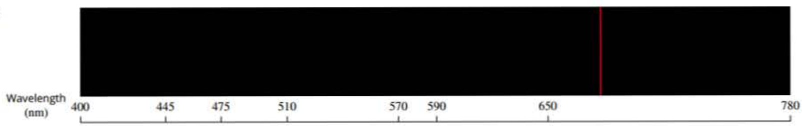
B



C



D



13 For an electron and a proton to have the same wavelength:

A the electron must have the same energy as the proton.

B the electron must have the same speed as the proton.

C the electron must have the same momentum as the proton.

D It is impossible for an electron and a proton to have the same wavelength.

14 If a particular atom in the sample has not decayed during the first half-life, which one of the following statements best describes its fate?

A It will definitely decay during the second half-life.

B It has a 50% chance of decaying during the second half-life.

C The probability that it will decay cannot be determined.

D If it does not decay during the first half-life, it will not decay at all.

15 Tritium (hydrogen-3) is radioactive and its decay equation is shown.

$$^{3}_{1}\text{H} \rightarrow \text{X} + {}^{0}_{-1}\text{Y}$$

Which of the following best describes the nature of Y in the decay equation?

A It is a positron.

B It is an electron.

C It is a proton.

D It is a neutron.

16 A nuclear scientist has prepared equal quantities of two radioisotopes of bismuth, ^{211}Bi and ^{215}Bi. These isotopes have half-lives of 2 minutes and 8 minutes respectively. Assume when answering these questions that each sample has the same number of atoms. Which one of the following statements best describes the activities of these samples?

A The samples start with an equal activity, then bismuth-211 has the greater activity.

B Bismuth-211 initially has four times the activity of bismuth-215.

C Bismuth-215 initially has four times the activity of bismuth-211.

D Bismuth-211 initially has twice the activity of bismuth-215.

17 Which of the following lists includes only subatomic particles that are classified as hadrons.

A protons and neutrons

B protons and electrons

C neutrons and electrons

D protons and neutrinos

18 What is the speed of an electron that has been accelerated across a potential difference of 5 kV? Take the mass of an electron to be 9.1×10^{-31} kg. Ignore any relativistic effects.

A $1.3 \times 10^{6}\,\text{m s}^{-1}$

B $4.2 \times 10^{7}\,\text{m s}^{-1}$

C $3.3 \times 10^{15}\,\text{m s}^{-1}$

D $1.0 \times 10^{17}\,\text{m s}^{-1}$

Short answer

19 Explain what is meant by the redshift of stellar spectra and why this is believed to be evidence that the universe is expanding.

20 An electron and a positron annihilate with the production of two photons.

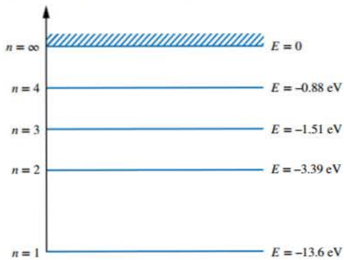
a Explain the source of the energy of the photons.

b Assuming the two leptons each had minimal kinetic energy, calculate the total photon energy. (Assume the mass of an electron or positron is 9.1×10^{-31} kg.)

21 a Explain why the development of the Bohr model of the hydrogen atom was significant in the development of a comprehensive understanding of the nature of light.

b How did Niels Bohr explain the observation that for the hydrogen atom, when the frequency of incident light was below a certain value, the light would simply pass straight through a sample of hydrogen gas without any absorption occurring?

22 Consider the energy-level diagram for the hydrogen atom shown below. A photon of energy 14.0 eV collided with a hydrogen atom in the ground state.



a Explain why this collision will eject an electron from the atom.

b Calculate the energy of the ejected electron in electronvolts and in joules.

c What is the momentum of the ejected electron?

d Determine the wavelength of the ejected electron.

e A hydrogen atom in the ground state collides with a 10.0 eV photon. Describe the result of such a collision.

23 In a Millikan oil-drop experiment, an oil drop of mass 1.96×10^{-14} kg is stationary between two horizontal parallel plates. The plates have a separation distance of 1.6 mm with 240 V between them.

a Determine the size and direction of the electric field strength between the plates.

b Calculate the size of the charge that must exist on the oil drop.

c How many excess electrons are on the drop?

24 In a double-slit interference experiment, an electron beam travels through two narrow slits, 20 mm apart, in a piece of copper foil. The resulting pattern is detected photographically at a distance of 2.0 m. The speed of the electrons is 0.1% of the speed of light.

a Calculate the de Broglie wavelength of the electrons used in the experiment.

b What do you expect to see on the photographic plate?

c Given that electrons are particles, how do you interpret the behaviour of the electrons in this experiment?

d If the experiment were to be repeated using neutrons, at what speed would a neutron need to travel to have the same de Broglie wavelength as the electrons in part a?

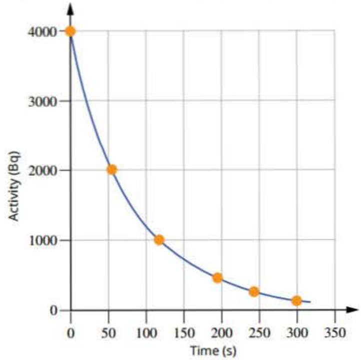
25 Physicists can investigate the spacing of atoms in a powdered crystal sample using electron diffraction. This involves accelerating electrons to known speeds using an accelerating voltage. In a particular experiment, electrons of mass 9.11×10^{-31} kg are accelerated to a speed of $1.75 \times 10^{7}\ \text{m s}^{-1}$. The electrons pass through a powdered crystal sample, and the diffraction pattern is observed on a fluorescent screen.

a Calculate the de Broglie wavelength (in nm) of the accelerated electrons.

b Describe the main features of the expected diffraction pattern.

c If the accelerating voltage is increased, what difference would you expect to see in the diffraction pattern produced? Explain your answer.

26 The graph below shows the data obtained in an experiment to determine the half-life of sodium-26.



a Use the graph to work out the half-life of sodium-26.

b If the initial sample contained 150 g of sodium-26, how much of this radioisotope will remain after 5 minutes?

c Sodium-26 is a beta emitter. Write the nuclear equation for its decay.

27 As a result of the disaster at the Fukushima nuclear power plant in 2011, the radioactive isotopes caesium-137 and iodine-131 were released into the atmosphere. Use a periodic table to determine the number of protons, neutrons and nucleons contained in a nuclide of each radioisotope.

28 Copy and complete the table about the properties of exchange particles, using the dot points below.

Particle	Property
gluon	
photon	
W^+, W^- and Z	
(graviton)	

- mediator of the electromagnetic force
- interacts with charged particles
- mediator of the gravitational force
- mediator of the weak nuclear force
- interacts with quarks
- causes nuclear decay
- mediator of the strong nuclear force

29 Explain how the single-slit diffraction experiment relates to the Heisenberg uncertainty principle.

30 The Heisenberg uncertainty principle states:

$$\Delta p \Delta x \geq \frac{h}{4\pi}$$

What does it mean if the value of Δp in the relation gets smaller? Explain your answer.

31 According to the big bang theory, space itself is expanding.

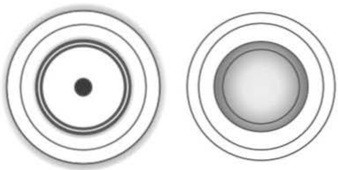
a Use Hubble's law to compare the recession velocity of a star at the position of our closest star, Proxima Centauri at 1.3 pc, with that of the edge of the observable universe, which is taken to be 4.4×10^{26} m away. (Proxima Centauri is in fact moving towards the Sun because of other factors.)

b The speed of the edge of the observable universe that you have calculated exceeds the speed of light. Is that a problem? Explain your answer.

c The Hubble constant is taken to be around $70\,\text{km s}^{-1}\,\text{Mpc}^{-1}$. Calculate the Hubble constant H_0 in SI units and explain why the redshift of a star at the distance of Proxima Centauri is not going to be measurable.

d Explain why, in contrast, stellar objects at much greater distances actually show measurable effects.

32 The image below shows diffraction images that have been obtained by scattering X-rays (left) and electrons (right) off the same sample, which is made up of many tiny crystals with random orientation. The X-rays have a frequency of 8.3×10^{18} Hz.



a Provide an explanation for the fact that the electrons and the X-rays have produced the same diffraction pattern.

b Determine the wavelength of the X-ray photons.

c Determine the de Broglie wavelength of the electrons.

d Calculate the momentum of the electrons.

e Do the X-rays and the electrons have the same energy? Explain your answer.

f How would de Broglie explain the light and dark rings produced when a beam of electrons is fired through a sodium chloride crystal?

g Describe how the wave–particle duality of electrons can be used to explain the quantised energy levels of the atoms.

33 A Geiger counter measures the radioactive disintegrations from a sample of a certain radioisotope. The count rate is recorded in the table.

Activity (Bq)	800	560	400	280	200	140
Time (min)	0	5.0	10	15	20	25

a Plot a graph of activity versus time.

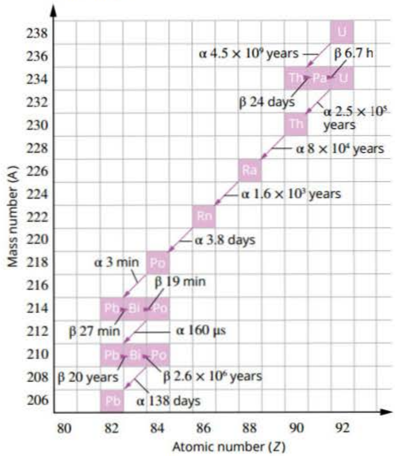
b Use your graph to estimate the activity of the sample after 13 minutes.

c What is the half-life of this element?

d What is the decay constant?

e Determine the activity of the sample after 30 minutes.

f The following diagram indicates the decay series for ^{238}U through the various decay products until the stable isotope ^{206}Pb is reached.



i Complete the following nuclear reactions:

$^{238}\text{U} \rightarrow {}^{234}\text{Th} +$ _____

$^{218}\text{Po} \rightarrow {}^{218}\text{At} +$ _____

ii Explain why both Po and At can have the same mass number, but still be distinct elements.

iii Describe the decay pathway from $^{210}\text{Bi} \rightarrow {}^{206}\text{Pb}$.

iv Explain the difference between the three different polonium isotopes on the chart.

MODULE 8 • REVIEW

34 **a** Copy and complete the table below, using the following list of particles:

electrons, positrons, neutrons, muons, (W^+, W^- and Z), neutrinos, gluons, antiprotons, photons, pions, protons and gravitons

Category	Particle type	Description	Particle name
gauge bosons		mediators of the fundamental forces	
fermions (make up all matter)	leptons	• experience the weak nuclear force, exchanging W and Z bosons • charged leptons experience the electromagnetic force, exchanging photons • do not experience the strong force	
	hadrons	• experience the strong force, exchanging gluons • made up of quarks	
	- baryons	made of three quarks	
	- mesons	made of two quarks	

b Successive generations of particle colliders have produced particles of ever-higher energies. High-energy collisions enabled scientists to determine that some subatomic particles have an internal structure. Which of the following particles have an internal structure: leptons, hadrons, quarks, gauge bosons? Give a reason for your answer.

c The Higgs boson was theorised to interact with other bosons, W and Z in particular, forcing them to go slower than the speed of light. Describe the essential role of the Higgs boson.

d The Standard Model is a very powerful theory and is well supported by experimental evidence from particle accelerators. It can't explain some predicted phenomena that are the subject of current collider experiments. Which of the following are not explained by the standard model?

Higgs boson, dark matter, strong nuclear force, antimatter, weak nuclear force

35 **a** A student is explaining the big bang theory to one of his friends. He tells her that at some point in time, the universe was concentrated in a very small dense mass and then, after the big bang, the mass spread out into space and continues to expand. This explanation contains several mistakes. Give a simple but more accurate explanation.

b Given the extraordinarily high temperatures of the early universe, the radiation from the big bang would have been expected to be very high frequency and short wavelength gamma radiation. Explain why physicists expected the CMB radiation to be detectable in the microwave range.

c When the CMB radiation was mapped it was found that the radiation was essentially uniform in all directions, but with small variations. Explain why these small variations are considered important in stellar evolution.

d What is 'pair production' and why is it important for understanding the creation of matter?

e Without the phase of extremely rapid inflation of the universe, any matter created would have immediately annihilated and the universe would be empty except for the CMB radiation. Explain why this would be the case.

f As the temperature of the universe dropped, it would be expected that less pair creation would occur. Explain why this would be the case.

g Once protons, neutrons and electrons had formed, one would think that atoms would follow. Why could atoms still not exist at this point?

h Fusion followed as protons and neutrons were forced together to form hydrogen, helium and lithium nuclei for a time. Then this process too stopped. Why?

i What condition would the background radiation photons have to satisfy before atoms could form?

j The universe continued to cool and expand. Explain the role of gravity in creating the first stars.

Answers

Chapter 1 Working scientifically

1.1 Questioning and predicting

1.1 Review

1 An inquiry question guides the investigation and ends in a question mark. A hypothesis is a cause-and-effect statement that can be tested through an investigation. The purpose is a statement outlining the aim of the investigation.

2 A

3 **a** independent variable: the surface area of the containers; dependent variable: time taken for the water to reach room temperature
b independent variable: the launch angle; dependent variable: the range of the projectile
c independent variable: the thickness of the foam bumper; dependent variable: force applied to come to a stop
d independent variable: the emf applied to the circuit; dependent variable: the total current

4 qualitative

5 B **6** A **7** A

1.2 Planning investigations

1.2 Review

1 **a** valid **b** reliable **c** accurate

2 B and D **3** B

4 **a** bumper foam density
b impact force
c cart mass, method of density measurement, equipment used to measure force, velocity of collision.

5 **i** Create a ramp that can be adjusted from 10° to 30° and place it on a table. Place two marks on the table, 20 cm apart. Start with the launcher at 10°.
ii Measure and record the height of the table.
iii Place some carbon paper face down on a piece of paper on the floor. Measure the distance between the edge of the table and the results paper.
iv Roll a ball down the ramp. Measure the time it takes for the ball to pass between the markers. Use this to calculate launch velocity. Ensure the ball lands on the carbon paper.
v Repeat step iv to get an accurate measurement of the launch velocity.
vi Increase the ramp angle by 5° and repeat steps iii to v.

1.3 Conducting investigations

1.3 Review

1 in a logbook

2 **a** Systematic error—due to improper calibration every recorded weight will be incorrect. Time should be taken to learn how to calibrate all equipment properly before starting an experiment.
b Mistake—the different masses were not correctly labelled or were incorrectly selected. All materials should be clearly labelled and double-checked before they are used in experiments.
c Random error—this result is an outlier. Repeating an experiment will make outliers obvious.
d Systematic error—due to poor definitions of the variables the results will be inconsistent. Variables should be clearly defined before starting an experiment.

3 0.02 s and 0.72 s are mistakes as they are significantly different from the other readings. The average time is 0.24 s. The mistakes are not included in the average.

4 **a** systematic error **b** mistake **c** random error

1.4 Processing data and information

1.4 Review

1 **a** 23 **b** 21 **c** 22

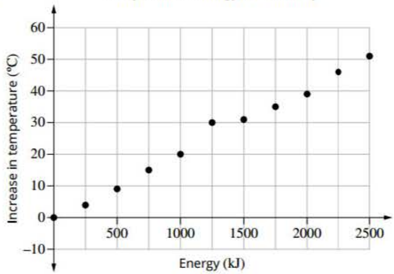
2 **a** 4 **b** the least number of significant figures
c the least number of decimal places

3 **a** *y*-axis **b** *x*-axis

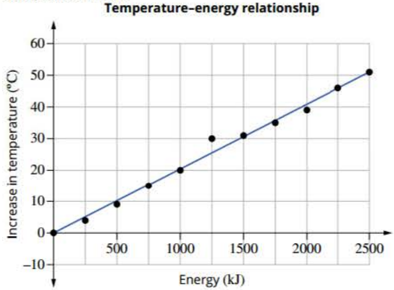
4 source A: 8, source B: 60

5 **a** The levels of carbon dioxide in the atmosphere slowly rose from 1805. After 1905 it increases at a higher rate.
Between 1805 and 1910, the average temperature decreased slightly. From 1910 onwards, the temperature has increased steadily, apart from in 1905, which had a slightly smaller increase.
b In general, both carbon dioxide and average global temperature have been rising from 1805.

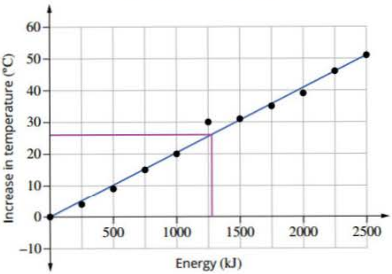
6 **a** **Temperature–energy relationship**



b Data that does not fit an observed pattern or trend.
c data point at 1250 kJ
d **Temperature–energy relationship**



e **Temperature–energy relationship**



Energy = 1280 kJ. If you are doing this by hand rather than using a spreadsheet program, there may be some variation in this value.

7 Include units on the *x*-axis (°C), include units on the *y*-axis (A), have a consistent scale on the *x*-axis, include a title, include scale on the *y*-axis.

8 **a** The length is least likely to be accurate or precise because it is difficult to take consistent strides and travel in a straight line.

b The measured length is likely to be more accurate than striding because a trundle wheel has a graduated scale that enables the length to be estimated to the nearest cm, but will not be precise as it is difficult to ensure the shortest distance is being measured.

c The length is most likely to be accurate because the tape measure has a graduated scale that measures to the cm, and it is more precise because it will measure in a straight line.

1.5 Analysing data and information

TY 1.5.1 -3.5×10^{-4} A

1.5 Review

1 bar graphs, histograms and pie charts

2 Five of: title, *y*-axis label, *y*-axis units, *y*-axis scale, *x*-axis label, *x*-axis scale, *x*-axis units, trend line.

3 The temperature increased at a steady rate for the first 4.5 min, reaching 42°C. From 4.5 min until 10 min, the temperature oscillates from 42°C to 38°C with a period of approximately 1.2 min.

4 E

5 qualified author, published in a reputable peer-reviewed journal, method detailed enough that the reader could replicate experiment, valid, reliable, accurate, precise, limitations outlined, assumptions outlined, includes suggested future improvements or research directions

1.6 Problem solving

1.6 Review

1 D

2 **a** Yes; $R = 1.2$ m

b No, because that is not within the range of experimental data. It is an extrapolation.

c No, that would be a generalisation that is beyond the scope of the experiment.

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

4 Five repeats of the procedure were conducted.

1.7 Communicating

1.7 Review

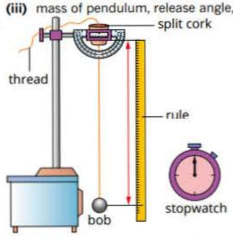
1 B

2 A, B and C

3 **a** **(i)** length of string **(ii)** period of motion

(iii) mass of pendulum, release angle, pendulum equipment

b



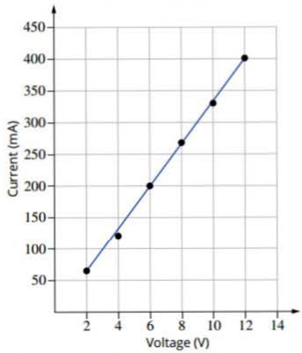
1. Set up equipment as shown. Record the mass of the bob and make a mark on the protractor at 10°.
2. Pull thread through cork until the length is 10 cm.
3. Keeping the thread taut, pull the bob back until the thread is in line with the angle marking on the protractor.
4. Release bob and start stopwatch. Measure the time for 10 periods and record it.
5. Repeat steps 3 and 4 for thread lengths 15 cm, 20 cm, 25 cm, 30 cm, 35 cm.

4 **a** 2.55×10^{5} **b** 4.32×10^{-7}

c It allows very large and very small numbers to be handled easily.

Chapter 1 Review

1 **The linear relationship between voltage and current**



2 B **3** C **4** A

5 An aim is a statement outlining the purpose of an investigation. A hypothesis is a cause-and-effect statement that can be tested through an investigation. The variables are the factors that can change in an investigation (the independent variable is the variable changed on purpose by the researcher, the dependent variable is the variable observed or measured in an investigation, controlled variables are all other variables that are maintained in the investigation).

6 **a** dependent **b** controlled **c** independent

7 independent variable: density of foam in shin pads; dependent variable: force on the ball; controlled variables: method of measurement, ball velocity, ball mass

8 **a** independent variable: material thickness of spring; dependent variable: spring coefficient
b independent variable: metal type; dependent variable: thermal expansion with a temperature increase of 50°C
c independent variable: mass; dependent variable: velocity and acceleration

9 $2 \times 10^{6}\,\mathrm{W\,m^{-2}\,K^{-1}}$ **10** 0.03000 L

11 **a** the average of a set of data
b the most frequent value in a set of data
c the middle value in a set of data

12 6.78 g ± 0.37 **13** the mean

14 **a** mistake **b** random error **c** systematic error

15 **a** For example, two significant figures: 0.032; three significant figures: 0.0302; four significant figures: 0.03020; five significant figures: 0.030200.
b Report the final calculation to the least number of significant figures of the measurement used in the calculation.
c Report the final calculation to the least number of decimal places of the measurement used in the calculation.

