IRANINOVIN

Cambridge
International AS and A Level

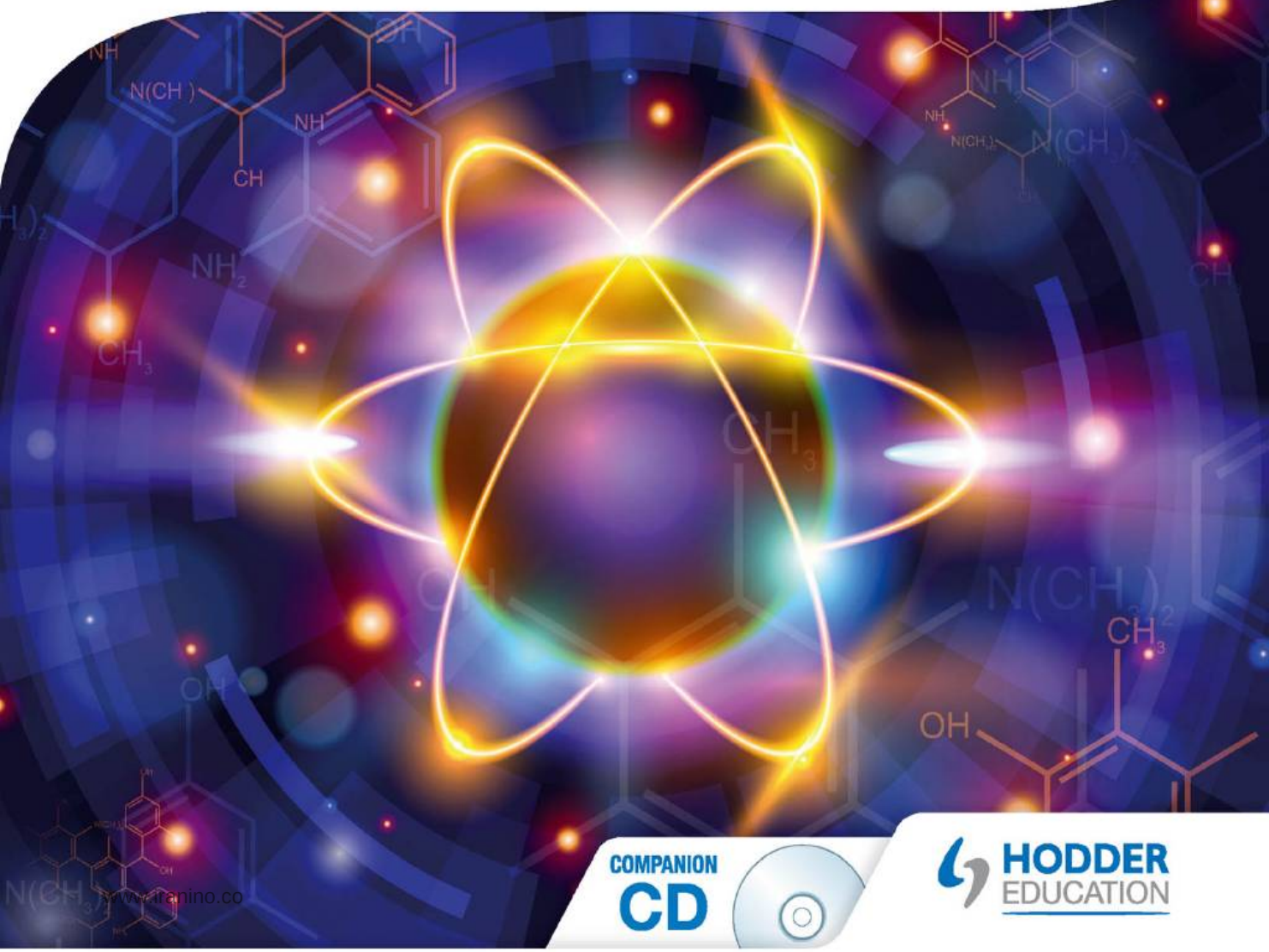
ENDORSED BY
CAMBRIDGE
International Examinations

NEW FOR 2014

Physics

Second Edition

Mike Crundell, Geoff Goodwin and Chris Mee



COMPANION CD

HODDER EDUCATION

www.iranino.co

IRANINOVIN

ENDORSED BY



CAMBRIDGE
International Examinations

Cambridge
International AS and A Level

Physics

Mike Crundell and Geoff Goodwin

HODDER EDUCATION
AN HACHETTE UK COMPANY

www.iranino.co

IRANINOVIN

Questions from Cambridge International AS and A Level Physics papers are reproduced by permission of Cambridge International Examinations

Cambridge International Examinations bears no responsibility for the example answers to questions taken from its past question papers which are contained in this publication.

Hachette UK's policy is to use papers that are natural, renewable and recyclable products and made from wood grown in sustainable forests. The logging and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

Orders: please contact Bookpoint Ltd, 130 Milton Park, Abingdon, Oxon OX14 4SB. Telephone: (44) 01235 827720. Fax: (44) 01235 400454. Lines are open 9.00–5.00, Monday to Saturday, with a 24-hour message answering service. Visit our website at www.hoddereducation.com

© 2014
First published in 2014 by
Hodder Education, a Hachette UK company,
338 Euston Road
London NW1 3BH

Impression number 5 4 3 2 1
Year 2018 2017 2016 2015 2014

All rights reserved. Apart from any use permitted under UK copyright law, no part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or held within any information storage and retrieval system, without permission in writing from the publisher or under licence from the Copyright Licensing Agency Limited. Further details of such licences (for reprographic reproduction) may be obtained from the Copyright Licensing Agency Limited, Saffron House, 6–10 Kirby Street, London EC1N 8TS.

Cover photo by © Uladzimir Bakunovich-Fotolia.com
Illustrations by Integra Software Services and Charon Tec Ltd.
Typeset in ITC Garamond Std Light 9/11 by Integra Software Services Pvt. Ltd., Pondicherry, India
Printed in Italy

A catalogue record for this title is available from the British Library
ISBN 978 1471 80921 7

www.iraninio.co

IRANINOVIN

Contents

Acknowledgements vii
Introduction viii
Syllabus structure relating to book topics ix

AS Level

Topic 1 Physical quantities and units 1
1.1 Physical quantities 2
1.2 SI quantities and base units 2
1.3 Scalars and vectors 7

Topic 2 Measurement techniques 15
2.1 Measurements 15
2.2 Errors and uncertainties 31

Topic 3 Kinematics 40
3.1 Speed, displacement, velocity and acceleration 40

Topic 4 Dynamics 55
4.1 Relationships involving force and mass 55
4.2 Weight 58
4.3 The principle of conservation of momentum 61

Topic 5 Forces, density and pressure 71
5.1 Types of force 71
5.2 Moment of a force 71
5.3 Equilibrium of forces 73
5.4 Density and pressure 76

Topic 6 Work, energy, power 80
6.2 Work 80
6.1 Energy 83
6.3 Kinetic energy 85
6.4 Power 89

Topic 9 Deformation of solids 94
9.1 & 9.2 Force and deformation 94

Topic 14 Waves 101
14.1 Wave motion 101
14.2 Graphical representation of waves 102
14.3 The determination of the frequency of sound using a calibrated c.r.o. 107
14.4 Doppler effect 108
14.5 The electromagnetic spectrum 109

www.iranino.co

Contents

IRANINOVIN

Topic 15 Superposition 114
15.3 Interference 114
15.1 Stationary waves 119
14.3 Measuring the speed of sound using stationary waves 124
15.2 & 15.4 Diffraction 126

Topic 17 Electric fields 136
17.1 Electric forces and fields 136
17.2 Electric field strength 139

Topic 19 Current of electricity 145
19.1 Charge and current 145
19.2 Potential difference 148
19.3 Resistance 149

Topic 20 D.C. circuits 156
20.1 Electrical circuits 156
20.2 Kirchhoff's first and second laws 159
20.3 Potential dividers and potentiometers 162

Topic 26 Particle physics 168
26.1 Atomic structure and radioactivity 168
26.2 Fundamental particles 178

A Level

Topic 7 Motion in a circle 184
7.1 Radian measure and angular displacement 184
7.2 Centripetal acceleration and centripetal force 185

Topic 8 Gravitational fields 191
8.1 Gravitational field 191
8.2 & 8.3 Gravitational field strength 192
8.4 Gravitational potential and gravitational potential energy 198

Topic 10 Ideal gases 202
10.1 & 10.3 Equation of state of an ideal gas 202
10.2 A microscopic model of a gas 205

Topic 11 Temperature 211
11.1 Temperature 211
11.2 Temperature scales 212
11.3 Thermometers 215

Topic 12 Thermal properties of materials 217
12.1 Solids, liquids and gases, and thermal (heat) energy 217
12.2 Internal energy 225

Topic 13 Oscillations 230
13.1 Oscillations 230
13.2 Energy changes in simple harmonic motion 235
13.3 Free and damped oscillations 237

www.iranino.co

IRANINOVIN

Contents

Topic 14 Ultrasound 247
14.6 The generation and use of ultrasound 247

Topic 16 Communication 255
16.1 Communication channels 255
16.2 Modulation 257
16.3 Analogue and digital signals 262
16.4 Relative merits of channels of communication 265
16.5 Signal attenuation 268

Topic 17 Electric fields 274
17.3 Point charges 274
17.4 Electric field strength due to a point charge 276
17.5 Electric potential energy and electric potential 276

Topic 18 Capacitance 281
18.1 Capacitors and capacitance 281
18.2 Energy stored in a capacitor 285