16 **a** accuracy: how close a measurement is to the true value; precision: how close measurements are to each other
b if more than one independent variable was changed at a time, if an inappropriate method was used, if outliers were included in data analysis, if an insufficient sample size was used
c many possible answers; for example, the experiment was only repeated three times

17 **a** reliability **b** accuracy **c** validity **d** precision

18 to give credit to others, to avoid plagiarism, to enable the reader to obtain further information

19 **a** To investigate the effect of temperature on the resistance of wire.
b independent variable: temperature of wire; dependent variable: resistance; controlled variables: wire diameter, wire material, wire length, measurement technique
c quantitative
d glass thermometer: low precision; non-contact thermometer: high precision; analog voltmeter and ammeter: low precision; depending on scale, voltage/current probe (data logger): high precision
e a graph that decreased with increasing temperature

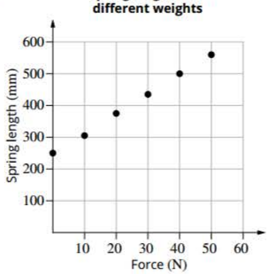
20 extension = 14 × mass

21 **a** $v = 3t - 8$
b The object started moving at $8\,\mathrm{m\,s^{-1}}$ in the negative direction. It then accelerated at a constant rate of $3\,\mathrm{m\,s^{-2}}$ in a positive direction.

22 The chart needs a title, and the *x*-axis needs a label. The trend line does not fit the data. The data is not linear.

23 **a**

Spring length with different weights



b length = 6.2 × force + 247.4
The spring starts with an unstretched length of 247.4 mm, and then stretches 6.2 mm for every N of force that is applied.

24 **a** independent variable: starting temperature, dependent variable: time to cool to room temperature, controlled variables: cup properties (surface area, insulation), coffee properties, room temperature, method of data collection
b Inquiry question: How does the starting temperature of a cup of black coffee, in a disposable paper cup 200 mL in volume, affect the time it takes to cool to room temperature?
Hypothesis: An increase in starting temperature will increase the time taken to reach room temperature exponentially.
c
1. Make 1 L of black coffee in a jug. Divide it into six different cups with 200 mL in each. Record the mass of each cup.
2. Heat each cup to a different starting temperature using a stove, water bath, or other heating element. Suggested temperatures are 50° C, 55° C, 60° C, 65° C, 70° C, 75° C.
3. Place each cup on the bench away from draughts, and at a distance from each other. Place a thermometer or temperature probe in each cup.
4. Record the starting temperature of each cup. Start a stopwatch. Record room temperature.
5. Record the temperature every minute. Record the time taken to reach room temperature.

d

	Cup 1	Cup 2	Cup 3	Cup 4	Cup 5	Cup 6
Mass (g)						
Time (s)						
Starting temperature (°C)						
1 min						
2 mins						

25 Student answers will vary depending on the cars chosen; see Fully Worked Solutions for an example.

26 Student responses will vary.

Chapter 2 Projectile motion

2.1 Projectiles launched horizontally

TY 2.1.1 a 4.0 s (to two significant figures) **b** $15\,\mathrm{m\,s^{-1}}$ down

TY 2.1.2 a 2.47 s (to three significant figures) **b** 49.5 m
c $31.4\,\mathrm{m\,s^{-1}}$ at 50.5° below the horizontal

2.1 Review

1 6.39 s **2** $32\,\mathrm{m\,s^{-1}}$ down **3** 20.4 m **4** B and C

5 **a** 1.0 s **b** 20 m **c** $9.80\,\mathrm{m\,s^{-1}}$ downwards **d** $21\,\mathrm{m\,s^{-1}}$
e $22\,\mathrm{m\,s^{-1}}$

2.2 Projectiles launched obliquely

TY 2.2.1 a $6.11\,\mathrm{m\,s^{-1}}$ horizontally to the right **b** 0.252 m
c 0.454 s

2.2 Review

1 B

2 **a** $14\,\mathrm{m\,s^{-1}}$ **b** $6.3\,\mathrm{m\,s^{-1}}$ **c** $9.80\,\mathrm{m\,s^{-2}}$ downwards
d $13.6\,\mathrm{m\,s^{-1}}$ horizontally

3 **a** $4.0\,\mathrm{m\,s^{-1}}$ **b** $6.9\,\mathrm{m\,s^{-1}}$ **c** 0.71 s **d** 3.9 m
e $4.0\,\mathrm{m\,s^{-1}}$ to the right

4 $24\,\mathrm{m\,s^{-1}}$

Chapter 2 Review

1 $17\,\mathrm{m\,s^{-1}}$ upwards **2** $12.3\,\mathrm{m\,s^{-1}}$ upwards

3 No. The squirrel reaches the ground after the nut.

4 B and D
5 **a** 0.49 s **b** 2.0 m (to two significant figures) **c** 9.80 m s^{-2} down
6 **a** 1.5 m **b** 7.4 m s^{-1} **c** 7.6 m s^{-1}
7 **a** 44.3 m s^{-1} **b** 60.6°
c 50.8 m s^{-1} at 60.6° to the horizontal
d The tourist would need to increase the initial speed by 1.6 m s^{-1}.
8 **a** 0.64 s **b** 0.64 s **c** 3.2 m
9 The hockey ball will travel further, because drag forces will have a greater effect on the horizontal velocity of the polystyrene ball.
10 **a** 54 m s^{-1} **b** 45°
11 **a** v_H = 10 m s^{-1} to the right, v_V = 4.4 m s^{-1} down
b 11 m s^{-1} at 24° to the horizontal
c 0.45 s **d** 4.5 m
12 **a** 2.5 m **b** 9.80 m s^{-2} vertically downwards
c 0.71 s **d** 7.4 m s^{-1}
13 No.
14 **a** 10 m s^{-1} **b** 12 m s^{-1} **c** 8.9 m
15 17.3 m s^{-1} **16** C
17 **a** **i** 23 m **ii** 22 m **iii** 22 m
b **i** Yes. By increasing or decreasing the angle away from 45° by the same amount, the displacement is equally reduced.
ii Students' answers may vary. An investigation will include changing the launch angle and measuring the distance travelled. The initial speed and the type of projectile need to be kept constant.
18 Both cannon balls travel the same distance.
19 **a** **i** 24.2 m s^{-1} **ii** 24.2 m s^{-1} **iii** 24.2 m s^{-1}
b **i** 14.0 m s^{-1} upwards **ii** 4.2 m s^{-1} upwards
iii 5.6 m s^{-1} downwards
c 24.8 m s^{-1} **d** 28 m s^{-1} **e** 40.0 m
20 **a** $t=\frac{s}{v}=\frac{20}{18\cos\theta}$ **b** $t=\frac{18\sin\theta}{4.9}$ **c** 19°
21 Student answers may vary; see Fully Worked solutions.

Chapter 3 Circular motion

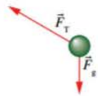
3.1 Circular motion

TY 3.1.1 7.5 km h^{-1} **TY 3.1.2** 300° s^{-1}
TY 3.1.3 **a** 520 m s^{-2} **b** 3.6 × 10^3 N

3.1 Review

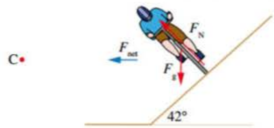
1 No. Phil's inertia made him stay where he was (stationary) as the tram moved forwards. This made it look like Phil was thrown backwards relative to the tram. This is an example of Newton's first law. Objects will remain at rest unless a net unbalanced force acts to change the motion.
2 Newton's third law ($\vec{F}_{AB} = -\vec{F}_{BA}$) says that every action has an equal and opposite reaction, and that action–reaction pairs must act on different bodies. The weight force and the normal force of an object both act on the same body, so they cannot be an action–reaction pair.
For a mug sitting on a table, the third law pairs for each will be:
Force on mug by table (F_N) with force on table by mug.
Gravitational force Earth exerts on mug (F_g) with gravitational force mug exerts on Earth.
3 B **4** 0.2 s
5 **a** A and D
b **i** 8.0 m s^{-1} **ii** 8.0 m s^{-1} south **iii** 7.0 m s^{-2} west
c 8.4 × 10^3 N west
d **i** 8.0 m s^{-1} north **ii** 7.0 m s^{-2} east
e The force needed to give the car a larger centripetal acceleration will eventually exceed the maximum frictional force that could act between the tyres and the road surface. At this time, the car would skid out of its circular path.
6 **a** 2.7 m s^{-2} (to two significant figures)
b Unbalanced, because the skater has an acceleration.
c 130 N (to two significant figures)
7 **a** 240° **b** 480° s^{-1} **c** 36 s

3.2 Circular motion on banked tracks

TY 3.2.1 **a** 1.53 m
b


There are two forces acting—the tension in the cord, $\vec{F}_T$, and gravity, $\vec{F}_g$.
c 2.34 N towards the left
d 3.05 N
TY 3.2.2 **a** 590 N towards the centre of the circle
b 17 m s^{-1}

3.2 Review

1 **a** 1.2 m
b The forces are her weight acting vertically downwards and the tension in the rope acting along the rope towards the top of the maypole.
c towards point B, the centre of her circular path
d 170 N towards B **e** 2.6 m s^{-1}
2 towards the centre of the circle
3 The architect could make the bank angle larger or increase the radius of the track.
4 Higher, as the greater speed means that a greater radius is required in the circular path. When travelling faster than the design speed the normal force is not sufficient to keep the car moving in a circle and causes the car to move outwards from the centre.
5 friction, normal, weight, balanced, normal, weight
6


7 47°
8 **a** **i** 4.9 kN **ii** 22°
b The driver would have to turn the front wheels slightly towards the bottom of the bank.

3.3 Work and energy

TY 3.3.1 **a** 4.85 m s^{-1} **b** 15.6 N up **c** 3.70 m s^{-1} **d** 6.76 N down

3.3 Review

1 0 J. There is no change in energy so there is no mechanical work.
2 1.7 N up
3 **a** 1.8 N down **b** 6.0 m s^{-1} **c** 4.9 N up
4 **a** 31.4 m s^{-1} **b** 19.9 m s^{-1} **c** 8300 N down **d** 12.1 m s^{-1}
5 **a** 18 m s^{-2} up **b** 1530 N up
6 **a** 9.8 m s^{-2} down **b** 2.2 m s^{-1}
7 2.8 m s^{-1}
8 **a** 2.8 m s^{-1} **b** 4.4 N towards the centre

3.4 Torque

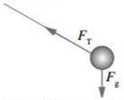
TY 3.4.1 39.5 N m anticlockwise **TY 3.4.2** 17.1 N m

3.4 Review

1 a The magnitude of the torque produced by a given force is proportional to the length of the force arm. By pushing the door at the handle, rather than the middle, the length to the force arm is increased.
  b A crowbar can be used to generate a large torque because the force can be applied at a large distance from the pivot.
2 a 200 N m anticlockwise b 0 N m
3 0.5 m (perpendicular to the force)
4 18 N (perpendicular to the force arm)
5 90 N m
6 a 4.9 N m b 9.8 N m c 4.9 N m
7 45 N m
8 a 3.4×10^4 N down b It remains the same.
  c 5.2×10^5 N m clockwise about the pivot

Chapter 3 Review

1 a $3.70\,m\,s^{-1}$ b $17.1\,m\,s^{-2}$ towards the centre of the circle
  c 0.430 N
2 a



  b 0.49 N
3 a $2.5\,m\,s^{-2}$ towards the centre of the circle
  b The centripetal force is created by the friction between the tyres and the ground.
4 50.0 Hz 5 $15.7\,m\,s^{-1}$
6 a $48\,km\,h^{-1}$ b 640 N
  c It is greater than the normal force on a flat track; in that case F_N is $mg = 539$ N.
7 $360°\,s^{-1}$ 8 4.1 Hz
9 a 28 s b 5.0 N
10 a 0.5 s b $10\,m\,s^{-1}$ c $130\,m\,s^{-2}$ (to two significant figures)
  d 320 N (to two significant figures)
11 a The centripetal acceleration would increase by a factor of four.
  b The centripetal acceleration would be one third of the original.
  c The centripetal acceleration would remain the same.
12 $3.40 \times 10^{-2}\,m\,s^{-1}$
13 a $10\,m\,s^{-1}$ south b $10\,m\,s^{-1}$ c 13 s d $5.0\,m\,s^{-2}$ west
  e 7.5×10^3 N west
14 0.146 m 15 A
16 a $1.02 \times 10^3\,m\,s^{-1}$ b $1.53 \times 10^{-4}\,°\,s^{-1}$ c 1.99×10^{20} N
17 5.8 kg
18 42 kg, 410 N (both to two significant figures)
19 0.38 m
20 Objects undergo circular motion when the net force acting on the object is radially inward with a magnitude of $F_{net} = F_c = \frac{mv^2}{r}$. See Fully Worked Solutions for more details.

Chapter 4 Motion in gravitational fields

4.1 Gravity

TY 4.1.1 7.1×10^{-9} N towards one another

TY 4.1.2 2.0×10^{20} N between the Earth and the Moon

TY 4.1.3 $a_{Earth} = 6.0 \times 10^{-3}\,N\,kg^{-1}$

$a_{Sun} = 1.8 \times 10^{-8}\,N\,kg^{-1}$

$\frac{a_{Sun}}{a_{Earth}} = 3.3 \times 10^5$

The acceleration of the Earth is 3.3×10^5 times greater than the acceleration of the Sun.

TY 4.1.4 $9.75\,N\,kg^{-1}$ towards the centre of the Earth

TY 4.1.5 $3.73\,N\,kg^{-1}$ towards the centre of Mars

4.1 Review

1 The force of attraction between any two bodies in the universe is directly proportional to the product of their masses and inversely proportional to the square of the distance between them.
2 1.8×10^{21} N
3 $2.8 \times 10^{-3}\,m\,s^{-2}$ (towards the Sun)
4 a 3.0×10^{16} N (attractive) b 3.4×10^{22} N (attractive)
  c 0.000088%
5 $\frac{1}{9}$ of the original
6 $8 \times 10^{-4}\,N\,kg^{-1}$ (towards the centre of 67P/Churyumov–Gerasimenko)
7 Mercury: $g = 3.7\,N\,kg^{-1}$
  Saturn: $g = 10.4\,N\,kg^{-1}$
  Jupiter: $g = 24.8\,N\,kg^{-1}$

4.2 Satellite motion

TY 4.2.1 $3.08 \times 10^3\,m\,s^{-1}$

TY 4.2.2 **a** 6.70×10^5 km **b** 1.90×10^{27} kg **c** $8.20\,km\,s^{-1}$

4.2 Review

1 C 2 B
3 a $0.22\,m\,s^{-2}$
  b 510 N, towards the Earth (to two significant figures)
4 15.6 days
5 All of these quantities are given as scalars, the direction for g will always be towards the centre of mass.
  a $5.67\,N\,kg^{-1}$ b $1.48\,N\,kg^{-1}$ c $0.564\,N\,kg^{-1}$ d $0.224\,N\,kg^{-1}$

4.3 Gravitational potential energy

TY 4.3.1 -5.5×10^8 J **TY 4.3.2** **a** -2.40×10^9 J **b** $7570\,m\,s^{-1}$

TY 4.3.3 -1×10^9 J **TY 4.3.4** $2.37 \times 10^3\,m\,s^{-1}$

4.3 Review

1 C
2 a g increases from point A to point D.
  b It will accelerate at an increasing rate as it approaches the Earth.
  c A, B and C
3 -1.85×10^{14} J 4 1.8×10^9 J 5 $5.03\,km\,s^{-1}$

Chapter 4 Review

1 730 N (attractive) 2 3.78×10^8 m 3 $2.1 \times 10^{-7}\,m\,s^{-2}$
4 The Moon has a smaller mass than the Earth and therefore experiences a larger acceleration from the same gravitational force.
5 $3.7\,m\,s^{-2}$ towards the centre of Mars
6 a 2.48×10^4 N b 2.48×10^4 N c $24.8\,m\,s^{-2}$
  d $1.31 \times 10^{-23}\,m\,s^{-2}$
7 C 8 240 N
9 a D b B c B
10 The direction of the arrowhead indicates the *direction* of the gravitational force and the space between the arrows indicates the *magnitude* of the field. The field lines point towards the sources of the field.
11 $9.3\,N\,kg^{-1}$
12 a $9.79\,N\,kg^{-1}$
  b The gravitational field is 0.7% stronger at the North Pole than the equator.
13 16 14 $N\,kg^{-1}$ 15 $2 \times 10^{12}\,N\,kg^{-1}$
16 The gravitational field strength at the poles is 1.4 times that at the equator.
17 10 Earth radii 18 B
19 Europa: $T = 85.2$ h
  Ganymede: $T = 172$ h
  Callisto: $T = 401$ h
20 $0.62\,N\,kg^{-1}$
21 a $5580\,m\,s^{-1}$ b 5.69×10^{26} kg
22 a -6.2×10^7 J b -3.1×10^7 J c 2.0×10^7 m
23 2.1×10^{11} J

24 **a** $2.45\,N\,kg^{-1}$ **b** $-1.5 \times 10^8\,J$ **c** $-3 \times 10^7\,J$
25 $4 \times 10^9\,J$ **26** $4.25\,km\,s^{-1}$
27 See Fully Worked Solutions.

Module 5 Review

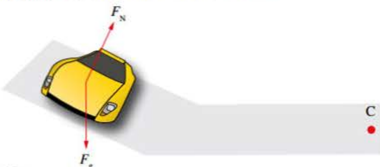
Advanced mechanics

Multiple choice

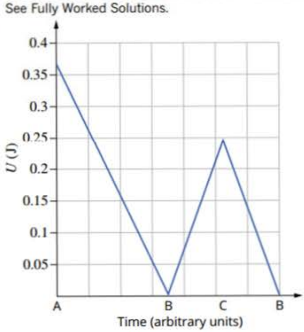
1 C	**2** C	**3** C	**4** D
5 A	**6** B	**7** B	**8** D
9 C	**10** D	**11** A	**12** C
13 B	**14** D	**15** C	**16** B
17 A	**18** C	**19** B	**20** C

Short answer

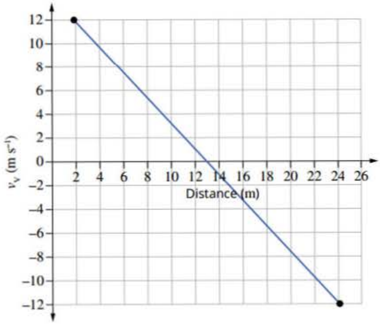
21 $2.5 \times 10^{-7}\,N$
22 Thomas should sit 4.6 m from the pivot opposite his siblings.
23 10 Earth radii
24 $3.5\,m\,s^{-2}$
25 The gravitational field strength at the poles is 1.4 times that at the equator.
26 $18\,m\,s^{-1}$ at 42° from the horizontal
27 **a** $1.06 \times 10^5\,N$ **b** $5.5 \times 10^3\,s$ **c** It has no effect.
28 **a** 4.9 m **b** $9.80\,m\,s^{-2}$ down **c** $10.6\,m\,s^{-1}$
29 **a** $18\,m\,s^{-2}$ up **b** $1.5 \times 10^3\,N$
c The apparent weight is given by the normal force of $1.5 \times 10^3\,N$. This is almost three times larger than the weight force and so the skater would feel much heavier than usual.
30 **a**



b 47°
31 **a** unbalanced, balanced **b** $12.6\,m\,s^{-1}$ **c** $31.8\,m\,s^{-2}$
d $1.9 \times 10^3\,N$ **e** 0.6 Hz
32 **a** $9.80\,m\,s^{-2}$ down **b** $2.2\,m\,s^{-1}$ **c** 0 **d** $4.9\,m\,s^{-1}$
e See Fully Worked Solutions.
f



33 **a** $4.25 \times 10^{10}\,J$ **b** $3.3 \times 10^3\,m\,s^{-1}$
c **i** $4.0 \times 10^4\,N$ (from graph) **ii** $8.1 \times 10^4\,N$ (from graph)
d It increases from $8.1\,m\,s^{-2}$ to $9.2\,m\,s^{-2}$.
e $3.4 \times 10^8\,m$
34 **a** 60 N **b** $3.6 \times 10^7\,J$
c Determine the energy associated with each grid square by multiplying each area by the mass of 20 kg. Calculate the altitude at which the total area starting from zero height is equal to 40 MJ.
d On Earth, weight is the gravitational force acting on an object near the Earth's surface whereas apparent weight is the contact force between the object and the Earth's surface. See Fully Worked Solutions for more details.
35 **a** 9 m **b** $6.0\,m\,s^{-1}$ **c** 68° **d** 11 m
e



Yes, the ball will reach the goal posts.

Chapter 5 Charged particles, conductors, and electric and magnetic fields

5.1 Particles in electric fields

TY 5.1.1 $9 \times 10^5\,m\,s^{-2}$ (to one significant figure)
TY 5.1.2 $2.16 \times 10^{-18}\,J$ of work is done on the field by the charge
TY 5.1.3 **a** $3.3 \times 10^7\,m\,s^{-1}$
b $2.6 \times 10^{16}\,m\,s^{-2}$ (towards the earthed plate)
c $8.8 \times 10^{-10}\,s$ **d** 1.8 cm to the right

5.1 Review

1 C **2** B
3 **a** True **b** False **c** False **d** True **e** True
f False **g** False
4 $1.25 \times 10^{-2}\,N$
5 $5.72 \times 10^{11}\,m\,s^{-2}$ (in the opposite direction to the field)
6 1200 V
7 *radial, static, non-uniform*
8 **a** $7.3 \times 10^{16}\,m\,s^{-2}$, in the opposite direction to the field
b $4.0 \times 10^{-10}\,s$

5.2 Particles in magnetic fields

TY 5.2.1 $4.8 \times 10^{-22}\,N$ **TY 5.2.2** into the screen
TY 5.2.3 **a** $1.3 \times 10^5\,V\,m^{-1}$ **b** $9.4 \times 10^7\,m\,s^{-1}$ **c** $1.8 \times 10^{-3}\,m$

5.2 Review

1 B **2** 2.4×10^{-24} south
3 **a** $9.6 \times 10^{-15}\,N$ **b** $4.6 \times 10^{-3}\,m$
4 A charged particle in a magnetic field will experience a force ($F = qv_{\perp}B$). As force ∝ velocity, the force will increase as the velocity increases. This will continue while the charge remains in the magnetic field, continuously accelerating the charge.

5 D

6 **a** south (S) **b** C **c** remains constant **d** A

e Particles with no charge, e.g. neutrons.

7 D **8** 0 N

Chapter 5 Review

1 0.0225 N **2** C

3 The electrical potential is defined as the work done per unit charge to move a charge from infinity to a point in the electric field. The electrical potential at infinity is defined as zero. When you have two points in an electric field (E) separated by a distance (d) that is parallel to the field, the potential difference V is then defined as the change in the electrical potential between these two points.

4 25V **5** C

6 *field, charged particle* **7** 4.2×10^{-18} J **8** 9.6×10^{-15} N

9 2.0×10^{-14} N

10 **a** work done by the field **b** no work is done

c work done on the field **d** no work is done

e work done on the field **f** work done by the field

11 **a** 1.09×10^{-19} J **b** Work was done on the field.

12 5.42×10^{4} m s^{-1} **13** $+1.63 \times 10^{-4}$ C **14** C

15 1.4×10^{-22} N **16** 2.0×10^{-7} T

17 **a** The electron will experience a force at right angles to its motion. This acts upwards in the initial moment and causes the electron to curve in an upwards arc from its starting position.

b The electron's velocity and the magnitude of the magnetic field

18 **a** 1.4×10^{4} V m^{-1} **b** 9.3×10^{6} m s^{-1}

19 4.8×10^{-23} N

20 See Fully Worked Solutions

Chapter 6 The motor effect

6.1 Force on a conductor

TY 6.1.1 2.5×10^{-3} N per metre of power line

TY 6.1.2 a 0 N **b** 2.25×10^{-3} N outwards **c** 9.4×10^{-5} N

6.1 Review

1 D **2** D **3** 0.40 N up

4 1.88×10^{-4} N away from the house

5 2.0×10^{-4} N north

6 **a** 0.18 N downwards **b** the same as before

6.2 Forces between conductors

TY 6.2.1 2.5×10^{-5} N m^{-1} attractive

TY 6.2.2 1.2×10^{-3} N m^{-1} repulsive

6.2 Review

1 attractive **2** C **3** 5.0×10^{-4} N m^{-1} repulsive

4 2.0×10^{-4} N m^{-1} attractive **5** 10 A

Chapter 6 Review

1 D **2** **a** attractive **b** repulsive

3 $\frac{1}{8}F$ **4** **a** palm **b** fingers **c** thumb

5 2.78 A

6 **a** 5.0×10^{-9} N out of the page **b** 2.0×10^{-3} N out of the page

7 An east–west line, as it runs perpendicular to the Earth's magnetic field.

8 C

9 perpendicular to the magnetic field

10 1.6 N upwards **11** 7.5×10^{-4} N

12 0.2 A

13 **a** 4.0×10^{-4} N north **b** 1.2 N north **c** 1.0 A

14 7.5×10^{-4} T **15** 0.6 N downwards

16 B

17 1.2×10^{-4} N m^{-1} repulsive

18 9.0×10^{-5} N m^{-1} attractive

19 15 A

20 See Fully Worked Solutions for derivation.

21 See Fully Worked Solutions.

Chapter 7 Electromagnetic induction

7.1 Magnetic flux

TY 7.1.1 8.0×10^{-5} Wb **TY 7.1.2** 1.6×10^{-4} Wb or 0.16 mWb

7.1 Review

1 A **2** 0 Wb **3** 3.2×10^{-6} Wb

4 The magnetic flux decreases from 3.2×10^{-6} Wb to 0 after one-quarter of a turn. Then it increases again to 3.2×10^{-6} Wb through the opposite side of the loop after half a turn. Then it decreases to 0 again after three-quarters of a turn. After a full turn it is back to 3.2×10^{-6} Wb again.

5 1.3×10^{-5} Wb **6** 0.25 mWb **7** 5.4×10^{-6} Wb

7.2 Faraday's and Lenz's laws

TY 7.2.1 a 5.0×10^{-4} Wb **b** 5.0×10^{-3} V **TY 7.2.2** 1000 turns

TY 7.2.3 clockwise when viewed from above

TY 7.2.4 (i) through the solenoid from Y to X (through the meter from X to Y)

(ii) There will be no induced emf or current in the solenoid.

(iii) through the solenoid from X to Y (through the meter from Y to X)

TY 7.2.5 anticlockwise

7.2 Review

1 1.2×10^{-6} Wb **2** 0 Wb **3** 3.0×10^{-5} V

4 C **5** 4×10^{-3} V

6 The effect of using multiple coils is similar to placing cells in series—the emf of each of the coils adds together to produce the total emf.

$\Delta\Phi = 2$ V

7 C

8 **a** A **b** A

7.3 Transformers

TY 7.3.1 4000 turns **TY 7.3.2** 0.013 A **TY 7.3.3** 3 W

TY 7.3.4 3.6×10^{5} W or 0.36 MW **TY 7.3.5** 500.6 kV

7.3 Review

1 B **2** D **3** 40 turns

4 **a** $P_1 = P_2$ **b** $\frac{I_1}{I_2} = \frac{N_2}{N_1}$

5 400 W **6** 4×10^{7} W or 40 MW

7 **a** 5000 A **b** 90 kV

8 B. A is incorrect because the ΔV in the formula indicates the voltage drop in the transmission lines; it does not refer to the voltage being transmitted.

Chapter 7 Review

1 **a** 3.2×10^{-3} V or 3.2 mV **b** clockwise

2 **a** 0.04 V **b** from Y to X

3 from X to Y

4 No current flows in S_2 between $t = 1$ s and $t = 4$ s. An increase in emf at a constant rate ($t = 0$ to $t = 1$ s) would produce a constant current, and a decrease in emf at a lower rate ($t = 4$ to $t = 7$ s) would produce a lower current in the opposite direction.

5 1.0 A **6** 10 **7** A **8** 8 V

9 It will be doubled to 16 V.

10 Any two of:
1. Using a DC power supply means that the voltage cannot be stepped up or down with transformers.
2. There will be significant power loss along the 8 Ω power lines.
3. Damage to any appliances operated in the shed that are designed to operate on 240 V AC and not on 240 V DC.

11 clockwise
12 AB and CD
13 15 A **14** 9970 V **15** 450 W
16 Without the first transformer, the power supplied = 105 kW. This represents a 30% power loss—bad idea!
17 anticlockwise
18 She must change the magnetic flux through the coil by changing the strength of the magnetic field (by changing the position of the magnet relative to the coil) or by changing the area of the coil (by changing the shape of the coil or by rotating the coil relative to the magnetic field). See Fully Worked Solutions for sample calculations.
19 0.010 m^2 **20** 0.125 s
21 See Fully Worked Solutions.

Chapter 8 Applications of the motor effect

8.1 Motors

TY 8.1.1 2.9×10^{-7} N m, clockwise as viewed from side 3
TY 8.1.2 5.41×10^{-7} N m

8.1 Review

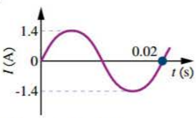
1 A
2 The torque is acting clockwise; see Fully Worked Solutions for diagram.
3 **a** 1.0×10^{-2} N into the page **b** 1.0×10^{-2} N out of the page **c** 0 N **d** anticlockwise **e** D **f** 2.0×10^{-4} N m
4 In an AC motor, the stator creates a rotating magnetic field that induces a current in the conductors of the rotor. Through application of Faraday's and Lenz's laws, the rotating magnetic field of the stator effectively pulls the rotor.

8.2 Generators

TY 8.2.1 2000 W

8.2 Review

1 B
2



3 30 W **4** 3.54 A
5 The resulting output of all three phases maintains an emf near the maximum voltage more continuously.
6 90 W

Chapter 8 Review

1 **a** down the page **b** up the page
2 anticlockwise
3 **a** down the page **b** up the page
c zero, as the forces are parallel to the coil (trying to pull the coil apart rather than turn it)
4 C
5 To reverse the current direction in the coil every half turn to keep the coil rotating in the same direction.
6 0.1 N
7 Current flows into brush P and around the coil from V to X to Y to W. The force on side VX is down and the force on side YW is up, so rotation is anticlockwise.
8 D
9 **a** 18 V **b** 375 W
10 D **11** 2.0×10^{-6} N m **12** 1.0×10^{-5} N m
13 1.2×10^{-5} N m
14 To prevent the coils in the rotor from becoming tangled as the rotor rotates, and to reverse the current at the point where the coil is perpendicular to the magnetic field.
15 Three-phase generators provide a more constant maximum voltage than a single-phase generator. In addition, each of the three phases can be distributed to different loads, allowing for a more balanced distribution. For a relatively small additional cost (two additional conductors), three times the power can be delivered for a three-phase system compared with a single-phase system.
16 14 A
17 Using multiple armature windings can produce a steadier DC output voltage (see Figure 8.2.5b of the Student Book).
18 4 kW
19 They are the electrical contacts to the split-ring commutator.
20 It is the result of current produced in response to the rotation of the rotor inside the motor in the presence of an external magnetic field. The back emf, following Lenz's law, opposes the change in magnetic flux that created it, so this induced emf will be in the opposite direction to the emf creating it. The net emf used by the motor is thus always less than the supplied voltage.
21 See Fully Worked Solutions.

Module 6 Review

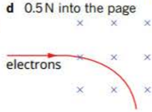
Electromagnetism

Multiple choice

1	C	**2**	C	**3**	B	**4**	A
5	C	**6**	A	**7**	C	**8**	B
9	C	**10**	A	**11**	D	**12**	C
13	B	**14**	A	**15**	B	**16**	D
17	A	**18**	C	**19**	C	**20**	D

Short answer

21 6.9×10^9 N C^{-1} to the left (away from the charge)
22 **a** 0.05 N **b** to the right **c** 0.01 N **d** to the left
23 **a** 9.9×10^7 m s^{-1} **b** 1.4×10^5 V m^{-1}
24 **a** 0 N **b** 0 N **c** 0.5 N out of the page
d 0.5 N into the page
25

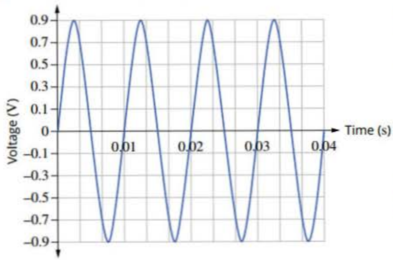


26 5.8×10^{-2} m
27 **a** 5×10^{-6} Wb **b** 0 Wb **c** 2.5×10^{-3} V **d** 1.25×10^{-3} A
e No. Once the loop is stationary, there is no change in flux and therefore no emf generated and no current flows in the loop.
28 **a** 1×10^{-3} N **b** from west to east **c** 4.9×10^4 A
d 2×10^{-3} N down
e The horizontal component of the current is now less and so there will be a smaller force per metre of cable.
29 **a** 0.4 A **b** 6000 V **c** 200 turns **d** 850 W **e** 1700 W
30 **a** 0 Wb
b Rotate the loop or the magnetic field so they are no longer parallel.
c Make them perpendicular (at right angles) to each other.
d $\Phi = B_{\perp}A$
$= 0.50 \times 0.2 \times 0.1$
$= 0.01$ Wb or 10^{-2} Wb

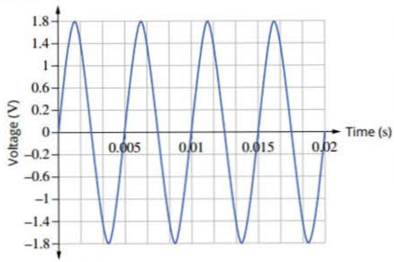
31 **a** AB: upwards; CD: downwards
b In the position shown (with the coil horizontal)
c When the coil is in the vertical position. It continues to rotate because: (i) Its momentum will carry it past the true vertical position. (ii) At the vertical position the commutator reverses the direction of the current through the coil and so the forces reverse, thus it continues to rotate for another half turn, at which point the current reverses again and the rotation continues.
d 4.0 A **e** 20 N **f** 64 N

32 **a** 500 Hz **b** 20 V **c** 7.1 V **d** 0.71 A **e** 5 W
f An alternator has a pair of slip-rings instead of a split-ring commutator.
g AC is generated in the coils of an alternator. Each slip ring connects to each end of the coil. The slip rings maintain the AC generated in the coil at the output.

33 **a**



b 0.64 V
c



d 50 Hz **e** 200 V

34 **a** from Y to X **b** 4×10^{-3} V **c** 8×10^{-3} A
d 3.2×10^{-5} W
e The external force that is moving the loop into the magnetic field.
f Zero. The loop will be totally within the magnetic field, hence there is no flux change and therefore no emf induced.
g Down through the loop; see Fully Worked Solutions for details.

35 **a** With little or no current in the power line there was almost no voltage drop. When the house appliances were turned on, there was a higher current in the power line and hence a voltage drop along the line, leaving a low voltage at the house.
b 218 V, 3488 W
c At the generator end, a 1:20 step-up transformer is required. There will be 20 times as many turns in the secondary as in the primary. At the house end, a 20:1 step-down transformer is required.
d 0.8 A **e** 1.6 V **f** 1.28 W **g** 249.92 V
h 3998.72 W
i Power loss without transformers: 12.8% of power generated; power loss with transformers: 0.03% of power generated.
j The power loss in the power line depends on the square of the current ($P = I^2R$). Since the current was reduced by a factor of 20 and the resistance remained constant, the power loss decreased by a factor of 20^2 or 400.

Chapter 9 Electromagnetic spectrum

9.1 Electromagnetism

TY 9.1.1 5.00×10^{14} Hz

9.1 Review

1 B 2 D 3 D
4 FM radio waves/infrared radiation/visible light/X-rays
5 **a** 4.57×10^{14} Hz **b** 5.09×10^{14} Hz

9.2 Spectroscopy

9.2 Review

1 If the element is given sufficient energy to excite its atoms, the energy is released as light, which forms the emission spectrum.
2 The lines have been blueshifted and they have been spread out. This means the object is moving towards Earth and it is rotating. The rotation of the object means that each side will be slightly redshifted or blueshifted, causing each of the spectral lines to spread out.
3 3, 2, 1

Chapter 9 Review

1 0.069% 2 B
3 **a** 600 nm **b** yellow
4 4.3 m 5 1.5×10^{18} Hz 6 0.33 m
7 The missing bands in the absorption band match the bands present in the emission spectrum.
8 **a** 6.17×10^{14} Hz **b** 7.56×10^{14} Hz
9 To generate electromagnetic radiation, a charge needs to oscillate. Wood is not conductive, therefore there are no charges in the wood that can move, or be induced to move.
10 Star B is 1.13 times hotter than star A.
11 UV and X-rays have shorter wavelengths (i.e. higher energy photons) and can penetrate skin. These photons have sufficient energy to disrupt biological molecules. Radio waves do not have sufficient energy to affect biological molecules.
12 Infrared radiation is absorbed as it travels through the Earth's atmosphere. More infrared radiation from the sources being studied can therefore be detected by a telescope operated outside the atmosphere, providing better measurements. As measurements are also being made in a cooler region than on Earth's surface, the measurements are less subject to interference from ground-based heat sources, again making the measurements more accurate.
13 From the data, $c = 2.9 \times 10^8\,\text{m s}^{-1}$
The theoretical value for the speed of light is $c = 3.00 \times 10^8\,\text{m s}^{-1}$. The value calculated in this experiment is slightly less than, but within 3% of, the theoretical value.
14 Emission spectra are produced when atoms are excited and release photons of particular wavelengths as they return to the ground state. Absorption spectra are produced when atoms absorb particular wavelengths as light passes through them. See Fully Worked Solutions for extension answer.
15 red, orange, yellow, blue
16 It is a young star as it has not had time to form heavier elements via nucleosynthesis.

17 The visible light can be broken up with a spectrograph, and the spectrum is compared to known spectra of other stars. Stars of the A class will have absorption spectra showing strong hydrogen lines. The temperature of the star can also be determined using Wien's law as stars will produce a specific radiation curve at a specific temperature—in this case in the range 7500–10000K. The colour of the star is related to its temperature.

18 According to Wien's law, the temperature is inversely proportional to the peak wavelength of the spectrum. Therefore, as star A has a lower peak wavelength (approximately 250 nm) than star B (at approximately 350 nm), star A must be hotter.

19 See Fully Worked Solutions.

Chapter 10 Light: wave model

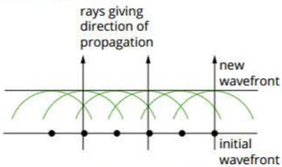
10.1 Diffraction and interference

TY 10.1.1

TY 10.1.2 550 nm

10.1 Review

1 The new wavefront should be a straight line across the front of the secondary wavelets.



2 D 3 A and D 4 460 nm

10.2 Polarisation

TY 10.2.1 25% **TY 10.2.2** 23 cd

10.2 Review

1 Polarisation occurs when transverse waves are allowed to vibrate in only one direction. Polarisation can only occur in transverse waves and cannot occur in longitudinal waves. Since light can be polarised, it must be a transverse wave.

2 12% 3 20 cd 4 18°

Chapter 10 Review

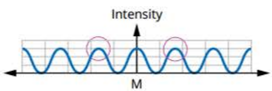
1 **a** wave model **b** wave model **c** particle model

2 C 3 A

4 the fringes would be more widely space, with a wider overall pattern

5 D 6 D

7 The central antinode occurs where both waves have travelled the same distance, i.e. the path difference is 0. The next antinode is on either side occurs when the path difference is 1λ.



8 Young's experiment demonstrated interference patterns, which are characteristic of waves. This led to scientists abandoning the prevailing particle theory and supporting a wave model of light.

9 **a** 0.44° **b** 580 nm

10 1.4° 11 61 µm

12 Young shone monochromatic light on a pair of narrow slits. Light passed through the slits and formed a pattern of bright and dark bands (or fringes) on a screen. Young compared this to interference patterns he had observed in water waves, and he identified that these lines corresponded to regions of constructive and destructive interference. This could only be explained by considering light to be a wave.

13 **a** increase **b** increase

14 The light reflected from water and snow is partially polarised. Snowboarders and sailors are likely to wear polarising sunglasses as these will absorb the polarised light reflected from the snow or water.

15 $I_2 = 41\%$ of I_1

16 1.0 cd 17 63°

18 The wavelength of light waves is very small; there are not many natural structures that are small enough to cause diffraction of light waves.

19 0.64°

20 See Fully Worked Solutions.

Chapter 11 Light: quantum model

11.1 Black-body radiation

TY 11.1.1 32000 K **TY 11.1.2** 9.66 µm **TY 11.1.3** 2.4×10^{-19} J

TY 11.1.4 1.5 eV **TY 11.1.5** 1.5 eV

11.1 Review

1 Maxwell's theory permits a continuous range of frequencies that can be emitted by a hot object. Planck proposed that light is radiated in discrete, quantised amounts or energy, rather than in a continuous unbroken wave.

2 3620 K 3 4140 K 4 322 nm

5 **a** 1.89 eV **b** 2.11 eV **c** 2.56 eV **d** 3.13 eV

11.2 The photoelectric effect

TY 11.2.1 8.0×10^{-19} J = 5.0 eV **TY 11.2.2** 2.1 eV

11.2 Review

1 In the photoelectric effect, a metal surface may become positively charged if light shining on it causes electrons to be released.

2 **a** True.
b False: When light sources of the same intensity but different frequencies are used, the higher frequency light has a higher stopping voltage, but it produces the same maximum current as the lower frequency.
c True.

3 **a** 4.1 eV **b** 4.6 eV **c** 6.2 eV

4 D

5 0.066 eV

Chapter 11 Review

1 Particle model: black-body radiation and photoelectric effect. Wave model: interference patterns and polarisation.

2 5770 K 3 6370 K 4 Rigel 5 263 nm

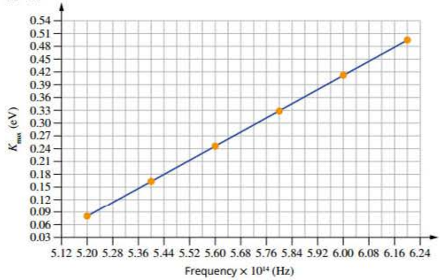
6 5.3×10^{-19} J 7 3.98×10^{-19} J = 2.49 eV 8 2.5 eV

9 8.0×10^{-19} J 10 photoelectrons 11 1.2×10^{15} Hz

12 2.9 eV 13 1.95 eV

14 Rb = 2.1 eV, Sr = 2.5 eV, Mg = 3.4 eV, W = 4.5 eV

15 **a**



b 4.1×10^{-15} eVs **c** 5.0×10^{14} Hz

d No. The frequency of red light is below the threshold frequency for rubidium.

16 **a** True.

b False: The stopping voltage is reached when the photocurrent is reduced completely to zero.

c True. **d** True.

17 C and D **18** 0.25 eV **19** 1.7 eV

20 See Fully Worked Solutions.

Chapter 12 Light and special relativity

12.1 Einstein's postulates

12.1 Review

1 D **2** A and D

3 The example should describe the movement of an object which highlights the non-inertial reference frame, i.e. a hanging pendulum in the spaceship will move from its normal vertical position when the spaceship accelerates.

4 The speed of the ball is greater for Jana than it is for Tom. The speed of the sound is greater forwards than it is backwards for Jana, while for Tom it is the same forwards and backwards. The speed of light is the same for Jana and Tom.

5 **a** $370\,\text{m s}^{-1}$ **b** $300\,\text{m s}^{-1}$ **c** $360\,\text{m s}^{-1}$ **d** $340\,\text{m s}^{-1}$

6 A

7 **a** $15\,\text{m s}^{-1}$ backwards **b** 3 m backwards **c** 0.2 s

8 **a** 0.1 s **b** $50\,\text{m s}^{-1}$ in all frames **c** 1 m

d $50\,\text{m s}^{-1}$ as always **e** 0.08 s

12.2 Evidence for special relativity

TY 12.2.1 520 s **TY 12.2.2** 3.90 m **TY 12.2.3** 20.5 m

12.2 Review

1 light, oscillation, time, constant

2 'Proper time' is the time measured at rest with respect to the event. Proper times are always less than any other times.