Topics 19 & 20 Electronic sensors 292
19.4 Sensing devices 292
20.3 The use of potential dividers 295

Topic 21 Electronics 298
21.1 The ideal operational amplifier (op-amp) 298
21.2 Operational amplifier circuits 299
21.3 Output devices 305

Topic 22 Magnetic fields 311
22.1 Concept of a magnetic field 311
22.2 Force on a current-carrying conductor 314
22.3 Force on a moving charged particle in a magnetic field 318
22.4 Magnetic fields due to currents 324
22.5 The use of (nuclear) magnetic resonance imaging 326

Topic 23 Electromagnetic induction 332
23.1 Magnetic flux and electromagnetic induction 332

Topic 24 Alternating currents 341
24.1 Characteristics of alternating currents 341
24.2 Transformers 343
24.3 Transmission of electrical energy 344
24.4 Rectification 345

Topic 25 Quantum physics 350
25.1 & 25.2 Photoelectric emission of electrons and energy of a photon 350
25.3 Wave–particle duality 355
25.4 Energy levels in atoms and line spectra 356
25.5 Band theory 360
25.6 The production and use of X-rays 363

www.iranino.co

IRANINOVIN

Topic 26 Nuclear physics 376
26.3 Mass defect and nuclear binding energy 376
26.4 The spontaneous and random nature of radioactive decay 381
AS Level Answers to Now it's your turn and Examination style questions 388
A Level Answers to Now it's your turn and Examination style questions 394
Index 399

Student's CD contents



Interactive tests

Topic summaries & Definitions and formulae

Revision checklists

Examination structure

Planning revision

Examination technique

Glossary of command words

Calculation of uncertainty in a result

Significant figures and decimal places

Proportionality and linearity

Straight-line graphs

www.iranino.co

Acknowledgements

The publishers would like to thank the following for permission to use copyright material

Photo credits:
p.2 *tl* © Royal Observatory Edinburgh/SPL; **p.2** *tr* © Juice Images / Alamy; **p.2** *bl* © Odua Images – Fotolia.com; **p.2** *bc* © age fotostock/Robert Harding; **p.2** *br* © EK aviation / Alamy; **p.3** © Alfred Pasieka/SPL; **p.6** *bl* © kyslynskyy – Fotolia.com; **p.6** *tl* © Corey Ford / Alamy; **p.7** © PCN/Corbis; **p.11** © Thierry GRUN - Aero / Alamy; **p.16** © Science Photo Library; **p.17** © Leslie Garland Picture Library / Alamy; **p.18** © ANDREW LAMBERT PHOTOGRAPHY/SCIENCE PHOTO LIBRARY; **p.21** *tl* © MARTYN F. CHILLMAID/SCIENCE PHOTO LIBRARY; **p.21** *bl* © foodfolio / Alamy; **p.21** *bc* © MICHAEL DALTON/FUNDAMENTAL PHOTOS/ SCIENCE PHOTO LIBRARY; **p.21** © ALEX BARTEL/SCIENCE PHOTO LIBRARY; **p.23** © Martyn Chillmaid; **p.24** *tl* © Synthetic Alan King / Alamy; **p.24** *bl* © burnel11 – Fotolia.com; **p.27** © SSPL via Getty Images; **p.28** *tl* © ANDREW LAMBERT PHOTOGRAPHY/SCIENCE PHOTO LIBRARY; **p.28** *tr* © sciencephotos / Alamy; **p.28** *br* © Noel Toone / PhotographersDirect.com; **p.29** © ANDREW LAMBERT PHOTOGRAPHY/SCIENCE PHOTO LIBRARY; **p.31** © ANDREW LAMBERT PHOTOGRAPHY/SCIENCE PHOTO LIBRARY; **p.45** © Martyn Chillmaid; **p.46** *r* © The Art Archive / Alamy; **p.46** *l* © frenta – Fotolia.com; **p.49** © AFP/Getty Images; **p.50** © John Doornkamp / Design Pics; **p.56** © Bill Sanderson/SPL; **p.57** © ChinaFotoPress via Getty Images; **p.60** © Dreamframer / Alamy; **p.63** © EDWARD KINSMAN/SCIENCE PHOTO LIBRARY; **p.64** © Kimberly Reinick – Fotolia.com; **p.73** © Stefan Sollfors / Alamy; **p.80** © AFP/SCIENCE PHOTO LIBRARY; **p.81** © Jurand – Fotolia.com; **p.82** © ***(all)*** Alex Bartel/SCIENCE PHOTO LIBRARY; **p.83** *tl* © Chrysler Group LLC; **p.83** *b* © Andrew Lambert Photography/SCIENCE PHOTO LIBRARY; **p.84** © Chrysler Group LLC; **p.85** © Dr Jeremy Burgess/SPL; **p.86** © hallucion_7 – Fotolia.com; **p.88** © J-L Charmet/SPL; **p.90** © dpullman – Fotolia.com; **p.115** © BRUCE COLEMAN INC. / Alamy; **p.119** *t* © GIPHOTOSTOCK/SCIENCE PHOTO LIBRARY; **p.119** *b* © imagebroker / Alamy; **p.120** © ANDREW LAMBERT PHOTOGRAPHY/SCIENCE PHOTO LIBRARY; **p.128** © GIPHOTOSTOCK/SCIENCE PHOTO LIBRARY; **p.148** © Wendy Brown; **p.151** © SCIENCE PHOTO LIBRARY; **p.163** © ANDREW LAMBERT PHOTOGRAPHY/SCIENCE PHOTO LIBRARY; **p. 171** © SSPL via Getty Images; **p.175** © PUBLIC HEALTH ENGLAND/SCIENCE PHOTO LIBRARY; **p.176** © Prof Erwin Mueller/SPL; **p.178** © Dr David Wexler/SPL; **p.188** *tl* © TEK IMAGE/SCIENCE PHOTO LIBRARY; **p.188** *tr* © Novosti/SPL; **p.188** *br* © Vitaliy Stepanenko – Thinkstock.com; **p.195** © JULIAN BAUM/SCIENCE PHOTO LIBRARY; **p.202** © SCIENCE PHOTO LIBRARY; **p.215** © ANDREW LAMBERT PHOTOGRAPHY/SCIENCE PHOTO LIBRARY; **p.216** © David J. Green / Alamy; **p.238** © Tawee Uthisanuraksakul / Alamy; **p.239** © ANDREW LAMBERT PHOTOGRAPHY/SCIENCE PHOTO LIBRARY; **p.240** *tl* © Blend Images / Alamy; **p.240** *br* © AP/AP/Press Association Images; **p.251** © Montgomery Martin / Alamy; **p.252** © SOVEREIGN, ISM/SCIENCE PHOTO LIBRARY; **p.257** © TEK IMAGE/SCIENCE PHOTO LIBRARY; **p.293** *tl* © MARTYN F. CHILLMAID/SCIENCE PHOTO LIBRARY; **p.293** *tr* © MARTYN F. CHILLMAID/SCIENCE PHOTO LIBRARY; **p.298** © RS Components Ltd http://uk.rs-online.com Quad op-amp HA 347415; **p.312** © Hodder Education**; p.313** © SCIENCE PHOTO LIBRARY; **p.319** © LAWRENCE BERKELEY LABORATORY/SCIENCE PHOTO LIBRARY; **p.320** © STEVE ALLEN/SCIENCE PHOTO LIBRARY; **p.326** © Pete Saloutos/Corbis; **p.327** © ZEPHYR/SCIENCE PHOTO LIBRARY; **p.344** Courtest of Tiles / Wikipedia Commons / GNU_Free_Documentation_License (http://en.wikipedia.org/wiki/File:PoleMountedDistributionTransformerNZ.jpg); **p.353** © The Granger Collection / TopFoto; **p.355** *t* WA Steer/Daresbury Lab/UCL; **p.355** *b* © ANDREW LAMBERT PHOTOGRAPHY/SCIENCE PHOTO LIBRARY; **p.356** © DEPT. OF PHYSICS, IMPERIAL COLLEGE/SCIENCE PHOTO LIBRARY; **p.358** © DEPT. OF PHYSICS, IMPERIAL COLLEGE/SCIENCE PHOTO LIBRARY; **p.359** *t* © PHYSICS DEPT., IMPERIAL COLLEGE/SCIENCE PHOTO LIBRARY; **p.359** *b* © Mike Guidry (b–d), Mike Guidry, University of Tennessee; **p.363** © GUSTOIMAGES/SCIENCE PHOTO LIBRARY; **p.366** © SOVEREIGN, ISM/SCIENCE PHOTO LIBRARY; **p.367** © SOVEREIGN, ISM/SCIENCE PHOTO LIBRARY; **p.369** *l* © MAURO FERMARIELLO/SCIENCE PHOTO LIBRARY; **p.369** *r* © ZEPHYR/SCIENCE PHOTO LIBRARY; **p.379** © European Space Agency/SPL

Acknowledgements:

Every effort has been made to trace all copyright holders, but if any have been inadvertently overlooked, the Publishers will be pleased to make the necessary arrangements at the first opportunity.

www.iranino.co

IRANINOVIN

Introduction

This book is a new edition of *International AS and A Level Physics* that has been revised and amended to be compatible with the Physics Syllabus 9702 of Cambridge International Examinations, published in 2014 for first examination in 2016.

The book has been fully endorsed by Cambridge International Examinations and is listed as an endorsed textbook for students studying this syllabus. New material has been included, where necessary, so that the book provides comprehensive cover of the subject content. The content of the book has been re-ordered so that students studying AS physics will find that part of the syllabus as being separate from the second part of the A Level course. In a few places, the content of the book goes slightly beyond the syllabus requirements, either to provide some background information or to arrive at a satisfactory termination of a topic.