3 1.29 s **4** 48.2 s **5** 1.15 s

6 The equator clock is moving faster relative to the poles. It is also accelerating and hence will run slower. The effect is well below what we can detect as the speed of the equator is 'only' about $460\,\text{m s}^{-1}$, which is about 1.5 millionths of c.

7 The length that a stationary observer measures in their own frame of reference. That is, the object (or distance) that is being measured is at rest with the observer.

8 0.812 m (to three significant figures)

12.3 Momentum and energy

TY 12.3.1 **a** $1.56 \times 10^{-21}\,\text{kg m s}^{-1}$ **b** 0.998c

TY 12.3.2 **a** 4.8×10^{-12} J **b** 5.3×10^{-29} kg

12.3 Review

1 $9.53 \times 10^{5}\,\text{kg m s}^{-1}$ **2** $9.65 \times 10^{-18}\,\text{kg m s}^{-1}$

3 $1.59 \times 10^{-23}\,\text{kg m s}^{-1}$ **4** 5.67×10^{14} J

5 3.11×10^{14} J **6** B **7** 3.59×10^{19} J

8 3.11×10^{31} J

Chapter 12 Review

1 No object can travel at or beyond the speed of light, so the value of $\frac{v^2}{c^2}$ will always be less than 1. The number under the square root sign will also, therefore, be a positive number less than one. The square root of a positive number less than one will always be less than one as well.

2 1.000000014

3 A (postulate 2) and C (postulate 1)

4 At the poles. **5** C

6 Space and time are interdependent—motion in space reduces motion in time.

7 $3.00 \times 10^{8}\,\text{m s}^{-1}$ **8** A and B **9** B

10 You could not tell the difference between (i) and (iii), but in (ii) you could see whether an object like a pendulum hangs straight down.

11 In your frame of reference time proceeds normally. Your heart rate would appear normal. As Mars is moving at a high speed relative to you, the clocks on Mars appear to be moving slowly as time for them, as measured by you, will be dilated.

12 26.8 s

13 **a** 0.992 s **b** 0.992 s

14 **a** 0.866c or $2.598 \times 10^{8}\,\text{m s}^{-1}$.

b No, it can't have doubled to over c! $v = 0.968c$ or $2.90 \times 10^{8}\,\text{m s}^{-1}$

15 **a** 1.67 s **b** length = 1.80 m, height = 1.0 m

16 **a** 5.6 years **b** 2.45 years

c Raqu measures the distance as only $l = \frac{l_0}{\gamma} = \frac{5}{2.29} = 2.183$ ly

17 **a** 1.4 mm **b** No.

18 **a** $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{1}{\sqrt{1-\frac{(0.995c)^2}{c^2}}} = \frac{1}{\sqrt{1-(0.995)^2}} = 10.01$

b No, they don't experience any difference in their own time frame.

c 25.1 years **d** 2.51 years

e No! They see the distance between Earth and Vega foreshortened because of the high relative speed, so to them the distance is only about 2.5 ly.

19 The muon's lifespan is greater according to an observer in the Earth's frame of reference, and the distance travelled is smaller in the muon's frame of reference. See Fully Worked Solutions for more details.

20 B

21 $2.60 \times 10^{8}\,\text{m s}^{-1}$ **22** 3.34×10^{-27} kg

23 1.23×10^{21} J **24** 2.20 s

25 **a** 1 m **b** 3.33×10^{-9} s **c** ct_c **d** 7.6×10^{-9} s **e** 2.3

26 **a** 1.74×10^{-5} s or 17.4 μs

b Non-relativistic: 655 m

Relativistic: 5178 m

27 Proper time, t_0, because the observer can hold a stopwatch in one location and start it when the front of the carriage is in line with the watch and stop it when the back of the carriage is in line with it.

28 23.5 m **29** 2.78×10^{-9} kg

30 See Fully Worked Solutions.

Module 7 Review

The nature of light

Multiple choice

1 B and E **2** A **3** **a** B **b** A **c** D **d** C

4 B **5** D **6** D **7** B

8 B 9 B 10 C 11 A
12 B, C, D, E 13 A 14 C
15 A (postulate 2) and C (postulate 1)
16 A or C 17 C 18 C 19 A
20 C

Short answer

21 a $3.00 \times 10^8 \text{ m s}^{-1}$ b $3 \times 10^{-8} \text{ m}$ c ultraviolet waves
d Any one of:
- UV lamps are used to sterilise surgical equipment in hospitals
- UV lamps are used to sterilise food and drugs
- UV rays help the body to produce vitamin D
- any other suitable use of UV.

22 It is tuned to produce electromagnetic waves at the resonant frequency of water molecules. This makes the water molecules within the food vibrate, and the energy of the water molecules is then transferred to the rest of the food, heating it up.

23 a Δx will be doubled.
b Δx will be doubled.
c Δx will be doubled.
d There will be a wider central band.

24 a when the polarising axes are parallel
b During a full 360° rotation, the polarising axes go from parallel, to perpendicular at 90°, to parallel at 180°, to perpendicular at 270° and back to parallel at 360°. After the start there are two instances of perpendicular polarising axes where the crossed polarising sheets transmit zero light and two instances of parallel polarising axes where the transmitted light intensity will be a maximum.

25 a The wave model and the particle (or corpuscular) model.
b Young's experiment resulted in bright and dark bands or fringes being seen on a screen. These can only be due to interference effects, which are a property of waves. The particle model could not explain the interference effects observed; it would predict just two bright bands.

26 500 nm

27 a K_{max} represents the maximum kinetic energy with which the electrons are emitted. f is the frequency of the light incident on the metal plate. ϕ is the work function, which is the minimum energy required to eject an electron. It is a property of the metal.
b K_{max} is not altered.
c More current will flow.

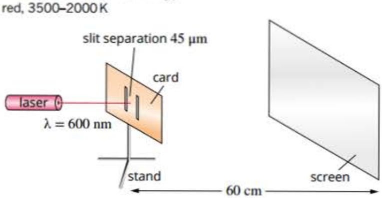
28 In your frame of reference time proceeds normally. As Mars is moving at a high speed relative to you, people on Mars appear to be in slow motion as time for them, as seen by you, will be dilated.

29 $2.93 \times 10^{-11} \text{ s}$

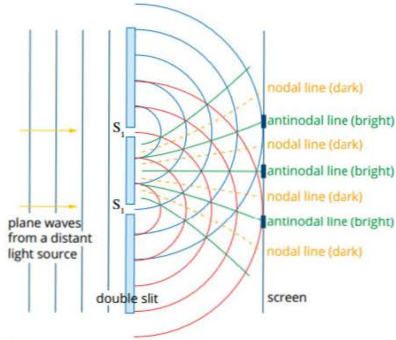
30 very short, μs, very similar, should not, do

31 a Redshift is the stretching out of waves due to the Doppler effect as a light source moves away from the observer.
Stellar spectra can be compared to the Sun's spectrum to see how much certain emission/absorption lines have been shifted. This will tell you whether the star is moving away from you and at what speed.
b Any of: surface temperature, density, chemical composition, rotational velocity spectrum. (A broadening of the lines will occur as an object is turning.)
c red, 3500–2000 K

32 a



b



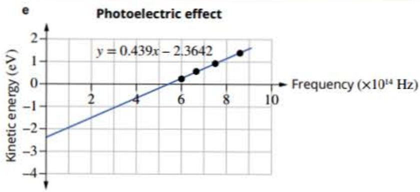
c 8.0 mm
d The fringe spacing decreased by 16%.

33 a In the particle model, the energy of the incident photons is set by their frequency according to $E = hf$. Each incident photon interacts with only one electron; therefore, the energy of the emitted electrons will depend only on the frequency of the incident light. Electron energy is not altered by altering the intensity because this only varies the number of photons, not their energy. Therefore, the energy of the emitted electrons is not affected, only the number emitted.
b The wave model predicts that altering the intensity of the light corresponds to waves of greater amplitude. Hence, the wavefronts should deliver more energy to the electrons and, therefore, the emerging electrons should have higher energy. This is not observed.
c The photoelectric effect supports the particle (photon) model of light because:
1 It predicts a minimum frequency (threshold frequency) and energy before electrons are emitted. The wave model predicts that any frequency should work.
2 The energy of the emitted electrons depends only on the frequency of the incident light. The wave model predicts that increasing the intensity of light would increase the energy of the emitted electrons.
3 It explains an absence of any time delay before electrons are emitted when weak light sources are used. This time delay is suggested by the wave model.

d

Wavelength (nm)	Maximum kinetic energy (eV)	Frequency (×10[14] Hz)
350	1.4	8.57
400	0.91	7.5
450	0.60	6.67
500	0.25	6.0

e



f From the trend line in the graph from part (e), $h = 4.4 \times 10^{-15}$ eV s, $\phi = 2.4$

g potassium

34 **a** 5.6 years **b** 2.4 years

c Relative to her, the distance appeared to be foreshortened by the factor γ, thus the distance she travelled was much less than 5 light years.

d Atomic clocks enabled extremely short events to be timed to many decimal places. Differences in time for the same event to occur, when measured by observers in different inertial frames of reference, indicate that time is not uniform between the two inertial frames. These measurements support Einstein's special theory of relativity.

35 **a** 4.23×10^{-12} J

b 9.3×10^{37} nuclei every second

c 3.8×10^{14} kg

d 1.49×10^{-23} kg m s^{-1}

Chapter 13 Origins of the elements

13.1 The big bang

TY 13.1.1 **a** 69 km s^{-1} Mpc^{-1} **b** 48000 km s^{-1}

13.1 Review

1 The steady state theory suggests that the universe is infinite and that matter is being created all the time at just the right rate to keep the density constant as it expands.

2 The cosmic microwave background radiation was generated in the initial very hot state of the universe, which the steady state theory did not support. The steady state theory had no explanation for the CMB radiation.

3 During inflation, pair production resulted in matter being created as pairs could not annihilate because of the rapid expansion. As the universe cooled, the energy of photons decreased which consequently decreased pair production and most particles annihilated with their antiparticles. However, there was a slight imbalance of matter and antimatter which left the universe mostly with matter particles.

4 The matter created during inflation would never have condensed to form atoms, and therefore, ultimately, galaxies or stars.

5 Annihilation, which involves the conversion of an electron and positron into two photons.

6 There was slightly more matter than antimatter in the early universe. Therefore, once all matter–antimatter pairs had annihilated only matter remained.

7 **a** Hydrogen, helium and lithium; they had been formed in the first few minutes while the universe was hot enough for fusion to occur.

b Other elements formed from the supernovae of the early, large stars. Elements heavier than iron are thought to be produced in neutron star mergers.

8 Hubble's observations of galaxies showed that the further away a galaxy was, the faster it appeared to be moving away from us. He concluded that this can be explained if the universe between the galaxies is expanding as this would cause them to appear to be moving away from each other. Therefore this supported earlier predictions of an expanding universe.

13.2 The life cycle of a star

TY 13.2.1 approximately 100 to 1000 times that of the Sun

13.2 Review

1 The luminosity of a star (which is derived from the absolute magnitude) against the spectral type of stars (from which the temperature of the star is derived).

2 Luminosity increases with the surface temperature.

3 The continuous spectrum provides information about the surface temperature of the star. The absorption spectrum gives information about the elements present.

4 A continuous black-body spectrum and an absorption spectrum both contain a distribution intensity across a range of wavelengths that is determined by the temperature of the object that emitted the radiation. The difference between these two is that the absorption spectrum has had radiation absorbed at various frequencies; these are called absorption lines. An emission spectrum, unlike a continuous black-body spectrum, only contains intensities at certain narrow bands of wavelength; these are called emission lines.

5 Less-broad spectral patterns indicate a large star. A peak in the red section of the visible spectrum coincides with a cooler star. A red giant would best fit these observations.

6 A

7 An A-type star is to the left of the Sun along the main sequence on the H–R diagram. This means it has a higher surface temperature and a greater luminosity than the Sun.

8 D

13.3 The life and death of stars

TY 13.3.1 5.6×10^{8} kg s^{-1}

13.3 Review

1 Nuclear fusion reactions in the Sun involve fusing hydrogen nuclei to produce helium nuclei. In contrast, hydrogen burning in oxygen involves only the electrons in the outer shell of the atoms. Fusion reactions are much more energetic than chemical reactions. The energy involved in nuclear reactions is about 100 million times greater than chemical reactions.

2 5×10^{26} J (to one significant figure)

3 protostar → main sequence star → red giant → white dwarf

4 The total mass of the products is less than the total mass of the reactants. This mass defect implies there was a release of energy from the system which can be quantified using Einstein's mass–energy equation.

5 planetary nebulae

6 The Sun is most likely to initially expand into a red giant before collapsing and becoming a white dwarf.

7 The force of pressure (mostly thermal pressure with some radiation pressure) pushing outwards is balanced by the inwards pull of gravity at hydrostatic equilibrium. The thermal pressure is a result of the energy the particles in the gas or plasma have due to their motion and is exerted through collisions. The radiation pressure is a result of the momentum that photons carry and impart during collisions. The force of gravity is a result of the gravitational field acting on objects that have mass.

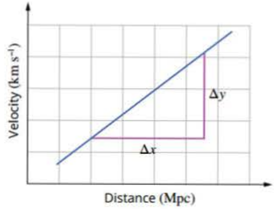
8 Nitrogen has an atomic number of 7, which means it has seven protons in the nucleus. Carbon only has six protons in the nucleus so one more proton must have been produced within the nucleus. From the CNO cycle in Figure 13.3.11, this process occurs through the addition of a proton (hydrogen nucleus) to a carbon atom.

Chapter 13 Review

1 Hubble measured the distance and redshift of many galaxies. He then calculated the velocity of recession for these galaxies and noticed that the further away a galaxy was the faster it was receding.

2 pair production and inflation

3



slope = $\frac{\Delta y}{\Delta x} = \frac{v}{d} = H_0$

$\therefore v = H_0 d$, which is Hubble's law

4 Predictions made by physicists as part of the work on the big bang theory say that the early universe should have been composed of mainly hydrogen and some helium. Observations of the cosmic microwave background radiation and the spectra of old stars yield results that agree closely with these predictions.

5 If the universe is expanding and the galaxies are getting further apart as the universe between them expands, this implies that in the past they were closer together. If we extrapolate back in time, the finite amount of mass and energy in the universe would have been contained in a smaller and smaller volume. This would mean the average energy of the particles (i.e. the temperature) would increase. At the very beginning of the universe, this would mean an infinite density and temperature—a hot big bang beginning for the universe.

6 A larger Hubble constant would suggest a smaller age of the universe as the age of the universe can be estimated by calculating the reciprocal of the Hubble constant.

7 the cosmic microwave background (CMB) radiation

8 spectral class, surface temperature and chemical composition

9 Their spectra show the same lines as our Sun, and these lines correspond to the 98 known elements in our periodic table.

10 Betelgeuse, Rigel, Polaris (pole star), Arcturus

11 Spectral class M stars have low surface temperatures of less than 3500 K. These are the only stars with temperatures in their atmospheres low enough for molecules to exist and therefore to produce absorption spectra with strong lines for those molecules.

12 **a** white dwarfs **b** main sequence **c** supergiant stars **d** red giant stars

13
- Like all stars Rigel would have started from a dust and gas cloud collapsing to form a protostar.
- Rigel would have spent most of its lifetime on the main sequence fusing hydrogen into helium.
- As Rigel starts to convert silicon to iron as the main fusion process, less energy will be produced than needed for the fusion process. It will begin to collapse.
- As a giant star, the core of Rigel can then be expected to heat to billions of degrees in a fraction of a second.
- An explosive supernova results.
- The final stage will be either a neutron star or a black hole.

14 Stars are born into the middle (approximately) of the main sequence, after rapidly igniting once the protostar collapses.

15 The Sun is close to the centre of the H–R diagram; that is, in terms of the overall range, it is of average temperature and average brightness. However, most stars are actually cooler and fainter than the Sun.

16 4×10^{26} J (to one significant figure)

17 Hydrogen fusing to helium is the main reaction in main-sequence stars, while helium fusing to carbon via the triple alpha process is the main reaction in red giants.

18 New elements are formed through nuclear fusion reactions in a process called nucleosynthesis. During their lives on the main sequence, stars like the Sun form helium. When they become a red giant they form elements like carbon, nitrogen and oxygen. Larger stars form elements up to iron while they are giants and other heavier elements when they explode as a supernova.

19 The predictions of the big bang theory are very accurate. Theoretical predictions indicate that the elements formed in the early universe should have been approximately 25% helium and the remainder hydrogen. Observations determining the abundance of these elements in stars and galaxies agree closely with the predicted values. The predicted abundance of other elements and isotopes predicted by the big bang theory were tested by the WMAP mission, and again observations closely matched the abundances from theoretical predictions.

20 Experimental evidence and observation are the most fundamental parts of the scientific method. Predictions are made by models and theories. These are then tested against experimental evidence or observation. If the predictions are supported by the evidence then scientists can be confident about their theory or model and can then test them in other ways. If predictions are not supported then the theory or model must be altered or new ones developed and tested against evidence. It is through this process that scientists can develop theories or models that are accepted, continue to be refined and become better at explaining what is being observed.

The steady state theory was another model that was proposed to describe the evolution of the universe. By using observational evidence for the big bang, such as the cosmic microwave background (CMB) radiation and the increasing acceleration of galaxies further away from the Milky Way (seen through cosmological redshift), astrophysicists were able to rule out the steady state theory and instead now use the big bang theory to model the expansion and evolution of the universe. Theories such as this always rely on the most current evidence and are able to change and adapt depending on what scientists have discovered.

21 Non-fusing helium builds up in the core of the star and a shell of hydrogen begins fusing around a non-fusing helium rich core. This causes the outer layers of the star to expand and cool forming a red giant.

22 Neutron stars are extremely dense objects comprised mainly of neutrons held up against gravity. In contrast, the gravity within a black hole has exceeded any force than can resist collapse and have obtained some much higher, possibly infinite, density. The escape velocity of neutron stars, unlike black holes, is not greater than the speed of light so they are visible through radiation that escapes their surface.

23 The expanding balloon is an analogy used to explain the expansion of the universe. The surface of the balloon represents a two-dimensional space. The universe is a three-dimensional space. The space within the balloon and outside the balloon does not represent anything in this analogy; only the two-dimensional surface is significant. As the balloon is inflated, the distance between the galaxies increases. From each galaxy, all the other galaxies are seen to move away. The further the galaxies are apart, the faster they seem to move away from each other. However, as the galaxies have not moved in the reference frame that was drawn on the balloon, they do not have relative velocities. Unlike the steady-state theory, where matter is being created all the time at just the right rate to keep the density of the universe constant, the big bang theory shows that the origin of matter occurred in the very early universe. Evidence of cosmological redshift (the expansion of space) supports the big bang theory, which helps us to understand the processes by which galaxies (and then stars and planets) were all formed. The heavier elements were then created due to the life cycles of stars. Fusion processes within stars create elements such as carbon, which, due to supernova events, then go on to create planets and even life.

Chapter 14 Structure of the atom

14.1 The electron

TY 14.1.1 a 1.8×10^8 m s^{-1} **b** 34 m

14.1 Review

1. **a** $5.93 \times 10^7 \text{ m s}^{-1}$ **b** 22.5 cm
2. $4.13 \times 10^6 \text{ m s}^{-1}$, 3.23 m
3. Water is more volatile (evaporates easily) than oil, so the mass and size of the droplets change more rapidly than oil drops.
4.

Atomic model	Description
solid-ball model	An indivisible ball. Where the name atom first appeared as it was referred to as atomos meaning indivisible.
plum pudding model	From Thomson's experiments with electrons, he proposed the plum pudding model in 1904. The atom in this model is a ball of positive charge with negative charges embedded within it.
nuclear model	The majority of the mass of an atom in this model is in a small positive nucleus which is surrounded by negative electrons.
planetary model	The electrons in this model orbit the positive nucleus in specific pathways, like planets orbiting the Sun.

14.2 Nuclear model of the atom

14.2 Review

1. The model predicts particles are affected only by the diffuse electrical charge. Electrical repulsion between the positively charged alpha particles and the diffuse positive atomic charge leads to a small amount of repulsion, observed as scattering of less than 2° from their initial path.
2. Higher energy (higher velocity) alpha particles will encounter the atoms with a greater momentum, and the effect of the electrical repulsion would be reduced with consequent reduced scattering.
3. The alpha particles and nucleus are both positively charged and experience electrical repulsion. The particles cannot touch as the repulsion becomes infinite as they approach.
4. A neutron—as a neutron decays to a proton and electron, the mass of the neutron is greater, with the difference being approximately the mass of an electron.