All the assessment objectives that are identified in the syllabus are covered in the book. The learning outcomes, as specified in the syllabus, are listed in each topic of the book using the same wording as in the syllabus so that students may identify easily the section of the syllabus that is being covered. The content of each topic is identified by learning outcome, not necessarily presented in sequential numerical order, but according to the most sensible order for learning. For example, in AS Level Topic 6, learning outcome 6.2 (Work) comes before 6.1 (Energy). The chart opposite shows how topics are arranged in the book and how this relates to syllabus coverage.

A new feature of the syllabus is Key concepts. These are the essential ideas, theories, principles or mental tools that help learners to develop a deep understanding of their subject, and make links between different topics. An icon indicates where each Key concept is covered:



Models of physical systems
Physics is the science that seeks to understand the behaviour of the Universe. The development of models of physical systems is central to physics. Models simplify, explain and predict how physical systems behave.



Testing predictions against evidence
Physical models are usually based on prior observations, and their predictions are tested to check that they are consistent with the behaviour of the real world. This testing requires evidence, often obtained from experiments.



Mathematics as a language and problem-solving tool
Mathematics is integral to physics, as it is the language that is used to express physical principles and models. It is also a tool to analyse theoretical models, solve quantitative problems and produce predictions.



Matter, energy and waves
Everything in the Universe comprises matter and/or energy. Waves are a key mechanism for the transfer of energy and are essential to many modern applications of physics.



Forces and fields
The way that matter and energy interact is through forces and fields. The behaviour of the Universe is governed by fundamental forces that act over different length scales and magnitudes. These include the gravitational force and the electromagnetic force.

Key points, definitions and equations are highlighted in coloured panels. There is a summary of the important features that have been covered after each section or topic. Throughout each topic, worked examples are provided so that students may familiarise themselves with the subject matter. The worked examples are followed by questions of similar difficulty, listed under the *Now it's your turn* headings. In addition, there are questions which have a broader context and are of the examination style as regards wording and level of difficulty. Answers to both types of question are provided at the back of the book.

This book has been written specifically for the Cambridge syllabus. However, its coverage of subject matter and style of questions make it suitable for students who are studying towards physics qualifications of other awarding bodies.

Mike Crundell
Geoff Goodwin
July 2014

www.iranino.co

Syllabus structure relating to book topics

Syllabus section	International A/AS Level Physics AS Level Topics	International A/AS Level Physics A Level Topics
1. Physical quantities and units	1. Physical quantities and units	
2. Measurement techniques	2. Measurement techniques	
3. Kinematics	3. Kinematics	
4. Dynamics	4. Dynamics	
5. Forces, density and pressure	5. Forces, density and pressure	
6. Work, energy, power	6. Work, energy, power	
7. Motion in a circle		7. Motion in a circle
8 Gravitational fields		8 Gravitational fields
9. Deformation of solids	9. Deformation of solids	
10. Ideal gases		10. Ideal gases
11. Temperature		11. Temperature
12. Thermal properties of materials		12. Thermal properties of materials
13. Oscillations		13. Oscillations
14. Waves	14. Waves *covers 14.1, 14.2, 14.3, 14.4 and 14.5*	14. Ultrasound *covers 14.6*
15. Superposition	15. Superposition *also covers 14.3b*	
16. Communication		16. Communication
17. Electric fields	17. Electric fields *covers 17.1 and 17.2*	17. Electric fields *covers 17.3, 17.4 and 17.5*
18. Capacitance		18. Capacitance
19. Current of electricity	19. Current of electricity *covers 19.1, 19.2 and 19.3*	19 & 20 Electronic sensors *covers 19.4*
20 D.C. circuits	20 D.C. circuits *covers 20.1, 20.2 and 20.3 a, b*	19 & 20 Electronic sensors *covers 20.3 c, d*
21. Electronics		21. Electronics
22. Magnetic fields		22. Magnetic fields
23. Electromagnetic induction		23. Electromagnetic induction
24. Alternating currents		24. Alternating currents
25. Quantum physics		25. Quantum physics
26. Particle and nuclear physics	26. Particle and nuclear physics *covers 26.1 and 26.2*	26. Particle and nuclear physics *covers 26.3 and 26.4*

www.iranino.co

This page intentionally left blank

www.iranino.co

IRANINOVIN

1 Physical quantities and units

By the end of this topic, you will be able to:

1.1 **(a)** understand that all physical quantities consist of a numerical magnitude and a unit

(b) make reasonable estimates of physical quantities included in the syllabus

1.2 **(a)** recall the following SI base quantities and their units: mass (kg), length (m), time (s), current (A), temperature (K), amount of substance (mol)

(b) express derived units as products or quotients of the SI base units and use the named units listed in the syllabus as appropriate

(c) use SI base units to check the homogeneity of physical equations

(d) use the following prefixes and their symbols to indicate decimal submultiples or multiples of both base and derived units:

- pico (p)
- nano (n)
- micro (μ)
- milli (m)
- centi (c)
- deci (d)
- kilo (k)
- mega (M)
- giga (G)
- tera (T)

(e) understand and use the conventions for labelling graph axes and table columns

1.3 **(a)** distinguish scalar and vector quantities and give examples of each

(b) add and subtract coplanar vectors

(c) represent a vector as two perpendicular components

Note: amount of substance (mol) is only used in the A level course but is included here for completeness.

Starting points

- Accurate **measurement** is very important in the development of physics.
- Physicists begin by observing, measuring and collecting data.
- These data are analysed to discover whether they fit into a pattern.
- If there is a pattern and this pattern can be used to explain other events, it becomes a **theory**.
- The process is known as the **scientific method** (see Figure 1.1).

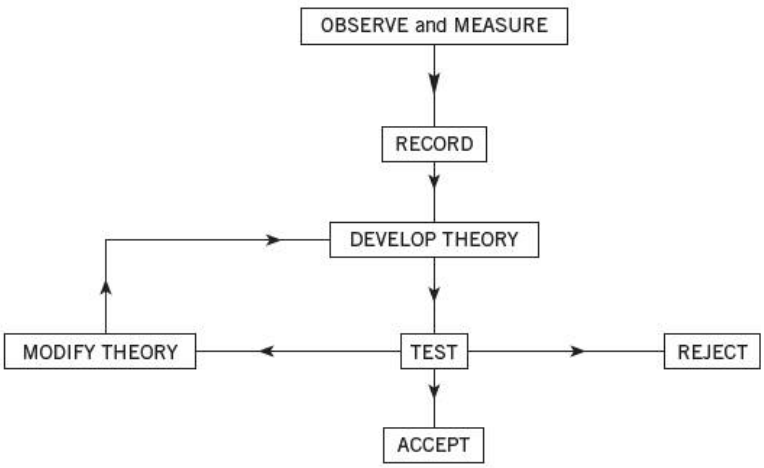


Figure 1.1 Block diagram to illustrate the scientific method

www.iranino.co

IRANINOVIN

1.1 Physical quantities





Figure 1.2 Brahe (1546–1601) measured the elevations of stars; these days a modern theodolite is used for measuring angular elevation.

A **physical quantity** is a feature of something which can be measured, for example, length, weight, or time of fall. Every physical quantity has a numerical value and a unit. If someone says they have a waist measurement of 50, they could be very slim or very fat depending on whether the measurement is in centimetres or inches! Take care – it is vital to give the unit of measurement whenever a quantity is measured or written down.

Large and small quantities are usually expressed in scientific notation, i.e. as a simple number multiplied by a power of ten. For example, 0.000 34 would be written as 3.4×10^{-4} and 154 000 000 as 1.54×10^{8}. There is far less chance of making a mistake with the number of zeros!



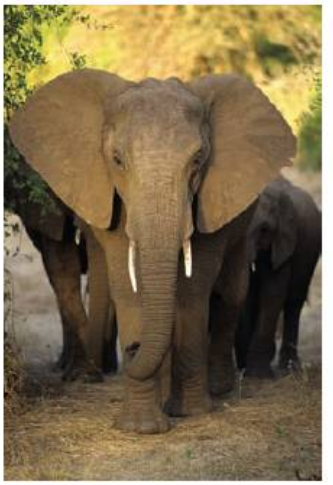




Figure 1.3 The elephant is large in comparison with the boy but small compared with the jumbo jet.

1.2 SI quantities and base units

In very much the same way that languages have developed in various parts of the world, many different systems of measurement have evolved. Just as languages can be translated from one to another, units of measurement can also be converted between systems. Although some conversion factors are easy to remember, some are very difficult. It is much better to have just one system of units. For this reason, scientists around the world use the **Système International (SI)** which is based on the metric system of measurement.

www.iranino.co

IRANINOVIN



Figure 1.4 The mass of this jewel could be measured in kilograms, pounds, carats, grains, etc.

If a quantity is to be measured accurately, the unit in which it is measured must be defined as precisely as possible.

SI is founded on seven fundamental or **base units**.

The base quantities and the units with which they are measured are listed in Table 1.1. For completeness, the candela has been included, but this unit will not be used in the A/AS course. The mole will only be used in the A Level course.

Table 1.1 The base quantities and units

quantity	unit	symbol
mass	kilogram	kg
length	metre	m
time	second	s
electric current	ampere (amp)	A
thermodynamic temperature	kelvin	K
amount of substance	mole	mol
luminous intensity	candela	cd

Each quantity has just one unit and this unit can have **multiples** and **sub-multiples** to cater for larger or smaller values. The unit is given a **prefix** to denote the multiple or sub-multiple (see Table 1.2). For example, one thousandth of a metre is known as a millimetre (mm) and 1.0 millimetre equals 1.0×10^{-3} metres (m).

Table 1.2 The more commonly used prefixes

prefix	symbol	multiplying factor
tera	T	10^{12}
giga	G	10^{9}
mega	M	10^{6}
kilo	k	10^{3}
deci	d	10^{-1}
centi	c	10^{-2}
milli	m	10^{-3}
micro	μ	10^{-6}
nano	n	10^{-9}
pico	p	10^{-12}

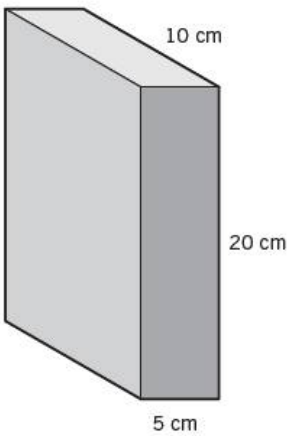


Figure 1.5 This box has a volume of $1.0 \times 10^{3}\,cm^{3}$.

Beware when converting units for areas and volumes!

$1\,mm = 10^{-3}\,m$

Squaring both sides $1\,mm^2 = (10^{-3})^2\,m^2 = 10^{-6}\,m^2$

and $1\,mm^3 = (10^{-3})^3\,m^3 = 10^{-9}\,m^3$

Note also that $1\,cm^2 = (10^{-2})^2\,m^2 = 10^{-4}\,m^2$

and $1\,cm^3 = (10^{-2})^3\,m^3 = 10^{-6}\,m^3$

The box in Figure 1.5 has a volume of $1.0 \times 10^{3}\,cm^3$ or $1.0 \times 10^{6}\,mm^3$ or $1.0 \times 10^{-3}\,m^3$.

A distance of thirty metres should be written as 30 m and not 30 ms or 30 m s. The letter s is **never** included in a unit for the plural. If a space is left between two letters, the letters denote different units. So, 30 m s would mean thirty metre seconds and 30 ms means 30 milliseconds.

www.iranino.co

IRANINOVIN

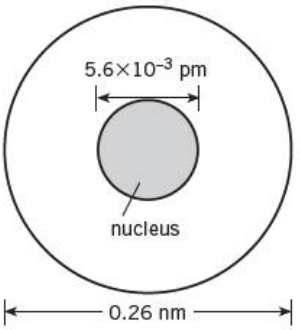


Figure 1.6 Atom of gold

Example

Calculate the number of micrograms in 1.0 milligram.

$1.0\,g = 1.0 \times 10^3\,mg$

and $1.0\,g = 1.0 \times 10^6$ micrograms (μg)

so, $1.0 \times 10^3\,mg = 1.0 \times 10^6\,\mu g$

and $1.0\,mg = (1.0 \times 10^6)/(1.0 \times 10^3) = \mathbf{1.0 \times 10^3\,\mu g}$

Now it's your turn

1 Calculate the area, in cm^2, of the top of a table with sides of 1.2 m and 0.9 m.
2 Determine the number of cubic metres in one cubic kilometre.
3 Write down, using scientific notation, the values of the following quantities:
   (a) 6.8 pF,
   (b) 32 μC,
   (c) 60 GW.
4 How many electric fires, each rated at 2.5 kW, can be powered from a generator providing 2.0 MW of electric power?
5 An atom of gold, Figure 1.6, has a diameter of 0.26 nm and the diameter of its nucleus is 5.6×10^{-3} pm. Calculate the ratio of the diameter of the atom to that of the nucleus.

Derived units

All quantities, apart from the base quantities, can be expressed in terms of **derived units**.

Derived units consist of some combination of the base units. The base units may be multiplied together or divided by one another, but never added or subtracted.

See Table 1.3 for examples of derived units. Some quantities have a named unit. For example, the unit of force is the newton, symbol N, but the newton can be expressed in terms of base units. Quantities which do not have a named unit are expressed in terms of other units. For example, specific latent heat (Topic 12) is measured in joules per kilogram ($J\,kg^{-1}$).

Table 1.3 Some examples of derived units which may be used in the A/AS course

quantity	unit	derived unit
frequency	hertz (Hz)	s^{-1}
velocity	$m\,s^{-1}$	$m\,s^{-1}$
acceleration	$m\,s^{-2}$	$m\,s^{-2}$
force	newton (N)	$kg\,m\,s^{-2}$
energy	joule (J)	$kg\,m^2\,s^{-2}$
power	watt (W)	$kg\,m^2\,s^{-3}$
electric charge	coulomb (C)	A s
potential difference	volt (V)	$kg\,m^2\,s^{-3}\,A^{-1}$
electrical resistance	ohm (Ω)	$kg\,m^2\,s^{-3}\,A^{-2}$
specific heat capacity	$J\,kg^{-1}\,K^{-1}$	$m^2\,s^{-2}\,K^{-1}$

Example

What are the base units of speed?

Speed is defined as $\frac{distance}{time}$ and so the unit is $\frac{m}{s}$.

Division by a unit is shown using a negative index, that is, s^{-1}.

The base units of speed are $\mathbf{m\,s^{-1}}$.

www.iranino.co

IRANINOVIN

Now it's your turn

Use the information in Tables 1.1 and 1.3 to determine the base units of the following quantities.

6 Density $\left(= \frac{mass}{volume}\right)$

7 Pressure $\left(= \frac{force}{area}\right)$

Checking equations

It is possible to work out the total number of oranges in two bags if one bag contains four and the other five (the answer is nine!). This exercise would, of course, be nonsense if one bag contained three oranges and the other four mangoes. In the same way, for any equation to make sense, each term involved in the equation must have the same base units. A term in an equation is a group of numbers and symbols, and each of these terms (or groups) is added to, or subtracted from, other terms. For example, in the equation

$$v = u + at$$

the terms are v, u, and at.

In any equation where each term has the same base units, the equation is said to be **homogeneous** or 'balanced'.

In the example above, each term has the base units m s^{-1}. If the equation is not homogeneous, then it is incorrect and is not valid. When an equation is known to be homogeneous, then the balancing of base units provides a means of finding the units of an unknown quantity.

Example

Use base units to show that the following equation is homogeneous.

work done = gain in kinetic energy + gain in gravitational potential energy

The terms in the equation are work, (gain in) kinetic energy, and (gain in) gravitational potential energy.

work done = force × distance moved

and so the base units are kg m s^{-2} × m = **kg m^2 s^{-2}**.

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

Since any pure number such as $\frac{1}{2}$ has no unit, the base units are kg × $(m\ s^{-1})^2$ = **kg m^2 s^{-2}**.

potential energy = mass × gravitational field strength g × distance

The base units are kg × m s^{-2} × m = **kg m^2 s^{-2}**.

Conclusion: **All terms have the same base units and the equation is homogeneous.**

Now it's your turn

8 Use base units to check whether the following equations are balanced:
(a) *pressure = depth × density × gravitational field strength,*
(b) *energy = mass × (speed of light)*^{2}.

9 The thermal energy Q needed to melt a solid of mass m without any change of temperature is given by the equation

$$Q = mL$$

where L is a constant. Find the base units of L.

10 Determine the base units of the following quantities:
(a) energy (= *force × distance*),
(b) specific heat capacity.
(*thermal energy change = mass × specific heat capacity × temperature change*)

11 Show that the left-hand side of the equation

$$pressure + \frac{1}{2} \times density \times (speed)^2 = \text{constant}$$

is homogeneous and find the base units of the constant on the right-hand side.

www.iranino.co

IRANINOVIN

Conventions for symbols and units

You may have noticed that when symbols and units are printed, they appear in different styles of type. The symbol for a physical quantity is printed in *italic* (sloping) type, whereas its unit is in roman (upright) type. For example, velocity v is italic, but its unit m s^{-1} is roman. Of course, you will not be able to make this distinction in handwriting.

At A/AS level and beyond, there is a special convention for labelling columns of data in tables and graph axes. The symbol is printed first (in italic), separated by a forward slash (the printing term is a solidus) from the unit (in roman). Then the data is presented in a column, or along an axis, as pure numbers. This is illustrated in Figure 1.7, which shows a table of data and the resulting graph for the velocity v of a particle at various times t.

t/s	v/m s^{-1}
0	2.5
1.0	4.0
2.0	5.5

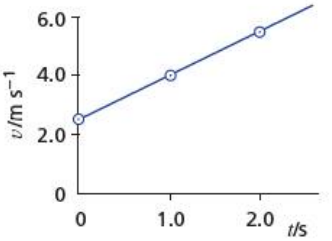


Figure 1.7 The convention for labelling tables and graphs

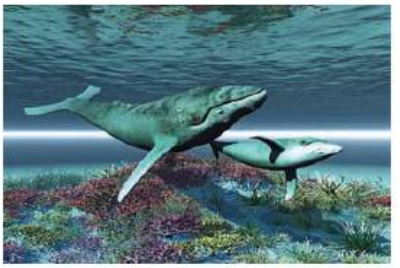




Figure 1.8 The ratio of the mass of the humpback whale to the mass of the mouse is about 10^4. That is minute compared to the ratio of the mass of a galaxy to the mass of a nucleus (10^{68})!

If you remember that a physical quantity contains a pure number and a unit, the reason for this style of presentation becomes clear. By dividing a physical quantity such as time (a number and a unit) by the appropriate unit, you are left with a pure number. It is then algebraically correct for the data in tables, and along graph axes, to appear as pure numbers.