Chapter 14 Review

1. The number of protons is equal to the number of electrons. The number of neutrons does not affect the electrical charge.
2. There is a gap or hole in the anode through which some of the electrons pass.
3. a magnetic field
4. $3.0 \times 10^7 \text{ m s}^{-1}$ **5** $4.0 \times 10^7 \text{ m s}^{-1}$
6. **a** 9.3×10^{-15} N **b** 4.0×10^{-3} m
7. 5.8×10^{-4} T
8. $1.58 \times 10^6 \text{ m s}^{-1}$
9. **a** $q = \frac{mgd}{V}$ **b** $q = \frac{4\rho\pi r^3 gd}{3V}$
   c 1.4×10^3 V **d** 480 V
10. C, A, B
11. The mass of an atom is concentrated in the nucleus, and that the nucleus occupies a tiny fraction of the volume of the atom.
12. Electrons. Atomic absorption spectra showed light could be absorbed and emitted by atoms, and in discrete bands. Rutherford's model did not address how this could occur, whether by the electrons or in the atomic nucleus.
13. Rutherford could not have concluded anything from this. The null result would have been consistent with the plum pudding model. However if the nucleus was very small, such that the backscattering may have occurred with one in one million collisions, it would have still occurred but have been harder to detect. Absence of evidence is not evidence of absence.
14. Team Proton: 28; Team Electron: 51 408
15. Divided atoms would have different numbers of protons in their nuclei, and therefore are different elements. This gives the daughter elements different chemical properties to the mother atom.
16. Protons have a single positive charge whereas neutrons are electrically neutral. They are approximately the same mass (neutrons are slightly heavier).
17. $7.6 \times 10^6 \text{ m s}^{-1}$
18. 2.21 m
19. **a** $8.2 \times 10^5 \text{ m s}^{-1}$
    b 0.17 T
20. See Fully Worked Solutions.

Chapter 15 Quantum mechanical nature of the atom

15.1 Bohr model

TY 15.1.1 0.42 eV **TY 15.1.2** 4.6×10^{-19} J

TY 15.1.3 103 nm, Lyman series **TY 15.1.4** 97.2 nm, Lyman series

TY 15.1.5 A photon of 6.7 eV corresponds to the energy required to promote an electron from the ground state to the second excited state ($n = 1$ to $n = 3$). The photon may be absorbed.

A photon of 9.0 eV cannot be absorbed.

A photon of 11.0 eV may ionise the mercury atom. The ejected electron will leave the atom with 0.6 eV of kinetic energy.

15.1 Review

1. An emission spectrum for an element typically consists of a series of spaced coloured lines on a black background.
2. The different coloured lines in the emission spectrum of an atom correspond to the possible electron transitions between energy levels within the atom.
3. The energy levels within an atom are commonly represented as horizontal lines on a graph.
4. 4.0×10^{-19} J **5** 3.0×10^{-6} m **6** 12.75 eV
7. 95 nm

15.2 Quantum model of the atom

TY 15.2.1 5.7×10^{-13} m **TY 15.2.2** 1.0×10^{-36} m

TY 15.2.3 0.17 nm

15.2 Review

1. 7.3×10^{-10} m **2** B
3. **a** 3.5×10^{-11} m **b** $2.1 \times 10^7 \text{ m s}^{-1}$
4. it would increase
5. Newtonian physics describes the position and velocity of an object as 'known', so its future position can be predicted. Quantum mechanics proposes that you cannot know the position and velocity of a particle at the same time. So the assumptions that Newtonian physics makes do not fit with what happens at the subatomic level.
6. D

Chapter 15 Review

1. 2.11 eV **2** 3.61 eV
3. No. The photon energy will be exactly equal to the energy difference between the electron's initial and final levels.
4. The light globe produces a continuous spectrum showing all the colours of the rainbow. The vapour lamp produces a discrete spectrum showing just coloured lines.
5. The temperature of the light globe filament increases when it is switched on. As the filament heats up, the free electrons in the tungsten atoms collide, accelerate and emit photons. A wide range of photon wavelengths are emitted due to a wide range of different collisions (some weak, some strong), and some visible light and a lot of infrared radiation are produced.

6 $1.8 \times 10^5\,m\,s^{-1}$
7 The location of electrons can't be restricted to specific orbital paths, because their precise location and velocity cannot be known simultaneously.
8 Sodium vapour is heated so that electrons are excited to higher energy levels, emitting light when they transition back to lower energy levels. The most common transitions in sodium produce orange light.
9 $1.7 \times 10^{-35}\,m$
10 No.
11 Electrons in the atom cannot assume a continuous range of energy values but are restricted to certain discrete values.
12 $2.9 \times 10^{15}\,Hz$
13 Bohr's work on the hydrogen atom convinced many scientists that a particle model was needed to explain the way light behaves in certain situations.
14 The lines in the emission spectrum of hydrogen correspond to the missing lines in hydrogen's absorption spectrum.
15 it will increase
16 it would increase
17 high-energy orbits of multi-electron atoms, the continuous emission spectrum of solids, Zeeman splitting and the two close spectral lines in hydrogen that are revealed at high resolution
18 –0.54 eV
19 The wavelength of a cricket ball is so small that its wave-like behaviour could not be seen by a cricket player.
20 $1.32 \times 10^5\,m\,s^{-1}$
21 $\lambda = \frac{h}{\sqrt{2qVm}}$ 22 $p = \frac{h}{\lambda}$
23 Because a high-speed electron has a shorter wavelength than a light wave.
24 See Fully Worked Solutions.

Chapter 16 Properties of the nucleus

16.1 Radioactive decay

TY 16.1.1 90 protons, 230 nucleons, 140 neutrons

TY 16.1.2 mass number = 214, atomic number = 82, lead

16.1 Review

1 nucleons
2 79 protons and 118 neutrons
3 235
4 In a neutral atom it is the same as the number of protons, i.e. the atomic number.
5 an alpha particle 6 beta-minus 7 beta-plus
8 **a** gamma **b** beta-minus **c** alpha **d** beta
e gamma

16.2 Half-life

TY 16.2.1 **a** $0.046\,hour^{-1}$ **b** 3.9×10^7 nuclei

16.2 Review

1 the count rate or the number of decays each second
2 4.0×10^{10} atoms 3 3.0×10^{11} nuclei
4 **a** 10 **b** 240 000 years
5 50% 6 192 μg 7 15 minutes
8 **a** 10 minutes **b** 50 Bq
9 beta decay, 20 years
10 seven alpha, four beta-minus

16.3 Nuclear fission and fusion

16.3 Review

1 A force of attraction that acts between every nucleon but only over relatively short distances. This force acts like a nuclear cement.
2 The waste products are radioactive and remain so for centuries or even millennia.
3 Attractive forces from other nucleons due to the strong nuclear force.
4 fissile—uranium-235 and plutonium-239
non-fissile—uranium-238 and cobalt-60
5 3
6 Fusion is the joining together of two small nuclei to form a larger nucleus. Fission is the splitting apart of one large nucleus into smaller fragments.
7 Electrostatic forces of repulsion act on the protons. If the protons are moving slowly they will not have enough energy to overcome the repulsive forces and they will not fuse together.
8 Initially, electrostatic forces of repulsion act on the protons, but they are travelling fast enough to overcome these forces. The protons will get close enough for the strong nuclear force to take effect and they will fuse together. These protons have overcome the energy barrier.
9 It is conserved—there are five nucleons on each side of the reaction.

16.4 Energy from nuclear reactions

TY 16.4.1 **a** 3 **b** $4.12 \times 10^{-11}\,J$, 258 MeV **c** 0.16%

16.4 Review

1 **a** $1.91 \times 10^{-11}\,J$ **b** $1.19 \times 10^8\,eV$
2 The mass of the products is less than the mass of the reactants. The mass difference is related to the energy released via $E = mc^2$.
3 The amount of energy released per nucleon during a single nuclear fission reaction is less than the amount for a single fusion reaction.
4 less than 1%
5 **a** $a = 4$, $b = 2$, ^4_2He **b** $5.9 \times 10^{-29}\,kg$
6 **a** atomic number = 2, mass number = 3, ^3_2He
b $3.7 \times 10^{-12}\,J$ **c** $4.1 \times 10^{-29}\,kg$
7 The binding energy per nucleon increases and the nucleus becomes more stable.

Chapter 16 Review

1 20 protons and 25 neutrons
2 A nuclide that is able to split in two when hit by a neutron is fissile.
3 27 protons, 33 neutrons and 60 nucleons
4 a gamma ray
5 beta-minus
6 No, only a few nuclides (e.g. uranium-235 and plutonium-239) are fissile.
7 The strong nuclear force causes the proton to be attracted to all other nucleons. It will also experience a smaller electrostatic force of repulsion between itself and other protons.
8 **a** beta-minus **b** proton **c** alpha **d** neutron
e gamma **f** beta-positive (positron)
9 atomic number = 3, mass number = 7, lithium ^7_3Li
10 **a** a proton **b** a neutron
11 Neutrons are uncharged and are not repelled by the nucleus as alpha particles are.
12 To give the hydrogen nuclei enough energy to overcome the electrostatic repulsion between them.
13 **a** $x = 208$, $y = 82$ **b** $x = 176$, $y = 78$
14 $a = 18$, $b = 9$, $c = 18$, $d = 8$
X is fluorine, F.
Y is oxygen, O.
15 carbon-12
16 Electromagnetic forces are balanced by the strong nuclear force acting between all nucleons in close proximity.
17 **a** gamma **b** gamma
18 gamma radiation
19 The bombarding electrons will be strongly repelled by the electron clouds of the atoms as they are all negatively charged. The small mass of the bombarding electrons also makes them relatively easy to repel compared to, for example, a proton.
20 3.0×10^{14} atoms
21 1.5×10^{10} atoms

22 The long half-life means that the source will not need to be replaced for many years. The gamma rays have a strong penetrating power so they are able to penetrate the skull and reach the tumour site.

23 4.49×10^{-11} J

24 a The combined mass of the hydrogen and helium-3 nuclei is greater than the combined mass of the helium-4 nucleus, positron and neutrino.

b The energy has come from the lost mass (or mass defect) via $E = mc^2$.

c 3.4×10^{-12} J   d 3.8×10^{-29} kg

25 Fission produces radioactive fission fragments, whereas fusion produces no radioactive waste products. Fusion creates more energy per nucleon than fission.

26 The binding energy per nucleon increases and the nucleus becomes more stable.

27 It is more stable.

28 8.0×10^{-15} J

29 Radioactive decay of a nucleus is a random process that can't be predicted for an individual atom. Scientists have discovered that the time it takes for half the nuclei to decay is constant for a particular atom; this is called a half-life. The half-life can be used to predict the radiation emitted so that it is usable in medical research, energy production and other industries.

Chapter 17 Deep inside the atom

17.1 The Standard Model

17.1 Review

1 between nucleons, i.e. protons and neutrons

2 weak nuclear: W^+, W^- and Z bosons; strong nuclear: gluons; electromagnetic: photons

3 Gauge bosons: force-carrier particles.
Leptons: fundamental particles that can be found individually and do not experience the strong force.
Quarks: experience the strong nuclear force and form composite particles called hadrons.

4 Quarks must exist in groups of two or three; leptons can exist individually. All quarks experience the strong force; leptons do not. Quarks have non-integer (fractional) charges; leptons have charges of −1 or 0.

5 the force-carrier particle being exchanged, i.e. a boson

6 Mesons: contain two quarks—one of normal matter and one of antimatter.
Baryons: contain three quarks; includes protons and neutrons.

7

Gauge boson	Lepton	Hadron
gluon, photon	electron, neutrino, muon	neutron, proton

17.2 Evidence for the Standard Model

17.2 Review

1 charge

2 They are passed through an electric field.

3 linac, booster ring, storage ring, beamlines

4 very close to the speed of light

5 from the outward spiralling circular path of the particles

6 Cyclotrons use a static magnetic field and an alternating electric field to accelerate particles, so that the particles spiral outwards as they gain energy. In a synchrotron, the magnetic field strength is increased as the particle's energy increases, allowing for a fixed radius even as the energy increases.

7 It has moveable tracking detectors that allow for studying a collision event at multiple points after the collision.

Chapter 17 Review

1 A particle collides with its antiparticle and mass is converted into energy.

2 protons: up, up, down; neutrons: up, down, down

3 electromagnetism and the strong and weak nuclear forces

4 A proton is made up of two up quarks $(2 \times +\frac{2}{3})$ and one down quark $(-\frac{1}{3})$ so $\frac{4}{3} - \frac{1}{3} = +1$. A neutron is made up of two down quarks $(2 \times -\frac{1}{3})$ and one up quark $(+\frac{2}{3})$ so $-\frac{2}{3} + \frac{2}{3} = 0$.

5 weak nuclear, electromagnetic, gravity, (strong nuclear)

6 The Standard Model is based on the assumption that forces arise through the exchange of particles called gauge bosons (or just bosons). Each of the three forces is mediated by a different particle: strong—gluon, electromagnetic—photon, weak—W^+, W^- and Z.

7 fundamental particle, lepton

8 Electrons are 'boiled' off a heated wire element acting as a cathode.

9 linear accelerator

10 A beamline is typically a stainless steel tube of 15–35 m in length along which synchrotron light travels from the storage ring, where it is produced, to its target for experimental work.

11 X-ray diffraction

12 in the electron gun

13 ranging from infrared to highest-frequency X-rays

14 An accelerator capable of generating the 2 TeV accelerated particles required wasn't available before the early 1990s.

15 Any two from the following:
- It is unable to explain dark matter.
- It is unable to explain the absence of antimatter in the universe today.
- It is unable to provide a completely unified model for all fundamental forces; interactions with gravity are unexplained.
- It is unable to explain why neutrinos have mass.

16 It allows X-ray diffraction techniques to be completed over considerably shorter times than with traditional X-ray sources.

17 Synchrotrons produce a continuous spectrum of radiation, and particular wavelengths within that spectrum are highly selectable or tunable.

18

X-ray tube	Synchrotron
single burst relatively divergent standard intensity	produced over hours highly collimated many times brighter

19 In the storage ring electrons orbit for hours at a time at speeds near that of light, before being channelled along the beamlines for experimentation.

20 in the booster ring

21 in the linac

22 hydrogen atom source, electrons are stripped away, linac, Proton Synchrotron Booster, Super Proton Synchrotron, storage ring, detectors

23 Reponses will vary; see the Fully Worked Solutions.

Module 8 Review questions

From the universe to the atom

Multiple choice

1 C	2 A	3 C	4 D
5 B	6 C	7 A	8 D
9 C	10 C	11 A	12 D
13 C	14 B	15 B	16 B
17 A	18 B		

Short answer

19 The lines in stellar spectra are shifted to longer wavelengths—towards the red end of the spectrum. The more distant objects are shifted the most and hence are receding the fastest. This is consistent with a model in which space–time and the whole universe is expanding.

20 **a** The energy of photons results from the mass of the leptons being converted to energy.
b 1.6×10^{-13} J

21 **a** Bohr's work on the hydrogen atom convinced many scientists that a particle model was needed to explain the way light behaves in certain situations.
b Neils Bohr would state that if incident light had an energy value less than the minimum energy difference between the lowest and next orbital levels within the hydrogen atom, the light would not result in any orbital changes.

22 **a** Photon energy > ionisation energy, i.e. the photon has enough energy to free the electron.
b $0.4\,\text{eV} = 6.4 \times 10^{-20}$ J
c $3.41 \times 10^{-25}\,\text{kg m s}^{-1}$
d 1.94×10^{-9} m
e Since there is no energy level 10.0 eV above the ground state, the photon cannot be absorbed.

23 **a** $1.5 \times 10^5\,\text{N C}^{-1}$ (or V m^{-1}) downwards
b 1.28×10^{-18} C
c 8

24 **a** 2.42×10^{-8} m
b A series of bright and dark fringes.
c The high-speed electrons are exhibiting wave-like behaviour.
d $164\,\text{m s}^{-1}$

25 **a** 0.0416 nm
b There would be circular bands or fringes of specific spacing around a common central point.
c less diffraction would occur; that is, the circular bands would be closer together

26 **a** 1 minute
b 4.7 g
c ${}^{26}_{11}\text{Na} \rightarrow {}^{0}_{-1}\beta + {}^{26}_{12}\text{Mg}$ + energy

27 Cs-137: 55 protons, 82 neutrons, 137 nucleons
I-131: 53 protons, 78 neutrons, 131 nucleons

28

Particle	Property
gluon	• mediator of the strong nuclear force • interacts with quarks
photon	• mediator of the electromagnetic force • interacts with charged particles
W^+, W^- and Z	• mediator of the weak nuclear force • causes nuclear decay
(graviton)	• mediator of the gravitational force

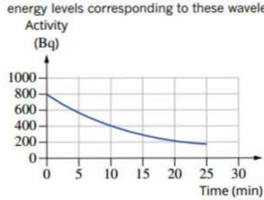
29 In the single-slit diffraction experiment, as the slit is made narrower the position of the particle becomes more precisely known. As a consequence, the direction, and therefore the momentum, of the particle becomes less precisely known, because with a narrower slit the diffraction pattern becomes wider.

30 The momentum (or velocity) of a particle is known more precisely, so the uncertainty in a particle's position, Δx, becomes greater.

31 **a** $v = H_0 d = 70\,\text{km s}^{-1}\,\text{Mpc}^{-1} \times d$
The speed is directly proportional to the distance.
Proxima Centauri is 1.3 pc away and so its speed is:
$70 \times 10^{-6}\,\text{km s}^{-1}\,\text{pc}^{-1} \times 1.3\,\text{pc} = 91\,\text{mm s}^{-1}$
The edge of the universe is:
$4.4 \times 10^{26}\,\text{m} = \frac{4.4 \times 10^{26}\,\text{m}}{3 \times 10^5} = 1.4 \times 10^{10}\,\text{pc}$
$\text{Speed} = 70 \times 10^{-6}\,\text{km s}^{-1}\,\text{pc}^{-1} \times 1.4 \times 10^{10}\,\text{pc} = 9.8 \times 10^8\,\text{m s}^{-1}$
b The edge of the visible universe is receding from us at a speed in excess of the speed of light. That is not a violation of the principles of special relativity, as no object is moving through space at a speed in excess of the speed of light, it is purely a relative velocity. If there is concern that technically this is not the edge of the visible universe, since light would never reach us, we are only concerned here with a factor of 3 or so and in astronomical terms, this is good enough!
c $2 \times 10^{-18}\,\text{s}^{-1}$. The relative proximity of a star like Proxima Centauri would mean redshift would not be measurable.
d The extremely small value for H_0 reflects the fact that the recession velocity is only significant for huge distances!

32 **a** To produce diffraction patterns with the same fringe separation, they must have equivalent wavelengths.
b 3.6×10^{-11} m
c 3.6×10^{-11} m
d $1.8 \times 10^{-23}\,\text{kg m s}^{-1}$
e No. The energy of the X-rays is given by $E = \frac{hc}{\lambda}$ and the energy of the electrons is given by $\Delta K = \frac{1}{2}mv^2$.
f de Broglie would say that the electrons (with their associated wavelengths) were diffracted as they passed through the gaps between the atoms in the crystal, creating a diffraction pattern.
g In addition to their particle properties, electrons have a de Broglie wavelength. The orbit must fit an integral number of wavelengths so that a standing wave is formed ($2\pi r = n\lambda$). Only energy levels corresponding to these wavelengths exist.

33 **a**



b about 320 Bq **c** 10 min **d** $1.2 \times 10^{-3}\,\text{s}^{-1}$ **e** 100 Bq
f **i** ${}^{4}_{2}\text{He}$, ${}^{0}_{-1}\beta$
ii They have different numbers of protons.
iii ^{210}Bi can undergo beta decay to form ^{210}Po and then this undergoes alpha decay to form ^{206}Pb. Alternatively, it an undergo an alpha decay first to form ^{206}Tl and then the subsequent beta decay results ion ^{206}Pb.
iv They have different numbers of neutrons: 214, 210 and 206, respectively.

34 a

Category	Particle type	Description	Particle name
gauge bosons		mediators of the fundamental forces	photons, gluons, gravitons, W^+, W^-, and Z
fermions (make up all matter)	leptons: positrons, electrons neutrinos, muons	• experience the weak nuclear force, exchanging W and Z • charged leptons experience the electromagnetic force, exchanging photons • do not experience the strong force	positrons, electrons, neutrinos, muons
	hadrons	experience the strong force, exchanging gluons made up of quarks	
	- baryons	made of three quarks	protons, neutrons, antiprotons
	- mesons	made of two quarks	pions

b Hadrons. The other particles listed are fundamental particles and so do not have an internal structure.

c It gives mass to all elementary particles.

d dark matter and antimatter

35 a The big bang is an expansion of space–time. Before the big bang there was no space, time or matter, so it is not a case of matter exploding out into space in a time continuum, but space and time itself being created at the big bang event as energy converted to matter, after which space expanded. The energy present allowed the creation of matter–antimatter pairs and the rapid inflation of the universe prevented annihilation taking place immediately, taking the created matter with it. While it is true that the early universe was extremely dense, the big bang theory would suggest that mass–energy, space and time all emerged at once from nothing.

b The radiation which when created would have had a very short wavelength would be expected to 'stretch out' with space itself, and so would have a much longer wavelength as space expanded. Calculations show that this would be in the microwave range today.

c The variations indicate a slightly uneven distribution of light and therefore matter. This allowed gravitational attraction to collect clumps of matter together, ultimately forming stars and galaxies. If there had been completely uniform radiation, there would have been no universe as we know it.

d Pair production is the creation of a matter and antimatter pair of particles, such as a positron and an electron from a photon. This is a mechanism for the creation of particles from photons.

e Normally pairs annihilate rapidly with the release of photons, but the rapid inflation moved the pairs apart so that the particles were able to persist.

f As the universe cooled, the average photon energy dropped to a level at which a photon no longer had the energy required to create a matter–antimatter pair.

g Any atoms formed would immediately be ionised as the photons, although not having enough energy for pair production, certainly possessed the ionisation energy for a hydrogen atom.

h Fusion requires very high densities, temperatures and pressures for charged particles to overcome their mutual repulsion and come close enough for the strong nuclear force to exceed the electrostatic repulsion. This happened in the first few seconds after the big bang, and then particle distances increased and energies dropped below the values required for fusion to be possible. Fusion reignited in stars much later when gravitational forces once again brought particles together at high densities.

i Photon energies had to be below the ionisation energy of the atoms.

j Gravity caused particles to aggregate. As the dust clouds collapsed under their mutual attraction, vast amounts of energy were released and this created the temperatures and pressures for fusion to reignite in the first stars.