You may also see examples in which the symbol for the physical quantity is followed by the slash, and then by a power of 10, and then the unit, for example $t/10^2$ s. This means that the column of data has been divided by 100, to save repeating lots of zeros in the table. If you see a table or graph labelled $t/10^2$ s and the figures 1, 2, 3 in the table column or along the graph axis, this means that the experimental data was obtained at values of t of 100 s, 200 s, 300 s.

Try to get out of the habit of heading table columns and graphs in ways such as 't in s', 't(s)' or even of recording each reading in the table as 1.0 s, 2.0 s, 3.0 s.

Order of magnitude of quantities

It is often useful to be able to estimate the size, or **order of magnitude**, of a quantity. Strictly speaking, the order of magnitude is the power of ten to which the number is raised. The ability to estimate is particularly important in a subject like physics where quantities have such widely different values. A *short* distance for an astrophysicist is a light-year (about 9.5×10^{15} m) whereas a *long* distance for a nuclear physicist is 6×10^{-15} m (the approximate diameter of a nucleus)! Table 1.4 gives some values of distance which may be met in the A/AS Physics course.

Table 1.4 Some values of distance

	distance/m
distance from Earth to edge of observable Universe	1.4×10^{26}
diameter of a galaxy	1.2×10^{21}
distance from Earth to the Sun	1.5×10^{11}
distance from London to Paris	3.5×10^{5}
length of a car	4
diameter of a hair	5×10^{-4}
diameter of an atom	3×10^{-10}
diameter of a nucleus	6×10^{-15}

The ability to estimate orders of magnitude is valuable when planning and carrying out experiments or when suggesting theories. Having an idea of the expected result provides a useful check that a silly error has not been made. This is also true when using a calculator. For example, the acceleration of free fall at the Earth's surface is about 10 m s^{-2}. If a value of 9800 m s^{-2} is calculated, then this is obviously wrong and a simple error in the power of ten is likely to be the cause. Similarly, a calculation in which the cost of boiling a kettle of water is found to be several dollars, rather than a few cents, may indicate that the energy has been measured in watt-hours rather than kilowatt-hours.

www.iran20.com

IRANINOVIN

Example

It is worthwhile to remember the sizes of some common objects so that comparisons can be made. For example, a jar of peanut butter has a mass of about 500 g and a carton of orange juice has a volume of 1000 cm^3 (1 litre).

Now it's your turn

12 Estimate the following quantities:
- (a) the mass of an orange,
- (b) the mass of an adult human,
- (c) the height of a room in a house,
- (d) the diameter of a pencil,
- (e) the volume of a small bean,
- (f) the volume of a human head,
- (g) the speed of a jumbo jet,
- (h) the temperature of the human body.



Figure 1.9 Although the athlete runs 10 km in the race, his final distance from the starting point may well be zero!

1.3 Scalars and vectors

All physical quantities have a magnitude and a unit. For some quantities, magnitude and units do not give us enough information to describe fully the quantity. For example, if we are given the time for which a car travels at a certain speed, then we can calculate the distance travelled. However, we cannot find out how far the car is from its starting point unless we are told the direction of travel. In this case, the speed and direction must be specified.

A quantity which can be described fully by giving its magnitude is known as a **scalar quantity**. A **vector quantity** has magnitude and direction.

Some examples of scalar and vector quantities are given in Table 1.5.

Table 1.5 Some scalars and vectors.

quantity	scalar	vector
mass	✔	
weight		✔
speed	✔	
velocity		✔
acceleration		✔
force		✔
pressure	✔	
temperature	✔	

Note: It may seem that electric current should be treated as a vector quantity. We give current a direction when we deal with, for example, the motor effect (see Topic 22) and when we predict the direction of the magnetic field due to current-carrying coils and wires. However, electric current does not follow the laws of vector addition and should be treated as a scalar quantity.

Example

A 'big wheel' at a theme park has a diameter of 14 m and people on the ride complete one revolution in 24 s. Calculate:
- (a) the distance a rider moves in 3.0 minutes,
- (b) the distance of the rider from the starting position.

www.iranino.co

IRANINOVIN

(a) In 3.0 minutes, the rider completes $\frac{3.0 \times 60}{24} = 7.5$ revolutions.

distance travelled = 7.5 × circumference of wheel

$= 7.5 \times 2\pi \times 7.0$

$= \mathbf{330\,m}$

(b) 7.5 revolutions are completed. Rider is ½ revolution from the starting point. The rider is at the opposite end of a diameter of the big wheel. So, the distance from starting position = **14 m**.

Now it's your turn

13 State whether the following quantities are scalars or vectors:
- (a) time of departure of a train,
- (b) gravitational field strength,
- (c) density of a liquid.

14 State whether the following quantities are scalars or vectors:
- (a) movement of the hands of a clock,
- (b) frequency of vibration,
- (c) flow of water in a pipe.

15 Speed and velocity have the same units. Explain why speed is a scalar quantity whereas velocity is a vector quantity.

16 A student states that a bag of sugar has a weight of 10 N and that this weight is a vector quantity. Discuss whether the student is correct when stating that weight is a vector.

Vector representation

When you hit a tennis ball, you have to judge the direction you want it to move in, as well as how hard to hit it. The force you exert is therefore a vector quantity and cannot be represented by a number alone. One way to represent a vector is by means of an arrow. The direction of the arrow is the direction of the vector quantity. The length of the arrow, drawn to scale, represents its magnitude. This is illustrated in Figure 1.10.

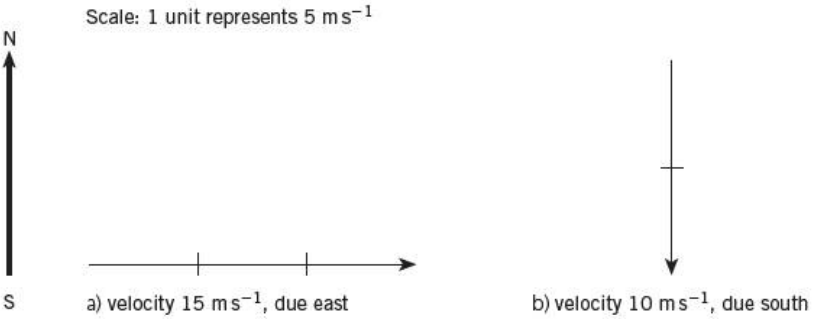


Figure 1.10 Representation of a vector quantity

Addition of vectors

The addition of two scalar quantities which have the same unit is no problem. The quantities are added using the normal rules of addition. For example, a beaker of volume 250 cm³ and a bucket of volume 9.0 litres have a total volume of 9250 cm³.

Adding together two vectors is more difficult because they have direction as well as magnitude. If the two vectors are in the same direction, then they can simply be added together. Two objects of weight 50 N and 40 N have a combined weight of 90 N because both weights act in the same direction (vertically downwards). Figure 1.11 shows the effect of adding two forces of magnitudes 30 N and 20 N which act along the same line in the same direction or in opposite directions. The angle between the forces is 0° when they act in the same direction and 180° when they are in opposite directions. For all other angles between the directions of the forces, the combined effect, or **resultant**, is some value between 10 N and 50 N.

www.iranino.co

IRANINOVIN

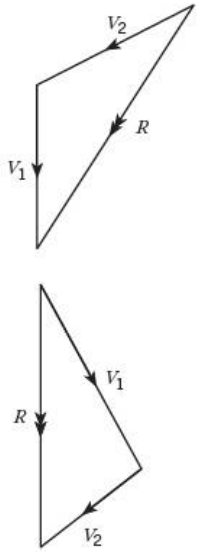


Figure 1.12 Vector triangles

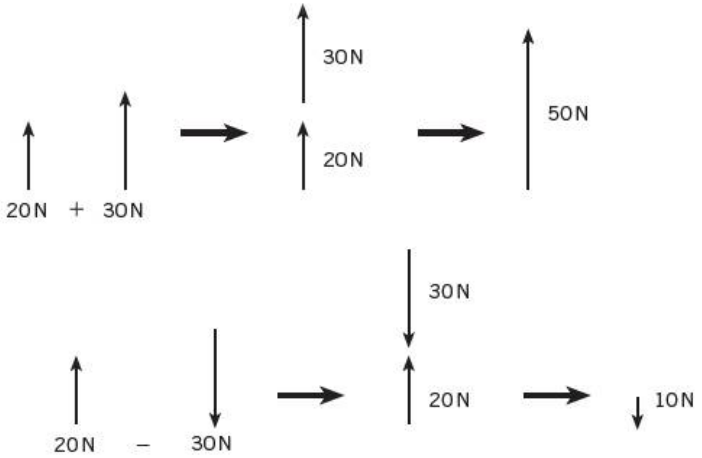


Figure 1.11 Vector addition

In cases where the two vectors do not act in the same or opposite directions, the resultant is found by means of a **vector triangle**. Each one of the two vectors V_1 and V_2 is represented in magnitude and direction by the side of a triangle. Note that both vectors must be in either a clockwise or an anticlockwise direction (see Figure 1.12). The combined effect, or resultant R, is given in magnitude and direction by the third side of the triangle. It is important to remember that, if V_1 and V_2 are drawn clockwise, then R is anticlockwise; if V_1 and V_2 are anticlockwise, R is clockwise.

The resultant may be found by means of a scale diagram. Alternatively, having drawn a sketch of the vector triangle, the problem may be solved using trigonometry (see the Maths Note on page 14).

Example

A ship is travelling due north with a speed of 12 km h^{-1} relative to the water. There is a current in the water flowing at 4.0 km h^{-1} in an easterly direction. Determine the velocity of the ship by:

(a) scale drawing,
(b) calculation.