Glossary

A

absolute error Equal to half the smallest unit of measurement.

absolute magnitude The absolute magnitude of a star or other celestial object corresponds to the apparent magnitude of the object if it was 10 parsecs (32.6 light-years) from Earth.

absolute value The magnitude of a variable ignoring its sign. The absolute value of a number is always positive.

absorb To take in (energy).

absorption The taking up and storing of energy, such as radiation, light or sound, without it being reflected or transmitted. During absorption, the energy may change from one form into another.

absorption spectrum Spectrum containing dark lines in the positions of the wavelengths that are absorbed by a gas as light passes through it. This is related to the emission spectrum of the gas.

acceleration The rate of change of velocity. Acceleration is a vector quantity.

acceleration due to gravity Rate at which a falling object will accelerate in a gravitational field. Equivalent to the gravitational field strength. Measured in $m\,s^{-2}$.

accuracy The ability to obtain the correct measurement.

activity The number of nuclei of a radioactive substance that decay each second, measured in Becquerels (Bq).

aether An invisible, massless, rigid substance that was proposed as the medium in which light waves propagate. There is no experimental evidence for the existence of the aether.

affiliation Connections or associations between two parties.

air resistance The frictional force that acts against moving objects as they travel through the air. Air resistance always acts in the opposite direction to the motion of an object.

alpha particle A particle consisting of two protons and two neutrons ejected from the nucleus of a radioactive nuclide.

alternating current In an alternating current (AC), electrons oscillate backwards and forwards around a mean position, as opposed to direct current (DC). Household power supplies usually operate at 240 V AC.

alternator An electric generator that produces alternating current (AC).

altitude Height above a planet's surface.

ammeter An ammeter is an instrument used to measure the electric current in a circuit. Electric current is measured in amperes (A), which is why it is called an ammeter.

ampere A unit of electric current equal to a flow of one coulomb per second.

Ampere's law The sum of all the magnetic field elements that make up the circle surrounding the wire is equal to the product of the current in the wire and the permeability of free space.

amplitude The maximum up or down displacement of a wave measured from its equilibrium position.

angular velocity A measure of how quickly an object is turning measured in radians per second.

annihilation The process in which matter is completely converted into energy. This is not a chemical process in which matter in one form is converted to matter in another form, as in burning.

anode A positively charged electrode, as of an electrolytic cell, storage battery or electron tube. Also, the negatively charged terminal of a primary cell or of a storage battery that is supplying current.

antimatter Particles that have the same mass as their ordinary matter equivalents but opposite properties like electromagnetic charge, spin, baryon number and lepton number.

antineutrino A neutral subatomic particle that interacts very weakly with other matter; the antimatter particle of a neutrino.

aphelion The point in an elliptical orbit that is furthest from the Sun.

apogee The furthest point in an elliptical orbit between a central mass and an orbiting body.

apparent brightness How bright a star or other celestial object appears from Earth measured in magnitudes. It is based on the brightest star in the northern sky being magnitude +1. A change of −1 corresponds to a brightening by about 2.5 times.

apparent magnitude An arbitrary scale based on the brightest star in the northern sky being magnitude +1. A change of −1 corresponds to a brightening by about 2.5 times.

apparent weight The weight felt by a person when their body is stationary or in motion. Sometimes it is higher or lower than their usual weight. Equivalent to the size of the normal reaction force acting on the person.

apparent weightlessness When an object is in free fall and there is no force between it and its surroundings. It appears to be weightless although it is still under the influence of gravity. When there is no normal reaction force acting between an object and a surface.

armature A revolving structure in an electric motor or generator, wound with the coils that carry the current. It rotates within a magnetic field to induce an emf.

artificial satellite Bodies, such as Sputnik, the Hubble Space Telescope or NOAA-19, made by humans and placed in orbit around a planet or the Moon.

atomic model A description of the structure of the atom.

atomic number The number of protons in a nucleus.

axis of rotation An imaginary line through the centre of mass or pivot point of an object, which is perpendicular to the plane of rotation of the object.

B

ballistic Object's movement due only to its weight; without powered flight.

banked track A track inclined at some angle to the horizontal enabling vehicles to travel at higher speeds when cornering compared with around a horizontal curved path.

bar graph A graph in which categorical data is represented by horizontal bars. Each bar represents one category of independent variable (such as a range of values, or a particular type of thing), and the length of the bar represents the value of the dependent value for that range or thing.

baryon A composite particle composed of three quarks. Baryons belong to the particles called hadrons. The most common examples are the proton and the neutron.

baryon number A quantum number conserved in particle interactions. This means the sum of the baryon numbers before an interaction is equal to the sum of the baryon numbers after the interaction. Baryons (particles containing three quarks) are assigned +1, antibaryons are assigned −1 and all other particles are assigned 0.

beamline A pathway in which the photon beam generated in a synchrotron travels from the storage ring to an experiment room or end station.

beta particle An electron or positron ejected from the nucleus of a radioactive nuclide.

bias A form of systematic error resulting from the researcher's personal preferences or motivations.

big bang theory The leading model for how the universe was created. It describes the universe starting in a high-density state and then expanding.

binding energy Energy required to split a nucleus into its separate nucleons.

black body A black body does not reflect any radiation. It does not necessarily have to be black; for instance, the Sun can be modelled as a black body.

black dwarf A black dwarf star is a white dwarf star that has cooled such that it no longer emits any significant heat or light.

black hole A collapsed star so massive that not even light can escape from its gravitational field.

blueshift The change of frequency that occurs in any wave phenomenon as the source of the waves moves towards the observer. Likewise, a redshift occurs if the source is moving away from the observer.

booster ring Part of a synchrotron where the electron energy and speed is further increased.

brushes Devices that transfer the current in the rotating coil to a stationary external circuit by pressing against the split ring commutator or the slip rings.

C

cathode In a cathode ray tube, the cathode is a filament which, when heated, produces electrons.

cathode ray tube A vacuum tube in which a hot cathode emits a beam of electrons that pass through a high voltage anode and are focused or deflected before hitting a fluorescent screen.

centre of mass Point at which the mass of an object is considered to be concentrated for the purpose of analysing motion.

centripetal acceleration Acceleration directed towards the centre of a circle when an object moves with constant speed in a circular path.

centripetal force The force that causes an object to travel in a circular path; this can include gravity, tension, normal force and friction.

chain reaction A series of nuclear fissions that may be controlled or uncontrolled.

charge A property of matter that causes electric effects. Protons have positive charge, electrons have negative charge and neutrons have no charge.

classical physics The physics of Galileo and Newton, in which the addition of velocities has no limit, and length and time are constant.

coherent Waves that are in phase, i.e. at the same stage at the same time.

collinear Lying on the same straight line.

colour charge A property of quarks that is related to how they bond together. (Note, this is not related to the normal interpretation of colour.)

component A vector that makes up one part of a two-dimensional vector.

commutator The rotating cylindrical copper segments in an electric motor that carry the best current from the brushes to the coils in the armature, which are in the best position for maximum torque.

conductor A material, usually metal, through which charges move freely. The electrons in conductors are only very slightly attracted to their respective nuclei and can therefore move easily from one atom to another throughout the material.

conserved When a quantity that exists before an interaction is exactly equal to the quantity that exists after the interaction.

conservation law A conservation law describes a condition that a measurable property must remain unchanged.

constructive interference The process in which two or more waves of the same frequency combine to reinforce each other. The amplitude of the resulting wave is equal to the sum of the amplitudes of the superimposed waves.

contact force Forces that exist when one object or material is touching another. Friction, drag and normal reaction forces are contact forces.

continuous variable A variable that can have any number value within a given range.

controlled variable A variable which is kept constant in order to reliably find the effect of changing the dependent variable.

conventional current Basically the same as electric current. Conventional current is in the opposite direction to electron flow.

coulomb The SI unit of charge; 1 C is equivalent to the combined charge of 6.2×10^{18} protons.

credible If a source is credible it will have reliable results which are unbiased.

crest The highest part or point of maximum amplitude of a transverse wave.

cross wind A wind that blows across the direction of motion.

current The net flow of electric charge. Current is measured in amperes (A), where $1\,\text{A} = 1\,\text{C}\,\text{s}^{-1}$. By convention, electric current is assumed to flow from positive to negative.

cyclotron A particle accelerator device that accelerates particles outward from the centre of their trajectory along a spiral path.

D

data analysis The processes used to find trends and gain meaning from measurements or observations collected during an investigation.

daughter nucleus A nucleus on the product side of a nuclear equation that results when a nucleus undergoes fission or radioactive decay.

de Broglie wavelength Wavelength associated with a particle due to it motion.

decay series A sequence of radioactive decays that results in the formation of a stable isotope.

dependent variable The variable which is to be measured.

design speed Relating to a banked track, the speed at which a vehicle experiences no sideways force as it travels around a track. It is dependent on angle.

destructive interference The process in which two or more waves of the same frequency combine to cancel each other out. The amplitude of the resulting wave is equal to the difference between the amplitudes of the superimposed waves.

diffraction A deviation in the direction of a wave at the edge of an obstacle or through a gap in its path.

diffraction pattern The pattern of dark and light bands that is seen when light passes through a single small gap. Areas of constructive interference appear as bright bands and areas of destructive interference appear as dark bands.

dimension Space can be considered to consist of three length dimensions. These length dimensions are arranged at degrees to each other with their point of intersection being the origin. The position of an object can be defined in relation to its position along each of the three dimensions. Typically, these three dimensions are labeled x, y and z. However, up-down, left-right and backward-forward are also appropriate.

dimensional analysis Using the units in a graph or formula to check that the derived term is correct.

dipole Two electric charges or magnetic poles that have equal magnitudes but opposite signs, usually separated by a small distance.

direct current A continuous electric current that flows in one direction only, without substantial variation in magnitude. Batteries are a source of direct current. Abbreviated to DC.

direction conventions Direction conventions are standardised systems for describing the direction in which an object is travelling. The use of cardinal points of a compass (N, S, E and W) is an example of a direction convention.

discharge tube A tube of gas where ionisation occurs due to the presence of an electric field. This causes the gas to emit light at a particular frequency.

discrete variable A variable that can have only certain values. For example, the number of individuals in a population can only be whole numbers.

dispersion The process of light splitting into its component colours to create a spectrum or rainbow.

displacement The change in position of an object in a given direction. Displacement is a vector quantity.

distance travelled How far an object travels during a particular motion or journey. Distance is a scalar value. Direction is not required when expressing magnitude. It is measured in metres (m).

Doppler effect A change in the observed frequency of a wave, such as sound or light, that occurs when the source and observer are in motion relative to each other.

driving frequency The frequency that an object is exposed to. When the driving frequency equals the resonant frequency, resonance occurs in the object.

E

effective resistance The total resistance of a circuit.

elastic potential energy Stored energy in a stretched or compressed material, measured in joules (J).

electric circuit A continuous conducting loop connected to an energy source that allows electric current to flow.

electric field A region around a charged particle in which a force is exerted on other charged particles or objects.

electric field strength A measure of the force per unit charge on a charged object within an electric field, with the units $\text{N}\,\text{C}^{-1}$. Field strength can also be a measure of the difference in electrical potential per unit distance, with the units $\text{V}\,\text{m}^{-1}$.

electrical potential Potential energy due to the concentration of charge in part of an electric circuit.

electricity A form of energy resulting from the existence of charged particles (electrons or protons). Electricity is fuelled by the attraction of particles with opposite charges and the repulsion of particles with the same charge.

electromagnet A magnet consisting of an iron or steel core wound with a coil of wire, through which a current is passed. The core only becomes magnetised when current is flowing.

electromagnetic induction The creation of an electric current, or an emf, in a loop of wire as the result of changing the magnetic flux through the loop.

electromagnetic radiation Energy emitted in continuous waves with two transverse, mutually perpendicular components: a varying magnetic field and a varying electric field.

electromagnetic spectrum The entire range of electromagnetic radiation. At one end of the spectrum are gamma rays, which have the shortest wavelengths and high frequencies. At the other end are radio waves, which have the longest wavelengths and low frequencies. Visible light is near the centre of the spectrum.

electron A negatively charged particle in the outer region of an atom; it can move from one object to another, creating an electrostatic charge. When electrons move in a conductor, they constitute an electric current.

electron flow The net flow of electrons. Although electric current is assumed to flow from the positive terminal to the negative terminal, electrons physically move from the negative terminal to the positive terminal.

electron gun Uses a heated cathode to produce an electron beam and a series of charged plates to accelerate the beam.

electronvolt Amount of energy equal to the charge of an electron multiplied by 1 volt, i.e. $1\,\text{eV} = 1.6 \times 10^{-19} \times 1 = 1.6 \times 10^{-19}\,\text{J}$. An alternative to the joule as a unit in which to measure energy.

electrostatic force The force between electrically charged particles or objects due to their charge; repulsion between like charges and attraction between unlike charges.

elementary charge The charge carried by a single proton, $1.602 \times 10^{-19}\,\text{C}$.

emf The electromotive force (known as the emf) is a source of energy that can cause a current to flow in an electrical circuit or device.

emission spectrum Spectrum of coloured lines in the positions of the wavelengths of light emitted when a gas is heated or has an electric current passed through it. This is related to the absorption spectrum of the gas.

emit Energy in the form of heat, light, radio waves etc. radiated from a source.

energy levels The orbital levels in which electrons orbiting the nucleus of an atom can remain stable.

equilibrium Equilibrium exists when the vector sum of all forces acting on an object results in a zero net force acting on the object.

escape velocity The velocity required for an object to escape a gravitational field.

excited state Higher energy state of an atom above the ground state ($n > 1$).

F

Faraday's law Law stating that the average emf generated in a coil is proportional to the rate of change of magnetic flux and the number of turns in the coil.

fermion A fermion can be an elementary particle, such as those that make up atoms, e.g. quarks and electrons, or it can be a composite particle, such as protons and neutrons. Bosons are particles that are not fermions. According to the spin-statistics theorem, particles with integer spin are bosons, while particles with half-integer spin are fermions.

field A region of space where objects experience a force due to a physical property related to the field.

field lines A two-dimensional graphic representation of a field, using arrows to indicate the direction of the field. The closer the field lines, the stronger the field.

fission When a nucleus splits into two or more pieces, usually after bombardment by neutrons.

fission fragments Nuclides formed during nuclear fission; these are usually radioactive.

force A vector quantity which measures the magnitude and direction of a pull or push. Force is measured in newtons (N).

force arm The perpendicular distance between the axis of rotation and the line of action of the force.

frame of reference A coordinate system that is usually fixed to a physical system that contains an object and/or an observer. There can be frames of reference within other frames of reference.

freefall The motion of a falling body under the effect of gravity only. No air resistance or propulsive forces are acting.

frequency The number of vibrations (or cycles) that are completed per second or the number of complete waves that pass a given point per second, measured in hertz (Hz).

friction A force that resists the direction of motion.

fusion A process taking place inside stars in which small nuclei are forced together to make larger nuclei. Energy is released in the process.

G

gamma ray High-energy electromagnetic radiation ejected from the nucleus of a radioactive nuclide.

gauge boson Gauge bosons are force-carrier particles which, according to the Standard Model of particle physics, mediate the four fundamental forces.

Geiger counter A device for measuring radioactive emissions.

generator An electrical device that converts kinetic energy into direct current (DC) electricity. Usually, a coil is rotated causing it to cut across a magnetic field.

geostationary satellite Satellite that remains in orbit above the same place on the Earth's surface. It has the same period as the Earth's rotation, i.e. 24 hours. Only occurs at an altitude of 36 000 km above the Earth.

giant A very large, bright non-main sequence star. Supergiants are the very largest stars, being thousands of times brighter than the Sun but with a much shorter lifetime due to the faster rate of fusion.

gluon Gluons are elementary particles that act as exchange particles for the strong nuclear force between quarks, similar to the exchange of photons in the electromagnetic force between two charged particles. Gluons themselves carry the colour charge of the strong interaction. Gluons can be considered to be the fundamental exchange particle underlying the strong interaction between protons and neutrons in a nucleus.

gravimeter Sensitive instrument used by geologists to detect small variations in gravitational field strength.

gravitational constant, *G* Universal constant of value $6.67 \times 10^{-11}\,\mathrm{N\,m^2\,kg^{-2}}$.

gravitational field The region around an object where other objects will experience a gravitational force.

gravitational field strength The strength of gravity, usually measured at the surface of a planet. Equivalent to the acceleration due to gravity, g. Measured in newtons per kilogram ($\mathrm{N\,kg^{-1}}$).

gravitational force The force of attraction acting between two objects that have mass.

gravitational potential energy The energy that a body possesses due to its position in a gravitational field. A scalar quantity that is measured in joules (J).

ground state Lowest energy state of an atom ($n = 1$).

H

hadron A composite particle that contains quarks held together by the strong force. Hadrons are subdivided into two families: baryons (e.g. the proton and neutron) and mesons (the pion and kaon).

half-life The time taken for half of the nuclei of a radioactive isotope to decay.

hard X-rays High-energy X-rays with a wavelength less than 0.1 nm and energy values greater than 10 keV. They have high penetrating capacity and frequency. They are commonly used for radiation therapy purposes.

heat exchanger Part of a nuclear reactor where heat drawn from the reactor core is used to turn water into steam.

heavy water Water that has a higher than normal proportion of water molecules that contain deuterium.

Heisenberg's uncertainty principle Concept that any measurement of a system creates a disturbance of the system with a resulting uncertainty in the measurement.

Hertzsprung–Russell (H–R) diagram A plot of the luminosity of stars against surface temperature that classifies stars by types.

Higgs boson Elementary particle discovered in 2012 at CERN that essentially gives mass to all elementary particles.

Hubble constant The unit of measurement used in Hubble's law, which describes the expansion of the universe. It has a value of around $70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

Hubble's law A law created by Edwin Hubble that states that the rate at which astronomical objects in the universe move apart from each other is proportional to their distance from each other.

hydrostatic equilibrium For the majority of the life of a star, the gravitational force from the mass of the star is in balance with the gas pressure due to energy generation in the core of the star. The star is 'hydrostatic equilibrium' during this phase of its life.

hypothesis A proposed explanation for an observed phenomenon.

I

ideal transformer Where the input power and the output power are equal and the transformer is 100% efficient. Real transformers obtain close to this value.

incandescent Emission of light due to very high temperature.

independent variable A variable that is varied during an experiment to test the effect on a dependent variable.

induced current Electric current produced by changing a magnetic flux in the region of a conductor or by moving the conductor in a magnetic field.

inertia A property of an object, related to its mass, that opposes changes in motion.

inertial frame of reference A frame of reference that is either moving with a constant velocity or is stationary. It is not accelerating.

inflation Period of time in the early universe that lasted 10^{-24} seconds where the size of the universe expanded to around 10^{50} times its original size.

inquiry question A question that defines the focus of an investigation.

insulator A material or an object that does not easily allow heat, electricity, light or sound to pass through it. Air, cloth and rubber are good electrical insulators; feathers and wool are good thermal insulators.

intensity A measure of the energy transmitted by a wave or radiation.

intrinsic brightness The actual brightness of a star, regardless of the distance from the observer.

interference The variation of wave amplitude that forms when two or more waves of the same or different frequencies come together. The amplitude of the resulting wave will be either larger or smaller than the amplitude of the individual waves, depending on whether or not their peaks and troughs match up. If the peaks and troughs match up, this is called constructive interference. If the peaks and troughs of the individual waves do not match up, he resulting amplitude is smaller. This interference is called destructive interference.

International System of Units The most commonly used system of measurement, including, for example, the metre, kilogram and Newton. These are known as SI units.

inverse square law A physical law in which some quantity (e.g. gravitational force) is inversely proportional to distance squared; for example: $F_g \propto \frac{1}{r^2}$.

ion An atom that has gained or lost electrons to become positively or negatively charged.

ionised To remove or add an electron from an atom, after which it becomes positively or negatively charged respectively.

ionising radiation Radiation with enough energy to alter the molecular structure of matter by displacing one or more electrons from an atom and thus creating electrically charged ions.

isolated system Situation where there should only be internal forces acting between the objects and no interaction with objects outside the system.

isotope Atoms with the same number of protons but with different numbers of neutrons.

K

kelvin An absolute temperature scale based on the triple point of water.

kinetic energy The energy of a moving body; measured in joules. Kinetic energy is a scalar quantity.

L

Large Hadron Collider The world's largest and most powerful particle accelerator, located at CERN laboratories near Geneva, Switzerland. It is a 27 km highly evacuated tube in which particles are accelerated to 99.999999% of the speed of light. They are held in place in the ring by a huge array of superconducting magnets with a number of accelerating structures to boost the energy of the particles along the way.

laser Source of a narrow beam of intense, monochromatic, polarised, coherent radiation.

length contraction Length in a moving frame of reference appears shorter when viewed by a stationary observer.

Lenz's law A law stating that the direction of the induced current in a conductor is such that its associated magnetic field opposes the change in flux that caused it.

lepton Leptons are the six fundamental particles of the Standard Model that do not experience the strong nuclear force. Examples are the electron, the muon and neutrinos. The charged leptons experience the electromagnetic force, but the neutral neutrinos do not.

light-emitting diode (LED) Semiconductor diode that uses the excitation of electrons to emit light.

linac Abbreviation for 'linear accelerator'. These are subatomic particle accelerator devices that are used to generate beams of high-energy X-rays to be used in radiation therapy for cancer treatment.

line of action of the force The line along which a force is acting. The line of action extends forwards and backwards from the force vector.

linear accelerator Type of particle accelerator in which particles are accelerated in a straight line.

literature review A summary of current published knowledge in a particular area.

longitudinal Extending in the direction of the length of something; running lengthwise, i.e. a wave vibration travels along the same direction as the wave.

Lorentz factor $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$

Lorentz force The force experienced by a point charge moving along a wire that is in a magnetic field; the force is at right angles to both the current and the magnetic field. Named for the Dutch physicist who shared a 1902 Nobel Prize for researching the influence of magnetism on radiation.

luminosity The absolute brightness of a star measured in watts of total energy output.

M

magnetic Of or relating to magnetism or magnets. Having the properties of a magnet. Capable of being magnetised or attracted by a magnet.

magnetic field A magnetic field is an area influenced by a magnet or something with the properties of a magnet.

magnetic flux The strength of a field in a given area expressed as the product of the area and the component of the field strength at right angles to the area (i.e. $\Phi = BA$).

magnetic flux density Amount of magnetic flux per unit area. In other words, it describes 'the closeness of magnetic field lines'. Same as magnetic field strength.

magnetism Magnetism is a physical phenomenon caused by magnets that results in a field that attracts or repels other magnetic materials.

magnitude The size or extent of something. In physics, this is usually a quantitative measure expressed as a number of a standard unit.

main sequence A group of stars lying on a line running from the top left to the bottom right of the Hertzsprung–Russell diagram.

mass defect The difference between the mass of an atom and the individual masses of its constituent parts, i.e. protons, neutrons and electrons. This is equivalent to the binding energy of the atom. This energy difference is harnessed during fusion reactions and is calculated using Einstein's equation, $E = mc^2$.

mass number The number of nucleons (protons and neutrons) in a nucleus.

mean Equal to the average of a set of data.

mechanical energy The energy that a body possesses due to its position or motion. Kinetic energy, gravitational energy and elastic potential energy are all forms of mechanical energy.

mechanical wave A mechanical wave is a wave that propagates as an oscillation of matter and therefore transfers energy through a medium.

median The middle number for a set of data.

medium A physical substance, such as air or water, through which a mechanical wave is propagated.

meson Mesons are unstable subatomic particles composed of one quark and one antiquark. They are in the hadron family—particles made of quarks. Baryons—subatomic particles composed of three quarks—are also part of the hadron family.

metal Material in which some of the electrons are only loosely attracted to their atomic nuclei.

metal vapour lamp Lamp that contains a low-pressure gas that becomes excited and emits photons with the colour characteristic of the element in the gas, e.g. sodium vapour lamp.

mistake An error which can be avoided.

mnemonic A mnemonic device is any learning technique that aids information retention. Mnemonics aim to translate information into a form that the brain can retain better than its original form. Even the process of learning this conversion might aid in the transfer of information to long-term memory.

mode A value that appears the most amount of times within a data set.

model Representations of structures or processes, such as physical models or digital models, that are used to create and test theories and explain concepts.

momentum Momentum, p, is the product of an object's mass, m, and its velocity, v. Objects with larger momentum require a larger force to stop them in the same time that an object with smaller momentum takes to stop.

monochromatic Light of a single colour, e.g. red light.

monopole A single mass or point electric charge. A mass is considered to be a monopole at its centre of mass. Magnetic poles only exist, as far as we currently know, as dipoles.

N

natural satellite A body such as the Moon or a planet (not made by humans) that is in orbit around another body.

nebula An interstellar cloud in outer space composed of dust and gasses. The Latin word for cloud is nebula.

net charge When the number of positive and negative charges in an object is not balanced.

net force The sum of all forces acting on an object.

neutral Carrying no net electric charge; a situation in which positive and negative charges are balanced.

neutrino An almost massless neutral particle released during some nuclear reactions.

neutron An uncharged subatomic particle.

neutron bombardment A physical process by which a stable atomic nucleus is bombarded with high-speed ions or neutrons inside a nuclear reactor. As a result, the atomic structure of the original nucleus changes and it becomes a different element.

neutron star The remnant of a supernova, consisting entirely of neutrons.

newton The SI unit for force (N).

Newton's first law An object will maintain a constant velocity unless an unbalanced, external force acts on it.

Newton's law of universal gravitation Law that states that the attractive gravitational force between two masses is directly proportional to the product of their masses and inversely proportional to the square of the distance.

Newton's second law Force is equal to the rate of change of momentum. This can be processed mathematically to: the acceleration of an object is directly proportional to the force on the object and inversely proportional to the mass of the object.

Newton's third law For every action (force), there is an equal and opposite reaction (force).

nominal variable A categorical variable in which there is no inherent order. Nominal variables can be counted but not ordered

non-contact force A force applied to an object by another body without any direct contact.

non-ionising radiation Radiation that does not have enough energy to break the molecular bonds within molecules and to alter the number of electrons in an atom. Lower forms of energy in the electromagnetic spectrum such as radio waves, microwaves, visible light and UVA radiation are non-ionising.

normal reaction force Force with which a surface pushes back on an object, at right angles to the surface. Same magnitude as the apparent weight of an object. Symbol $\vec{F}_N$ or $\vec{N}$, measured in newtons.

nuclear fission reactor A device built to control and harness the energy from nuclear fission reactions. Nuclear power stations are one type of nuclear reactor; other uses include medical, industrial, weapons and research purposes. Australia has one nuclear reactor (OPAL) at Lucas Heights in Sydney.

nuclear transmutation The changing of one element into another.

nucleon A particle located in the nucleus of an atom.

nucleosynthesis Process of fusion that results in the formation of the nuclei of heavier elements.

nucleus The central part of an atom.

nuclide The range of atomic nuclei associated with a particular atom, which is defined by its atomic number, and the various isotopes of that atom as identified by the mass number.

O

ohmic Behaving according to Ohm's law; resistance of the material is constant regardless of the applied potential difference.

orbit The circular path an object takes around a central body; for example, the satellite motion of the moon around the earth or the path an electron takes around a nucleus in Rutherford's atomic model.

orbital Unlike the orbits of electrons predicted by Rutherford and Bohr, an orbital is the complex, three-dimensional region of space that an electron occupies. The orbital is represented by a wave function.

ordinal variable A categorical variable in which there is an inherent order. Ordinal variables can be counted as well as ordered.

outlier A value that lies outside the main group of data of which it is a part.

overtone Any of the higher-level harmonics, except for the fundamental frequency.

P

pair production The creation of a particle and its antiparticle. This is commonly the result of two photons interacting or a photon interacting with an atomic nucleus.

paradox A situation that appears to have contradictory elements.

parallax movement The apparent movement of the closer stars relative to the background stars which is actually due to the motion of the Earth around the Sun.

parent nucleus A nucleus on the reactant side of a nuclear equation that when struck by a neutron undergoes fission or simply decays by natural means.

parsec The distance to a star that has a parallax angle of one arcsecond; equal to 206 265 AU.

particle accelerator A machine that can accelerate a charged particle (proton, electron) or an atomic nucleus to very high speeds, including speeds that approach the speed of light.

particle radiation Refers to a form of energy that is given off by atoms, in the form of small, high-energy subatomic particles, such as alpha particles, beta particles and positrons.

path difference The difference in the lengths of the paths from each slit to the screen in a double-slit experiment.

penetrating ability A measure of how easily radiation passes through matter.

percentage uncertainty The absolute uncertainty divided by the measurement, expressed as a percentage.

perigee The closest point in an elliptical orbit between a central mass and an orbiting body.

perihelion The point in an elliptical orbit that is closest to the Sun.

period The time interval taken to complete one cycle of a regularly repeating phenomenon, such as a rotating object or in a soundwave. The SI unit for period is seconds (s).

personal protective equipment (PPE) Clothing that should be worn to minimise risk in an investigation. Examples include eye goggles or a laboratory coat.

persuasion A style of writing which aims to persuade the audience. Scientific reports are generally written in an objective style and avoid using persuasion.

phenomenon A fact which is described scientifically.

photocurrent Current caused by the flow of photoelectrons during the photoelectric effect.

photoelectric effect Spontaneous emission of electrons from a metal surface when it is illuminated by light of particular frequencies and energies.

photoelectron An electron released from an atom due to the photoelectric effect.

photon Packet or bundle of electromagnetic radiation (light).

photosphere A thin layer from where a star's visible light is emitted.

photovoltaic cell A cell that converts visible light from the Sun into direct current electricity.

pie diagram A circular diagram divided into sectors, with each sector representing the value of one set of data as a proportion of the total data set.

pivot point A point about which an object can rotate.

plane wave A constant frequency wave with wavefronts that are infinite parallel lines or planes.

planetary nebula The usually ring-shaped nebula formed by an expanding shell of gas around an ageing star. These are basically layers formed as a result of the various stages of fusion reactions and may involve a considerable proportion of the star's mass.

point charge An ideal situation in which all of the charge on an object is considered to be concentrated at a single point. The point size is negligible in relation to the distance between it and another point charge.

polarisation The phenomenon in which transverse waves are restricted in their direction of vibration.

pole The north-seeking magnetic pole of a magnet. The north pole of a freely suspended magnet is attracted to the Earth's geographic North Pole (a magnetic south). The south pole is attracted to the Earth's geographic South Pole (a magnetic north).

position The location of an object with respect to a reference point. Position is a vector quantity.

positron The antimatter pair of the electron. It shares the same mass as an electron but has opposite properties such as electromagnetic charge and spin.

postulate A suggestion that is put forward as a fact as a basis for further discussion or reasoning.

potential difference The difference in electric potential between two points in a circuit, measured by a voltmeter when placed across a circuit component. A battery creates the potential difference across a circuit, which drives the current.

potential energy Energy that can be considered to be 'stored' within a body due to its position, composition or molecular arrangement.

power The rate at which work is done; a scalar quantity measured in watts (W).

precision The ability to consistently obtain the same measurement. To obtain precise results, you must minimise random errors.

primary source Original sources of data or evidence generated personally.

projectile Object moving freely through the air without an engine or power source driving it.

proper length A measurement of length made from the frame of reference in which the object being measured is stationary.

proper time A measurement of time made with a clock that doesn't move relative to the point at which the start and end of the event occurs.

proton A positively charged subatomic particle.

protostar A collapsing cloud of gas that will eventually become a star.

pulsar A celestial object that is thought to be the rapidly rotating neutron star left after the explosion of a giant star. It emits regular pulses of radio waves and other electromagnetic radiation as it revolves at rates of up to 1000 pulses per second.

pulse A short burst; for example, one wave.

purpose A statement describing in detail what will be investigated.

Q

qualitative variable Variables that can be observed but not measured. Examples include types of animals and brightness.

quantitative variable Variables that can be measured. Examples include wavelength and temperature.

quantum Plural quanta. According to the quantum model, electromagnetic radiation is emitted from objects as discrete packets called quanta. Each quantum has an energy proportional to its frequency according to the equation $E = hf$.

quantum mechanics Area of modern particle physics where the wave properties of electrons are studied.

quantum number A number that describes a property of a particle; for example, charge or spin.

quark Quarks are six of the fundamental particles in the universe. They cannot be found individually and only exist bound by the strong force within hadrons.

R

radian The measure of a central angle subtending an arc equal in length to the radius; 1 full revolution is equivalent to 2π radians.

radiation The process by which energy is emitted by one object or system, transmitted through an intervening medium or space, and absorbed by another object or system.

radioactive Something that spontaneously emits radiation in the form of alpha particles, beta particles and gamma rays.

radioisotope An isotope of a chemical element that emits radioactivity due to its unstable combination of neutrons and protons in the nucleus.

random error Errors in measurement that occur in an unpredictable manner.

raw data Data which has yet to be processed.

ray The straight line path of a wave. (Also a narrow beam of light.)

ray diagram Ray diagrams are created to follow the path of light rays as they interact with either a mirror or a lens.

redshift The change (lowering) of frequency that occurs in any wave phenomenon as the source of the waves moves away from the observer. Likewise, a blueshift occurs if the source is moving towards the observer.

reference point A point about which a rotational equilibrium can be calculated for an object that is in static equilibrium. This point can be anywhere, but is best selected to cancel the torque of an unknown force in a problem.

refraction The bending of the direction of travel of a ray of light, sound or other wave as it enters a medium of differing density.

reliability The consistency of the results received from an experiment or collection of data. Reliable results are also repeatable, meaning another scientist performing the same analysis will come up with the same results.

reputation Describes how somebody is known or perceived.

resultant One vector that is the sum of two or more vectors.

rhetoric Style of communicating that aims to be persuasive.

right-hand grip rule Used to find the direction of a magnetic field induced around a current-carrying wire.

right-hand rule Tells us the force (palm) on a current (thumb) in a magnetic field (straight fingers).

root mean square The square root of the arithmetic mean of the squares of the numbers in a given set of numbers. In terms of alternating power, the root mean square value, $P_{rms} = \frac{P_{peak}}{2}$. Alternatively, it is the effective mean (average) value of an AC supply.

rotational equilibrium A situation in which the sum of all the clockwise torques is equal to the sum of all the anticlockwise torques.

S

satellite Object in a stable orbit around a central body. Could be natural, like a planet, or artificial, like a communications satellite.

scalar A physical quantity that is represented by a magnitude and units only. Mass, time and speed are examples of scalar quantities.

scatter plot A graph in which two variables are plotted as points. The *x* coordinate of a particular point is one measured value of the independent variable and the *y* coordinate is the corresponding measured value of the dependent variable.

scientific method The process scientists apply to construct theories to explain practical observations.

scientific theory The theory that results from a tested and supported hypothesis, as per the scientific method.

secondary source Outside sources of data or evidence such as from other people's scientific reports, textbooks or magazines.

significant figures The number of digits used. For example, 5.1 has two significant figures, whereas 5 has just one significant figure.

simultaneous When two events occur at exactly the same time.

sinusoidally Anything that varies in the form of a sine wave.

slip rings Components of alternators (AC generators) that allow a constant electrical connection to be made between the rotating armature and the static external circuit through which the generated alternating current flows.

soft X-rays X-rays with wavelengths greater than 0.1 nm. They have lower energy and penetrative power compared to hard X-rays. They are used mostly for diagnostic imaging purposes.

solar wind A continuous stream of charged particles ejected by the Sun. It consists mostly of protons and electrons and has enough energy to escape the Sun's gravitational field, at speeds ranging from about 300 to 800 kms^{-1}, which allows it to reach the Earth in about 3.9 days. The speed and intensity of the charged particles depend on magnetic activity at different regions of the Sun.

solenoid A coil of wire that acts as an electromagnet when electric current is passed through it due to the magnetic field that is set up by the current passing through it.

spacetime A term used to describe the situation in which the three-dimensional space coordinate system (x, y, z) is linked to the one-dimensional time system.

spectroscope A device, optical or digital, for producing and recording spectra from sources of electromagnetic radiation for examination.

spectroscopy Spectroscopy is an area of study that describes the interaction of all forms of electromagnetic radiation with matter, including refraction, diffraction and absorption. This information can be further utilised to identify and quantify different types of matter.

speed The ratio of distance travelled to time taken. Speed is a scalar quantity.

split ring commutator A component of DC generators and motors that typically resembles a ring that has been cut into two equal pieces or shells. Each part of the ring has a fixed connection to the ends of the coil, while also making contact with the stationary brushes. This means the connection between the rotating coil and the static circuit is reversed every half turn, which ensures the direction of current in the circuit is constant (in the case of the generator) or the direction of rotation is constant (in the case of the motor).

Standard Model The Standard Model of particle physics is a mathematical description of all known particles and three of the forces acting on them. It is currently the most successful theory for predicting the behaviour and properties of the particles that exist in nature.

standing wave Also called a stationary wave, the periodic disturbance in a medium resulting from the combination of two waves of equal frequency and intensity travelling in opposite directions.

stator A portion of a machine that remains stationary with respect to rotating parts, especially the collection of stationary parts in the magnetic circuits of a motor or generator.

steady state theory A model of the universe based on an idea called the perfect cosmological principle. This states that the universe on the largest scales looks essentially the same everywhere at all times. Therefore, the universe maintains the same average density of matter forever. The steady state theory is not the accepted model of the universe.

stellar parallax The difference in direction of a celestial object as seen by an observer from two widely separated points. The two points generally used coincide with the position of Earth at opposite sides of the Sun during its annual rotation since these are the most widely separated locations available to astronomers.

step-down transformer Device that decreases the secondary voltage compared to the primary voltage.

step-up transformer Device that increases the secondary voltage compared to the primary voltage.

stopping voltage The applied voltage required to stop all photoelectrons from reaching the collector electrode. For a particular frequency of incident light on a particular metal, the stopping voltage is a constant.

storage ring Part of a synchrotron where the electrons are held for a long period of time while they continuously emit synchrotron light.

strong nuclear force A short-range but powerful force of attraction that acts between all the nucleons in the nucleus. The strong nuclear force acts on quarks and binds them together in hadrons. It also acts at larger distances to bind protons and neutrons together within atomic nuclei. In the Standard Model of particle physics the strong force is described by quantum chromodynamics and is mediated by an exchange of gluons.

subcritical mass A quantity of fissile material that is too small to sustain a chain reaction.

supercritical mass A quantity of fissile material that is large enough to sustain a chain reaction.

supernova A giant explosion that occurs when a star many times larger than our Sun runs out of nuclear fuel.

synchrotron Large particle accelerator in a circular shape producing a very intense, very narrow beam of electromagnetic radiation called synchrotron light.

systematic error Errors that are consistent and will occur again if the investigation is repeated in the same way.

T

tangential Describes a direction forming a tangent to a curve.

temperature The amount of heat in an object.

terminal velocity Velocity at which the force of drag is equal to the weight of an object.

thermal energy A form of energy transferred as a result of a difference in temperature or average kinetic energy within a system.

threshold frequency The minimum frequency of electromagnetic radiation for which the photoelectric effect can occur for a given material.

time dilation When one observer watches events in a frame of reference that is moving (very fast) relative to him/her, time in that frame of reference will appear to go more slowly. People in the moving frame do not experience any difference in the rate at which time passes. This effect is one of the strange consequences of Einstein's theory of special relativity.

torque Any force or system of forces that causes or tends to cause rotation.

torsion balance Device used to measure very small twisting forces. Cavendish used this device to measure the force of attraction between lead balls held a small distance apart.

transducer A device that receives a signal in the form of one type of energy and converts it into another form of energy.

transfer The conversion of energy from one system to another.

transform To change from one thing to another; for example, to change energy from electrical potential energy to kinetic energy.

transformer A device that transfers an alternating current from one circuit to one or more other circuits, usually with an increase (step-up transformer) or decrease (step-down transformer) in voltage. The input goes to a primary coil; the output is taken from a secondary coil or windings linked by induction to the primary coil.

transmit To cause light, heat, or sound etc. to pass through a medium.

trend The general orientation. A trend line can be added to a graph to describe the way data is developing.

trend line A line drawn on a graph to show the general relationship between the independent and dependent variables.

U

uncertainty A description of the range of data obtained; calculated as $\pm \frac{x_{max} - x_{min}}{2}$.

uniform Constant, unvarying.

units Properties related to physical measurements. Units can be fundamental like metres (m), seconds (s) or kilograms (kg). Units can also be derived by combining fundamental units like metres per second ($m s^{-1}$).

V

validity The reasonableness of the results received from an experiment or collection of data. Valid results meet all the requirements of the criteria of the scientific method.

variable A factor or condition that can change.

vector A physical quantity that requires magnitude, units and a direction in order to be fully defined. Velocity, acceleration and force are examples of vector quantities.

vector notation Vector notation within this course is shown by adding an arrow above the variable. For example, the variable for velocity will be written as $\vec{v}$.

velocity The ratio of displacement to time taken. Velocity is a vector quantity.

voltaic pile An early form of battery consisting of a pile of paired plates of dissimilar metals, such as zinc and copper, each pair being separated from the next by a pad moistened with an electrolyte (mild acid).

voltmeter A device used to measure the voltage change between two points in a circuit.

volts Volts (V) are the unit of electrical potential. One volt is equal to one joule of potential energy given to one coulomb of charge in a source of potential difference. The voltage (V), or number of volts, is another name for the potential difference.

W

wave front The set of points reached by a wave or vibration at the same instant as the wave travels through a medium. Wave fronts generally form a continuous line or surface.

wavelength The distance between one peak or crest of a wave of light, heat or other energy and the next corresponding peak or crest. Represented by the symbol λ.

wave–particle duality The theory that, in some experiments, light and matter behave like waves and, in other experiments, they behave like particles.

weak nuclear force The weak nuclear force is the interaction between quarks responsible for changing from one type of quark to another. This force is described by electroweak theory and is responsible for radioactive decay and nuclear fission. In the Standard Model of particle physics this force is mediated by the W and Z bosons.

weight The force of attraction on a body due to gravity.

white dwarf A small, very dense star around the size of a planet. A white dwarf is formed when a low-mass star, such as our Sun, exhausts all its fuel and loses its outer layers. It will eventually cool to become a black dwarf as no nuclear fuel remains to generate additional heat and energy.

work The transfer of energy as a result of the application of a force; measured by multiplying the force and the displacement of its point of application along the line of action. Measured in joules (J).

work function The energy required to remove an electron from its state of being bound to an atom; measured in joules or electronvolts.

X

X-ray diffraction X-rays diffract from a crystalline structure to produce a particular pattern. Analysis of this pattern reveals information about the spacing of atoms in the sample.

Index

Attributions

Cover image: Eric Heller/Science Photo Library.

The following abbreviations are used in this list: t = top, b = bottom, l = left, r = right, c = centre.

123RF: 1xpert, p. 112; Pashkov Andrey, p. 100t; claudiodivizia, p. 275; eshma, p. 174cl; Pawel Gaca, p. 100c; Patricia Hofmeester, p. 316; jewhyte, p. 69tl; lianem, p. 97cr; Felix Lipov, p. 300; manganganath, p. 306; verdateo, p. 94tl.

AAP: Adam Davy/PA, p. 82.

Age Fotostock: David Nunuk, p. 395.