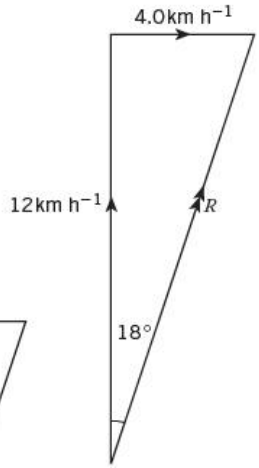


Figure 1.13

(a) By scale drawing (Figure 1.13):
Scale: 1 cm represents 2 km h^{-1}
resultant R
The velocity is:
6.3 × 2 = **12.6 km h^{-1}** in a direction **18° east of north**.

(b) By calculation:
Referring to the diagram (Figure 1.14) and using Pythagoras' theorem,

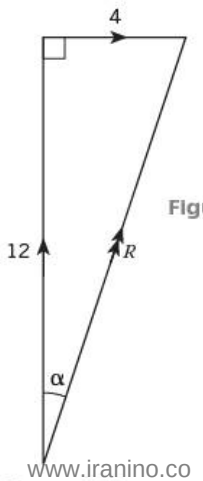
$$R^2 = 12^2 + 4^2 = 160$$

$$R = \sqrt{160} = 12.6$$

$$\tan\alpha = \frac{4}{12} = 0.33$$

$$\alpha = 18.4°$$

The velocity of the ship is **12.6 km h^{-1}** in a direction **18.4° east of north**.



www.iranino.co

Figure 1.14

Now it's your turn

17 Explain how an arrow may be used to represent a vector quantity.

18 Two forces are of magnitude 450 N and 240 N respectively. Determine:
(a) the maximum magnitude of the resultant force,
(b) the minimum magnitude of the resultant force,
(c) the resultant force when the forces act at right angles to each other.
Use a vector diagram and then check your result by calculation.

IRANINOVIN

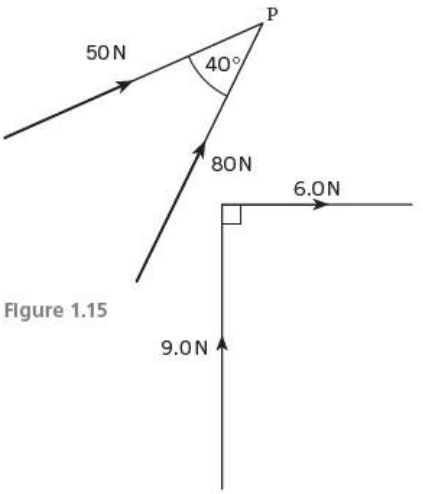


Figure 1.15

Figure 1.16

19 A boat can be rowed at a speed of 7.0 km h^{-1} in still water. A river flows at a constant speed of 1.5 km h^{-1}. Use a scale diagram to determine the angle to the bank at which the boat must be rowed in order that the boat travels directly across the river.

20 Two forces act at a point P as shown in Figure 1.15. Draw a vector diagram, to scale, to determine the resultant force. Check your work by calculation.

21 A swimmer who can swim in still water at a speed of 4 km h^{-1} is swimming in a river. The river flows at a speed of 3 km h^{-1}. Calculate the speed of the swimmer relative to the river bank when she swims:
(a) downstream,
(b) upstream.

22 Draw to scale a vector triangle to determine the resultant of the two forces shown in Figure 1.16. Check your answer by calculating the resultant.

The use of a vector triangle for finding the resultant of two vectors can be demonstrated by means of a simple laboratory experiment. A weight is attached to each end of a flexible thread and the thread is then suspended over two pulleys, as shown in Figure 1.17. A third weight is attached to a point P near the centre of the thread. The string moves over the pulleys and then comes to rest. The positions of the threads are marked on a piece of paper held on a board behind the threads. This is easy to do if light from a small lamp is shone at the board. Having noted the sizes W_1 and W_2 of the weights on the ends of the thread, a vector triangle can then be drawn on the paper, as shown in Figure 1.18. The resultant of W_1 and W_2 is found to be equal in magnitude but opposite in direction to the weight W_3. If this were not so, there would be a resultant force at P and the thread and weights would move. The use of a vector triangle is justified. The three forces W_1, W_2 and W_3 are in equilibrium. The condition for the vector diagram of these forces to represent the equilibrium situation is that the three vectors should form a **closed triangle**.

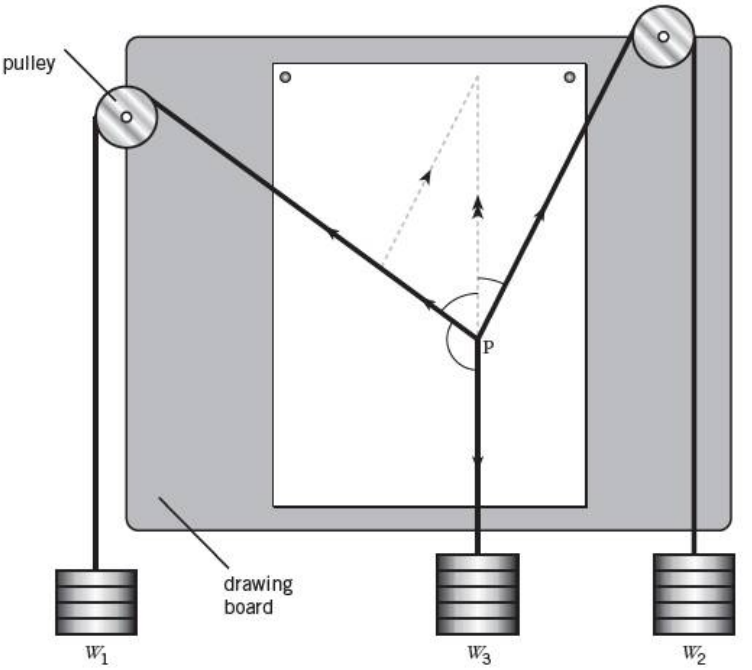


Figure 1.17 Apparatus to check the use of a vector triangle.

www.iranino.co

IRANINOVIN

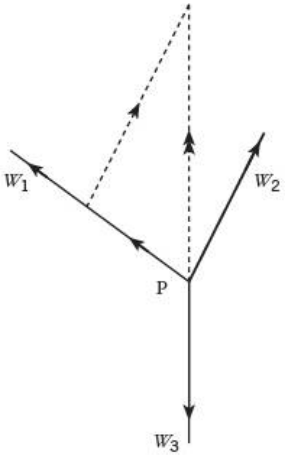


Figure 1.18 The vector triangle

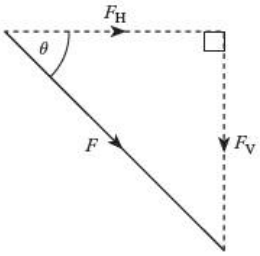


Figure 1.19 Resolving a vector into components



Figure 1.20

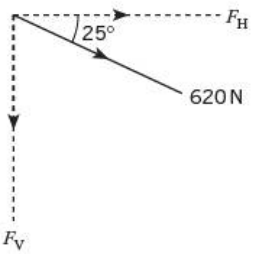


Figure 1.21

We have considered only the addition of two vectors. When three or more vectors need to be added, the same principles apply, provided the vectors are **coplanar** (all in the same plane). The vector triangle then becomes a vector polygon: the resultant forms the missing side to close the polygon.

To subtract two vectors, reverse the direction (that is, change the sign) of the vector to be subtracted, and then add.

Resolution of vectors

On pages 8–10 we saw that two vectors may be added together to produce a single resultant. This resultant behaves in the same way as the two individual vectors. It follows that a single vector may be split up, or **resolved**, into two vectors, or **components**. The combined effect of the components is the same as the original vector. In later chapters, we will see that resolution of a vector into two perpendicular components is a very useful means of solving certain types of problem.

Consider a force of magnitude F acting at an angle of θ below the horizontal (see Figure 1.19). A vector triangle can be drawn with a component F_H in the horizontal direction and a component F_V acting vertically. Remembering that F, F_H and F_V form a right-angled triangle, then

$$F_H = F\cos\theta$$

$$\text{and } F_V = F\sin\theta$$

The force F has been resolved into two perpendicular components, F_H and F_V. The example chosen is concerned with forces, but the method applies to all types of vector quantity.

Example

A glider is launched by an aircraft with a cable, as shown in Figure 1.20. At one particular moment, the tension in the cable is 620N and the cable makes an angle of 25° with the horizontal (see Figure 1.21). Calculate:

(a) the force pulling the glider horizontally,
(b) the vertical force exerted by the cable on the nose of the glider.

(a) horizontal component $F_H = 620 \cos 25 =$ **560 N**
(b) vertical component $F_V = 620 \sin 25 =$ **260 N**

Now it's your turn

23 An aircraft is travelling 35° east of north at a speed of 310km h^{-1}. Calculate the speed of the aircraft in:
(a) the northerly direction,
(b) the easterly direction.