Alamy Stock Photo: Archive Pics, pp. 285r, 381l, 381r; blickwinkel/McPHOTO/Bioquatic, p. 369; ClassicStock, p. 327; dpa picture alliance, p. 440; GL Archive, p. 215; The Granger Collection, p. 264l; Granger, NYC/Granger Historical Picture Archive, p. 466; Henglein and Steets/Cultura Creative (RF), p. 48–9; imageBROKER, p. 199; INTERFOTO, p. 410; Carsten Leuzinger/imageBROKER, p. 237, 334–7; NASA/RGB Ventures/SuperStock, p. 105; Photos 12, p. 295; Photo Researchers/Science History Images, pp. 338–9, 473–8; Photo Researchers, p. 437br; PJF Military Collection, pp. 61bl, 178; PRISMA ARCHIVO, p. 119br; Sacramento Bee/ZUMA Press Inc, p. 75bl; Scherl/Sueddeutsche Zeitung Photo, p. 287; Stocktrek Images, Inc., p. 144–145; VintageCorner, p. 106tr.

ANSTO: pp. 157tl, 161, 461.

Cross, Malcolm: pp. 266cr, 266tl, 268.

DK Images: Dorling Kindersley, p. 266c.

EcoSwitch®: p. 203.

European Space Agency: pp. 119, 126; AOES Medialab, p. 118cr.

European Southern Observatory: ESO/Digitized Sky Survey 2, p. 350.

Fotolia: BillionPhotos.com, p. 251b; Generalfmv, p. 416; Michael Ireland, p. 147; oceandigital, p. 347; tadamichi, p. 24; sumire8, p. 285b.

Getty Images: Bettmann, pp. 292, 434br, 434c; Gianni Tortoli, p. 432; Quinn Rooney, p. 78; Stocktrek images, p. 366br.

Hecker, Richard: p. 243.

Martayan, Dr. Christophe: p. 257.

NASA: Don Davis, Acrylic on board for NASA, JPL., pp. 133, 244b; ESA/Hubble Telescope/NASA/courtesy of nasaimages.org, p. 124; ISS Expedition 12 Crew/NASA/courtesy of nasaimages.org, p. 123; JPL, p. 110; JPL/University of Arizona/University of Idaho, p. 252; courtesy of nasaimages.org, pp. 113, 116cr, 327; Johnson Space Centre, pp. 142–3, 231–6; X-ray: NASA/CXC/PSU/L. Townsley et al, Optical: UKIRT, Infrared: NASA/JPL-Caltech, pp. 340–1; NASA and WMAP Science Team, pp. 348l, 348r.

Pearson Education Asia Ltd: Tsz-shan Kwok, p. 15bl.

Pniok, Heinrich: Photo of a deuterium discharge tube by Heinrich Pniok, p. 252.

Retrospect Photography: Dale Mann, p. 147.

S P Andrew Ltd: Portrait negatives, Ref: 1/2-043316-F, Alexander Turnbull Library, Wellington, New Zealand, p. 434bc.

Science Photo Library: p.320; American Institute of Physics, p. 434bl, 434cl; Carl Anderson, p. 387; Andrew Lambert Photography, pp. 189, 278, 299, 377l, 377r; Arscimed, p. 452; Atlas Collaboration, p. 459; Dr Tony Brain, p. 407; Andrew Brookes, National Physical Laboratory, p. 309; A. Barrington Brown, Gonville and Caius College, p. 434cr; Martyn F. Chillmaid, p. 11r; Carlos Clarivan, pp. 257, 397; CERN, p. 463t; Alex Cherney/Terrastro.com, pp. 46–7, 137–41; I. Curie & F. Joliot, p. 385t; Department of Physics, Imperial College, p. 393b; Luke Dodd, p. 358; Prof. George Efstathiou, p. 353; Equinox Graphics, p. 449tl; European Southern Observatory, p. 370; Mark Garlick, p. 117bl; Tony & Daphne Hallas, p. 343; Adam Hart-Davis, p. 51br; Eric Heller, p. i; James King-Holmes, pp. 468, 470; Dorling Kindersley, p. 251t; Russell Kightley, p. 269; Mehau Kulyk, p. 373; Lawrence Berkeley Laboratory, pp. 456, 463b; LMSAL/Stanford University/NASA, p. 247; Patrice Loiez, CERN, p. 390–1; Claus Lunau, p. 6; Walt Myers, p. 323; NASA/ESA/STSCI/J. Hester & P. Scowen, ASU, p. 366cl; NASA/ESA/STSCI/J. Hester & A. Loll, ASU, p. 371; NASA/ESA/STSCI/R. Windhorst & S. Pascarelle, ASU, p. 352; NASA Goddard Space Flight Center, pp. 238–9; NASA/JPL, p. 286; NASA/SDO/Goddard Space Flight Center, p. 253; NOAO, p. 362; N.A.Sharp, NOAO/NSO/KITT PEAK FTS/AURA/NSF, p. 249; David Parker, pp. 66–7, 404, 415, 447, 449br; Jose Antonio Peñas, pp. 365, 380; Physics Dept., Imperial College, p. 393t; Public Health England, p. 424; Detlev Van Ravensway, pp. 255, 259, 342; Harald Ritsch, p. 247; Royal Institution of Great Britain, p. 184; Royal Obersavatory, Edinburgh/AAO, p. 366tl; Science Photo Library, pp. 61br, 151, 264r, 420b; 437bl; Sputnik, p. 55cr; Take 27 Ltd., p. 299; Sheila Terry, pp. 277, 337; Trevor Clifford Photography, p. 169; US Department of Energy, p. 434; US Navy, p. 438; Mark Williamson, p. 267r.

Science Source - Photo Researchers, Inc: Charles D. Winters, p. 157bl; GIPhotoStock, pp. 271,273

Shutterstock: 270058, p. 37; AGorohov, p. 94bl, 94cl, 95, 96; Miss Kanithar Aiumla-Or, p. 201t; aniad, p. 322; Gunnar Assmy, p. 57b; Beverly Speed, p. 244t; Christian Delbert, p. 220; Designua, p. 356; Diyana Dimitrova, p. 280–1; Dinga, p. 174bl; Everett Historical, p. 420t; Filipe Frazao, p. 107; Konstantin Faraktinov, p. 385b; Iryna1, p. 106tl; Dmitry Kalinovsky, p. 99; saied shahin kiya, p. 246; koya979, p. 402; ktsdesign, pp. 198b, 375; langdu, p. 266cl; LittleMiss, p. 179; fuyu liu, p. 166–7; Joana Lopes, p. 29; muzsy, p. 69tr; Nikolay 53, p. 263; oneo, p. 15br; Claire Plumridge, p. 94cr; Pichitchai, p. 201b; Keith Publicover, p. 109; Fouad A. Saad, p. 9; Carlos E. Santa Maria, p. 250; SnvvSnvvSnvv, p. 212–3; solarseven, p. 240; ssuaphotos, p. 220; Surakit, pp. 97tr, 98; yelantsevv, p. 183; zstock, p. 294.

The Picture Source: Terry Oakley, p. 406.

Wikimedia Commons: Bielasko, p. 100b; Twostaricebox, p. 198t.

Vigan, Tyler: Spurious Correlations, tylervigan.com. Data sources: U.S. Department of Agriculture and Centers for Disease Control & Prevention, p. 24b.

Groups

Key:

- ATOMIC NUMBER → 12
- RELATIVE ATOMIC MASS ( ) indicates most stable isotope → 24.31
- SYMBOL → Mg
- BOILING POINT °C → 1090
- MELTING POINT °C → 650
- ELECTRONEGATIVITY → 1.3
- ELECTRON STRUCTURE → [Ne]$3s^2$
- Magnesium

s BLOCK (groups 1–2); *d* BLOCK (groups 3–12); *p* BLOCK (groups 13–18)

Each cell: atomic number, relative atomic mass / symbol, boiling point °C, melting point °C, electronegativity / electron structure / name.

Periods	1	2	3	4	5	6	7	8	9
1									1 1.008 H -252.9 -259 2.2 $1s^1$ Hydrogen
2	3 6.941 Li 1342 180 1.0 $1s^2 2s^1$ Lithium	4 9.012 Be 2468 1287 1.6 $1s^2 2s^2$ Beryllium							
3	11 22.99 Na 882.9 97.8 0.9 [Ne]$3s^1$ Sodium	12 24.31 Mg 1090 650 1.3 [Ne]$3s^2$ Magnesium							
4	19 39.10 K 758.8 63.4 0.8 [Ar]$4s^1$ Potassium	20 40.08 Ca 1484 842 1.0 [Ar]$4s^2$ Calcium	21 44.96 Sc 2836 1541 1.4 [Ar]$3d^1 4s^2$ Scandium	22 47.87 Ti 3287 1670 1.5 [Ar]$3d^2 4s^2$ Titarium	23 50.94 V 3407 1910 1.6 [Ar]$3d^3 4s^2$ Vanadium	24 52.00 Cr 2671 1907 1.7 [Ar]$3d^5 4s^1$ Chromium	25 54.94 Mn 2061 1246 1.6 [Ar]$3d^5 4s^2$ Manganese	26 55.85 Fe 2861 1538 1.8 [Ar]$3d^6 4s^2$ Iron	27 58.93 Co 2927 1495 1.9 [Ar]$3d^7 4s^2$ Cobalt
5	37 85.47 Rb 687.8 39 0.8 [Kr]$5s^1$ Rubidium	38 87.61 Sr 1377 769 1.0 [Kr]$5s^2$ Strontium	39 88.91 Y 3345 1522 1.2 [Kr]$4d^1 5s^2$ Yttrium	40 91.22 Zr 4406 1854 1.3 [Kr]$4d^2 5s^2$ Zirconium	41 92.91 Nb 4741 2477 1.6 [Kr]$4d^4 5s^1$ Niobium	42 95.96 Mo 4639 2622 2.2 [Kr]$4d^5 5s^1$ Molybdenum	43 (98) Tc 4262 2157 2.1 [Kr]$4d^5 5s^2$ Technetium	44 101.1 Ru 4147 2333 2.2 [Kr]$4d^7 5s^1$ Ruthenium	45 102.9 Rh 3695 1963 2.3 [Kr]$4d^8 5s^1$ Rhodium
6	55 132.9 Cs 670.8 28.5 0.8 [Xe]$6s^1$ Caesium	56 137.3 Ba 1845 727 0.9 [Xe]$6s^2$ Barium	57 138.9 La 2464 920 1.1 [Xe]$5d^1 6s^2$ Lanthanum	72 178.5 Hf 4600 2233 1.3 [Xe]$4f^{14} 5d^2 6s^2$ Hafnium	73 180.9 Ta 5455 3017 1.5 [Xe]$4f^{14} 5d^3 6s^2$ Tantalum	74 183.9 W 5555 3414 1.7 [Xe]$4f^{14} 5d^4 6s^2$ Tungsten	75 186.2 Re 5596 3454 1.9 [Xe]$4f^{14} 5d^5 6s^2$ Rhenium	76 190.2 Os 5008 3033 2.2 [Xe]$4f^{14} 5d^6 6s^2$ Osmium	77 192.2 Ir 4428 2446 2.2 [Xe]$4f^{14} 5d^7 6s^2$ Iridium
7	87 (223) Fr 676.8 27 0.7 [Rn]$7s^1$ Francium	88 (226) Ra 1140 699.8 0.9 [Rn]$7s^2$ Radium	89 (227) Ac 3200 1050 1.1 [Rn]$6d^1 7s^2$ Actinium	104 (261) Rf - - - [Rn]$5f^{14} 6d^2 7s^2$ Rutherfordium	105 (262) Db - - - [Rn]$5f^{14} 6d^3 7s^2$ Dubnium	106 (266) Sg - - - [Rn]$5f^{14} 6d^4 7s^2$ Seaborgium	107 (264) Bh - - - [Rn]$5f^{14} 6d^5 7s^2$ Bohrium	108 (267) Hs - - - [Rn]$5f^{14} 6d^6 7s^2$ Hassium	109 (268) Mt - - - [Rn]$5f^{14} 6d^7 7s^2$ Meitnerium

Periods	10	11	12	13	14	15	16	17	18
1									2 4.003 He -268.9 -272 - $1s^2$ Helium
2				5 10.81 B 4000 2077 2.0 $1s^2 2s^2 2p^1$ Boron	6 12.01 C 4827 3500 2.6 $1s^2 2s^2 2p^2$ Carbon	7 14.01 N -195.8 -210 3.0 $1s^2 2s^2 2p^3$ Nitrogen	8 16.00 O -183 -219 3.4 $1s^2 2s^2 2p^4$ Oxygen	9 19.00 F -188.1 -220 4.0 $1s^2 2s^2 2p^5$ Fluorine	10 20.18 Ne -246 -249 - $1s^2 2s^2 2p^6$ Neon
3				13 26.98 Al 2519 660.3 1.6 [Ne]$3s^2 3p^1$ Aluminium	14 28.09 Si 3265 1414 1.9 [Ne]$3s^2 3p^2$ Silicon	15 30.97 P 280.5 44.2 2.2 [Ne]$3s^2 3p^3$ Phosphorus	16 32.07 S 444.6 115.2 2.6 [Ne]$3s^2 3p^4$ Sulfur	17 35.45 Cl -34 -102 3.2 [Ne]$3s^2 3p^5$ Chlorine	18 39.95 Ar -185.8 -189 - [Ne]$3s^2 3p^6$ Argon
4	28 58.69 Ni 2913 1455 1.9 [Ar]$3d^8 4s^2$ Nickel	29 63.55 Cu 2560 1085 1.9 [Ar]$3d^{10} 4s^1$ Copper	30 65.38 Zn 907 419.5 1.6 [Ar]$3d^{10} 4s^2$ Zinc	31 69.72 Ga 2229 27.8 1.8 [Ar]$3d^{10} 4s^2 4p^1$ Gallium	32 72.64 Ge 2833 938.2 2.0 [Ar]$3d^{10} 4s^2 4p^2$ Germanium	33 74.92 As 613 816.8 2.2 [Ar]$3d^{10} 4s^2 4p^3$ Arsenic	34 78.96 Se 684.8 220.8 2.6 [Ar]$3d^{10} 4s^2 4p^4$ Selenium	35 79.90 Br 58.8 -7.1 3.0 [Ar]$3d^{10} 4s^2 4p^5$ Bromine	36 83.80 Kr -153.4 -157 - [Ar]$3d^{10} 4s^2 4p^6$ Krypton
5	46 106.4 Pd 2963 1555 2.2 [Kr]$4d^{10} 5s^0$ Palladium	47 107.9 Ag 2162 961.8 1.9 [Kr]$4d^{10} 5s^1$ Silver	48 112.4 Cd 766.8 321.1 1.7 [Kr]$4d^{10} 5s^2$ Cadmium	49 114.8 In 2072 156.6 1.8 [Kr]$4d^{10} 5s^2 5p^1$ Indium	50 118.7 Sn 2602 231.9 2.0 [Kr]$4d^{10} 5s^2 5p^2$ Tin	51 121.8 Sb 1587 630.6 2.0 [Kr]$4d^{10} 5s^2 5p^3$ Antimony	52 127.6 Te 987.8 449.5 2.1 [Kr]$4d^{10} 5s^2 5p^4$ Tellurium	53 126.9 I 184.4 113.7 2.7 [Kr]$4d^{10} 5s^2 5p^5$ Iodine	54 131.3 Xe -108.1 -112 2.6 [Kr]$4d^{10} 5s^2 5p^6$ Xenon
6	78 195.1 Pt 3825 1768 2.2 [Xe]$4f^{14} 5d^9 6s^1$ Platinum	79 197.0 Au 2836 1064 2.4 [Xe]$4f^{14} 5d^{10} 6s^1$ Gold	80 200.6 Hg 356.6 -38.8 1.9 [Xe]$4f^{14} 5d^{10} 6s^2$ Mercury	81 204.4 Tl 1473 303.8 1.8 [Xe]$4f^{14} 5d^{10} 6s^2 6p^1$ Thallium	82 207.2 Pb 1749 327.5 1.8 [Xe]$4f^{14} 5d^{10} 6s^2 6p^2$ Lead	83 209.0 Bi 1564 271.4 1.9 [Xe]$4f^{14} 5d^{10} 6s^2 6p^3$ Bismuth	84 (210) Po 962 253.8 2.0 [Xe]$4f^{14} 5d^{10} 6s^2 6p^4$ Polonium	85 (210) At 366.8 301.8 2.2 [Xe]$4f^{14} 5d^{10} 6s^2 6p^5$ Astatine	86 (222) Rn -61.8 -71.2 - [Xe]$4f^{14} 5d^{10} 6s^2 6p^6$ Radon
7	110 (271) Ds - - - [Rn]$5f^{14} 6d^8 7s^2$ Darmstadtium	111 (272) Rg - - - [Rn]$5f^{14} 6d^9 7s^2$ Roentgenium	112 (285) Cn - - - [Rn]$5f^{14} 6d^{10} 7s^2$ Copernicium	113 (284) Nh - - - [Rn]$5f^{14} 6d^{10} 7s^2 7p^1$ Nihonium	114 (289) Fl - - - [Rn]$5f^{14} 6d^{10} 7s^2 7p^2$ Flerovium	115 (289) Mc - - - [Rn]$5f^{14} 6d^{10} 7s^2 7p^3$ Moscovium	116 (292) Lv - - - [Rn]$5f^{14} 6d^{10} 7s^2 7p^4$ Livermorium	117 (294) Ts - - - [Rn]$5f^{14} 6d^{10} 7s^2 7p^5$ Tennessine	118 (294) Og - - - [Rn]$5f^{14} 6d^{10} 7s^2 7p^6$ Oganesson

f BLOCK

58 140.1 Ce 3443 799 1.1 [Xe]$4f^2 5d^0 6s^2$ Cerium	59 140.9 Pr 3520 931 1.1 [Xe]$4f^3 5d^0 6s^2$ Praseodymium	60 144.2 Nd 3074 1016 1.1 [Xe]$4f^4 5d^0 6s^2$ Neodymium	61 (145) Pm 3000 1042 - [Xe]$4f^5 5d^0 6s^2$ Promethium	62 150.4 Sm 1794 1072 1.1 [Xe]$4f^6 5d^0 6s^2$ Samarium	63 152.0 Eu 1794 1072 - [Xe]$4f^7 5d^0 6s^2$ Europium	64 157.3 Gd 3273 1313 1.2 [Xe]$4f^7 5d^1 6s^2$ Gadolinium	65 158.9 Tb 3230 1359 - [Xe]$4f^9 5d^0 6s^2$ Terbium
90 232.0 Th 4875 1750 1.3 [Rn]$5f^0 6d^2 7s^2$ Thorium	91 231.0 Pa - 1572 1.5 [Rn]$5f^2 6d^1 7s^2$ Protactinium	92 238.0 U 4131 1135 1.7 [Rn]$5f^3 6d^1 7s^2$ Uranium	93 (237) Np 3902 644 1.3 [Rn]$5f^4 6d^1 7s^2$ Neptunium	94 (244) Pu 3228 640 1.3 [Rn]$5f^6 6d^0 7s^2$ Plutonium	95 (243) Am 2011 1176 - [Rn]$5f^7 6d^0 7s^2$ Americium	96 (247) Cm 3110 1340 - [Rn]$5f^7 6d^1 7s^2$ Curium	97 (247) Bk - 986 - [Rn]$5f^9 6d^0 7s^2$ Berkelium

66 162.5 Dy 2567 1412 1.2 [Xe]$4f^{10} 5d^0 6s^2$ Dysprosium	67 164.9 Ho 2700 1472 1.2 [Xe]$4f^{11} 5d^0 6s^2$ Holmium	68 167.3 Er 2868 1529 1.2 [Xe]$4f^{12} 5d^0 6s^2$ Erbium	69 168.9 Tm 1950 1545 1.3 [Xe]$4f^{13} 5d^0 6s^2$ Thulium	70 173.1 Yb 1196 824 - [Xe]$4f^{14} 5d^0 6s^2$ Ytterbium	71 175.0 Lu 3402 1663 1.0 [Xe]$4f^{14} 5d^1 6s^2$ Lutetium
98 (251) Cf - 900 - [Rn]$5f^{10} 6d^0 7s^2$ Californium	99 (252) Es - - - [Rn]$5f^{11} 6d^0 7s^2$ Einsteinium	100 (257) Fm - - - [Rn]$5f^{12} 6d^0 7s^2$ Fermium	101 (258) Md - 827 - [Rn]$5f^{13} 6d^0 7s^2$ Mendelevium	102 (259) No - - - [Rn]$5f^{14} 6d^0 7s^2$ Nobelium	103 (262) Lr - - - [Rn]$5f^{14} 6d^1 7s^2$ Lawrencium

Pearson Australia
(a division of Pearson Australia Group Pty Ltd)
707 Collins Street, Melbourne, Victoria 3008
PO Box 23360, Melbourne, Victoria 8012
www.pearson.com.au

eBook credits
Resource authors: Doug Bail, Amber Dommel, Norbert Dommel, Mark Hamilton, Kristen Hebden, Richard Hecker, David Madden, Jeff Stanger
Lead publishers: Misal Belvedere, Malcolm Parsons
Project Manager: Michelle Thomas
Production Editors: Elizabeth Gosman, Laura Pietrobon, Virginia O'Brien
Lead Development Editor: Fiona Cooke
Development Editors: Naomi Campanale, Haeyean Lee
Editors: Sam Trafford, iEnergizer Aptara Ltd
Rights and Permissions Editor: Samantha Russell-Tulip
Illustrator: DiacriTech
eBook Production Coordinator: Louise Grundy
Digital Administrator: Jenny Smirnov

ISBN 978 1 4886 1942 7

Pearson Australia Group Pty Ltd ABN 40 004 245 943

eBook attributions
Video: *Reviewing the scientific method*, boclips/Mazzarella Media, Chapter 1.