24 A cyclist is travelling down a hill at a speed of 9.2m s^{-1}. The hillside makes an angle of 6.3° with the horizontal. Calculate, for the cyclist:
(a) the vertical speed,
(b) the horizontal speed.

www.iranino.co

IRANINOVIN

Summary

- All physical quantities have a magnitude (size) and a unit.
- The SI base units of mass, length, time, electric current, thermodynamic temperature and amount of substance are the kilogram, metre, second, ampere, kelvin and mole respectively.
- Units of all mechanical, electrical, magnetic and thermal quantities may be derived in terms of these base units.
- Physical equations must be homogeneous (balanced). Each term in an equation must have the same base units.
- The convention for printing headings in tables of data, and for labelling graph axes, is the symbol for the physical quantity (in *italic*), followed by a forward slash, followed by the abbreviation for the unit (in roman). In handwriting, one cannot distinguish between italic and roman type.
- The order of magnitude of a number is the power of ten to which the number is raised. The order of magnitude can be used to make a check on whether a calculation gives a sensible answer.
- A scalar quantity has magnitude only.
- A vector quantity has magnitude and direction.
- A vector quantity may be represented by an arrow, with the length of the arrow drawn to scale to give the magnitude.
- The combined effect of two (or more) vectors is called the resultant.
- Coplanar vectors may be added (or subtracted) using a vector diagram.
- The resultant may be found using a scale drawing of the vector diagram, or by calculation.
- A single vector may be divided into two separate components.
- The dividing of a vector into components is known as the resolution of the vector.
- In general, a vector is resolved into two components at right angles to each other.

Examination style questions

1 **a** **i** Explain what is meant by a *base unit*.

ii Give four examples of base units.

b State what is meant by a *derived unit*.

c **i** For any equation to be valid, it must be *homogeneous*. Explain what is meant by a homogeneous equation.

ii The pressure p of an ideal gas of density ρ is given by the equation

$$p = \frac{1}{3}\rho \langle c^2 \rangle$$

where $\langle c^2 \rangle$ is the mean-square-speed (i.e. it is a quantity measured as [speed]^2).

Use base units to show that the equation is homogeneous.

2 The period T of a pendulum of mass M is given by the expression

$$T = 2\pi \sqrt{\frac{I}{Mgh}}$$

where g is the gravitational field strength $\left(\frac{\text{force}}{\text{mass}}\right)$ and h is a length.

Determine the base units of the quantity I.

3 **a** Determine the base units of:

i work done,

ii the moment of a force.

b Explain why your answers to **a** mean that caution is required when the homogeneity of an equation is being tested.

4 **a** Distinguish between a *scalar* and a *vector* quantity.

b A mass of weight 120 N is hung from two strings as shown in Fig. 1.22.

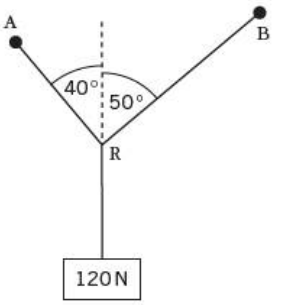


Fig. 1.22

Determine, by scale drawing or by calculation, the tension in:

i RA,

ii RB.

www.iranino.co

IRANINOVIN

c Use your answers in **b** to determine the horizontal component of the tension in:

i RA,
ii RB.
Comment on your answer.

5 A fielder in a cricket match throws the ball to the wicket-keeper. At one moment of time, the ball has a horizontal velocity of 16 m s^{-1} and a velocity in the vertically upward direction of 8.9 m s^{-1}.

a Determine, for the ball:

i its resultant speed,
ii the direction in which it is travelling relative to the horizontal.

b During the flight of the ball to the wicket-keeper, the horizontal velocity remains unchanged. The speed of the ball at the moment when the wicket-keeper catches it is 19 m s^{-1}. Calculate, for the ball just as it is caught:

i its vertical speed,
ii the angle that the path of the ball makes with the horizontal.

c Suggest with a reason whether the ball, at the moment it is caught, is rising or falling.

6 **a** The spacing between two atoms in a crystal is 3.8×10^{-10} m. State this distance in pm. *[1]*

b Calculate the time of one day in Ms. *[1]*

c The distance from the Earth to the Sun is 0.15 Tm. Calculate the time in minutes for light to travel from the Sun to the Earth. *[2]*

d Identify all the vector quantities in the list below.

distance energy momentum
weight work *[1]*

e The velocity vector diagram for an aircraft heading due north is shown to scale in Fig. 1.23. There is a wind blowing from the north-west.

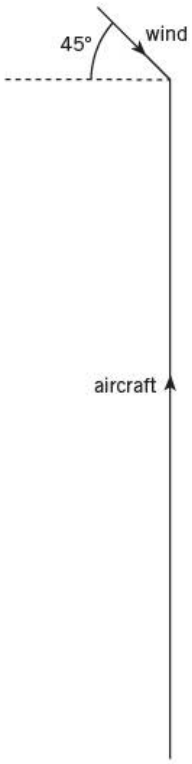


Fig. 1.23

The speed of the wind is 36 m s^{-1} and the speed of the aircraft is 250 m s^{-1}.

i Make a copy of Fig. 1.23. Draw an arrow to show the direction of the resultant velocity of the aircraft. *[1]*

ii Determine the magnitude of the resultant velocity of the aircraft. *[2]*

Cambridge International AS and A Level Physics, 9702/23 Oct/Nov 2012 Q 1

7 **a** **i** State the SI base units of volume. *[1]*

ii Show that the SI base units of pressure are kg m^{-1} s^{-2}. *[1]*

b The volume V of liquid that flows through a pipe in time t is given by the equation

$$\frac{V}{t} = \frac{\pi P r^4}{8Cl}$$

where P is the pressure difference between the ends of the pipe of radius r and length l. The constant C depends on the frictional effects of the liquid.

Determine the base units of C. *[3]*

Cambridge International AS and A Level Physics, 9702/21 May/June 2012 Q 1

8 Make reasonable estimates of the following quantities.

a the frequency of an audible sound wave *[1]*

b the wavelength, in nm, of ultraviolet radiation *[1]*

c the mass of a plastic 30 cm ruler *[1]*

d the density of air at atmospheric pressure *[1]*

Cambridge International AS and A level Physics, 9702/02 May/June 2008 Q 1

9 **a** Two of the SI base quantities are mass and time. State three other SI base quantities. *[3]*

b A sphere of radius r is moving at speed v through air of density ρ. The resistive force F acting on the sphere is given by the expression

$$F = Br^2\rho v^k$$

where B and k are constants without units.

i State the SI base units of F, ρ and v. *[3]*

ii Use base units to determine the value of k. *[2]*

Cambridge International AS and A Level Physics, 9702/21 Oct/Nov 2010 Q 1

www.iranino.co

IRANINOVIN

Maths Note

Sine rule

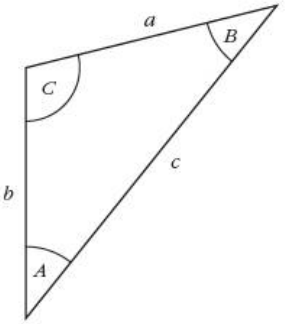


Figure 1.24

For any triangle (Figure 1.24),

$$\frac{a}{\sin A} = \frac{b}{\sin B} = \frac{c}{\sin C}$$

Cosine rule

For any triangle,

$a^2 = b^2 + c^2 - 2bc \cos A$

$b^2 = a^2 + c^2 - 2ac \cos B$

$c^2 = a^2 + b^2 - 2ab \cos C$

Pythagoras' theorem

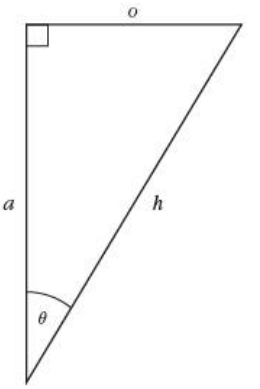


Figure 1.25

For a right-angled triangle (Figure 1.25),

$h^2 = o^2 + a^2$

Also for a right-angled triangle:

$\sin \theta = \frac{o}{h}$

$\cos \theta = \frac{a}{h}$

$\tan \theta = \frac{o}{a}$

www.iranino.co

IRANINOVIN

2 Measurement techniques

By the end of this topic, you will be able to:

2.1 **(a)** use techniques for the measurement of length, volume, angle, mass, time, temperature and electrical quantities, all to appropriate ranges of magnitude. In particular, you should be able to:
- measure lengths using a ruler, calipers and micrometer
- measure weight and hence mass using balances
- measure an angle using a protractor
- measure time intervals using clocks, stopwatches and the calibrated time-base of a cathode-ray oscilloscope (c.r.o.)
- measure temperature using a thermometer
- use ammeters and voltmeters with appropriate scales
- use a galvanometer in null methods
- use a cathode-ray oscilloscope (c.r.o.)
- use a calibrated Hall probe

(b) use both analogue scales and digital displays

(c) use calibration curves

2.2 **(a)** understand and explain the effects of systematic errors (including zero errors) and random errors in measurement

(b) understand the distinction between precision and accuracy

(c) assess the uncertainty in a derived quantity by simple addition of absolute, fractional or percentage uncertainties

Note: use of a calibrated Hall probe is required only in the A level course but is included here for completeness.

Starting points

- Theoretical ideas in physics are generally tested by experiment before being fully accepted.
- Experimental work is an important part of a physics course.
- Make a sensible choice of the instrument to use to measure a particular physical quantity.
- There are sources of error and uncertainty in experimental work.

2.1 Measurements

All experiments that are designed to obtain a quantitative result for a physical quantity involve measurements. These measurements must be of some combination of the base quantities length, mass, time, temperature and current. (To complete the list, we should include quantity of substance and luminous intensity, but these are not encountered in experimental work in A/AS Physics.) In the following sections we will look at the methods available for measuring the base quantities in a school or college laboratory. By understanding the principles of the available methods, we will be able to make an informed decision about the choice of a particular technique, with respect to making the experiment as precise and reproducible as possible, and avoiding sources of systematic error. For all of the quantities, the effective choice will be limited by what instruments are available in your laboratory. However, in one type of examination question, on planning and design, you may be asked to devise an experiment and draw on your theoretical, rather than practical, knowledge of various types of apparatus.

At A/AS level, students generally assume that the calibration of the instruments they use is correct. However, it is worth thinking about how to compare the calibration of one instrument against another, even if this is a check you will very seldom make.

www.iranino.co

IRANINOVIN



Figure 2.1 Students doing an experiment in a physics laboratory

In a planning/design question, you might be asked to suggest a method of calibration. Generally it is easy to compare the calibration of different instruments, but not so easy to determine which of two instruments giving different readings is correct.

On the pages that follow, we will spend most time on methods for measuring length, because length-measuring instruments of several different types will be available in your laboratory.

Methods of measuring length

The metre rule

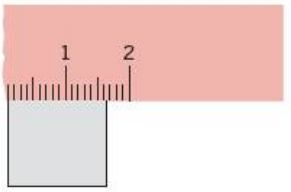
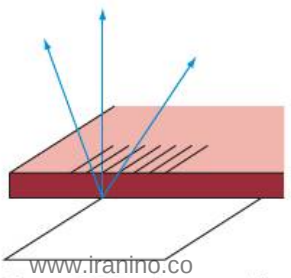


Figure 2.2 Zero error with a metre rule

The simplest length-measuring instrument to be found in your laboratory is a metre (or half-metre) rule. It has the great advantages of being cheap, convenient and simple to use. A relatively unskilled student should have no difficulty in taking a reading with an uncertainty of 0.5 mm. However, you should be aware of three possible sources of error in using a metre rule.

The first may arise if the end of the rule is worn, giving rise to a zero error (Figure 2.2). For this reason, it is bad practice to place the zero end of the rule against one end of the object to be measured and to take the reading at the other end. You should place the object against the rule so that a reading is made at each end of the object. The length of the object is then obtained by subtraction of the two readings. A zero error like this is a systematic error, because it is involved every time a reading is taken from the zero end of the rule. (A more detailed explanation of systematic errors will be found on page 35.) In general, the zero reading of any instrument may be subject to an error. We shall meet this type of error again in the micrometer screw gauge and in the ammeter.

The calibration of the metre rule may give rise to another systematic error because the markings are incorrect. Try comparing the 30 cm graduated length of one plastic rule with the same nominal length on another. You are quite likely to find a discrepancy of one or two millimetres. One of the reasons why wooden or plastic metre rules are cheap is that the manufacturer does not claim any great accuracy for the scale markings. The calibration may be checked by laying the rule alongside a more accurate rule, such as a steel 30 cm or metre rule, and noting any discrepancy. If you compare an engineer's steel rule with a plastic rule, you will see at once that the engraved marks on the steel rule are much finer than the impressed marks on the plastic. Of course, the extra care which has been taken in engraving the steel rule has to be paid for. A one-metre steel rule is many times more expensive than a plastic metre rule.



www.iranino.co

Figure 2.3 Parallax error with a metre rule

Another source of error with the metre rule is parallax error. If the object to be measured is not on the same level as the graduated surface of the rule, the angle at which the scale is viewed will affect the result (Figure 2.3). This is a random error (see page 35), because the angle of view may be different for different readings. It may

IRANINOVIN

be reduced by arranging the rule so that there is no gap between the scale and the object. Parallax error is also important in reading any instrument in which a needle moves over a scale. A rather sophisticated way of eliminating parallax error is to place a mirror alongside the scale. When the needle and scale are viewed directly, the needle and its image in the mirror coincide. This ensures that the scale reading is always taken at the same viewing angle.

The smallest division on the metre rule is 1 mm. If you take precautions to avoid parallax error, you should be able to estimate a reading to about 0.5 mm. If you are measuring the length of an object by taking a reading at each end, the uncertainties add to give a total uncertainty of 1 mm. The range of the metre rule is from 1 mm to 1000 mm. To measure a length of more than 1 m with a metre rule will introduce a further uncertainty, of perhaps 1 mm or 2 mm, because of the difficulty of making a reference mark at the 1 m end of the rule and moving the rule so that the zero exactly corresponds with this reference. It is usually better to use a steel tape to measure lengths of more than 1 m.

The micrometer screw gauge

The type of micrometer screw gauge available in a school or college laboratory may be used to measure the dimensions of objects up to a maximum of about 50 mm. Measurements can easily be made with an uncertainty of 10 μm or less. The principle of the instrument is the magnification of linear motion using the circular motion of a screw. The instrument consists of a U-shaped piece of steel with a fixed, plane, end-piece A (see Figure 2.4). Opposite this is a screw with a corresponding end-piece B. The position of the screw can be adjusted using the ratchet C which is connected to the thimble D. There are graduations along the barrel of the instrument (the bearing in which the screw turns), and round the circumference of the thimble. The purpose of the ratchet is to ensure that the same torque (that is, the amount of twist) is applied to the thimble for each reading. If this torque is exceeded, the ratchet slips. The object to be measured is placed in the jaws of the gauge between end-pieces A and B, and B is screwed down on to the object, using the ratchet C, until the ratchet slips.

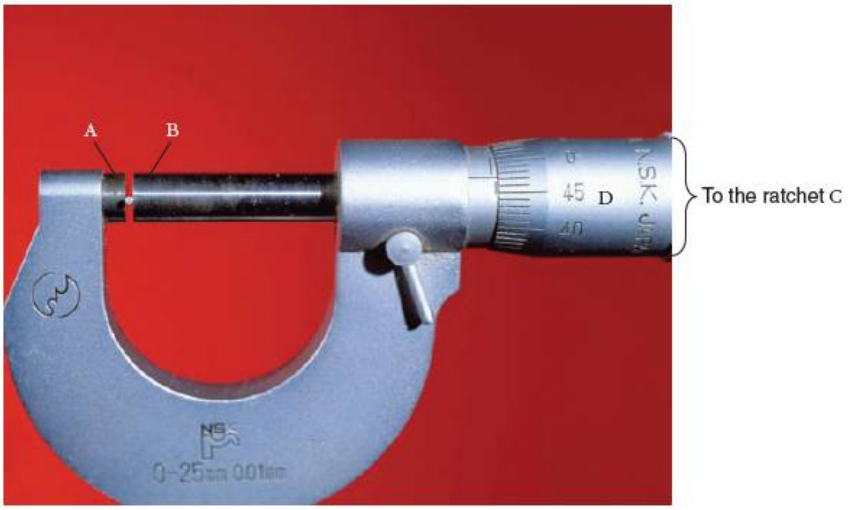


Figure 2.4 Micrometer screw gauge

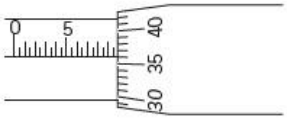


Figure 2.5 Screw gauge scales with a reading of 9.86 mm

The screw advances exactly 1 mm for two revolutions. That is, the pitch of the screw is 0.5 mm or 500 μm. If you look at the graduations on the barrel of the screw bearing you will see that there are divisions every 0.5 mm. The reading on the barrel corresponds to the position of the edge of the thimble (see Figure 2.5). When taking a reading it is important to check which half of the millimetre the edge of the barrel is in. The graduations round the circumference of the thimble run from 0 to 50. Each division corresponds to one-hundredth of a mm, or 10 μm. The reading on the thimble

www.iraninov.co

IRANINOVIN

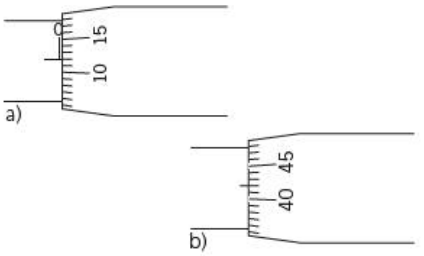


Figure 2.6 a) zero error of + 0.12 mm b) zero error of – 0.08 mm

is added to the reading on the barrel. Thus, Figure 2.5 shows a reading of 9.5 mm on the barrel plus 0.36 mm on the thimble, 9.86 mm in total. You can easily read to the nearest division on the thimble; that is, to the nearest 0.01 mm (10 µm).

The micrometer screw gauge is very likely to have a systematic zero error. Every time you use the gauge, you should check the zero error by moving face B so that it makes contact with face A. The screw must be tightened with the ratchet C, so that a reproducible zero is obtained. Then take the reading on the barrel and the thimble. This gives the zero error, which must be allowed for in all subsequent readings. Figure 2.6a shows a screw gauge with a zero error of +0.12 mm. If this were the error which applied when the reading of 9.86 mm was obtained in Figure 2.5, the true length of the object would be 9.86 mm – 0.12 mm = 9.74 mm. Figure 2.6b shows a zero error of –0.08 mm. In this case, the true length of the object in Figure 2.5 would be 9.86 mm + 0.08 mm = 9.94 mm.

In the case of a wooden or plastic metre rule, it is good practice to check the calibration against an engineer's steel rule, if one is available in your laboratory. It would be unusual to do the same with a micrometer screw gauge, but if there is doubt about the calibration of a particular gauge, it can be checked by measuring the dimensions of a series of gauge blocks. A gauge block is a rectangular steel block with faces which are accurately plane and parallel. The length of the gauge block is known to an uncertainty of less than 1 µm. However, not many school laboratories possess gauge blocks.

The vernier caliper

A vernier caliper is a versatile instrument for measuring the dimensions of an object, the diameter of a hole, or the depth of a hole. Its range is up to about 100 mm, and it can be read to 0.1 mm or 0.05 mm depending on the type of vernier with which it is fitted. It consists of a steel mm scale A with two reference posts at the zero mark (see Figure 2.7).

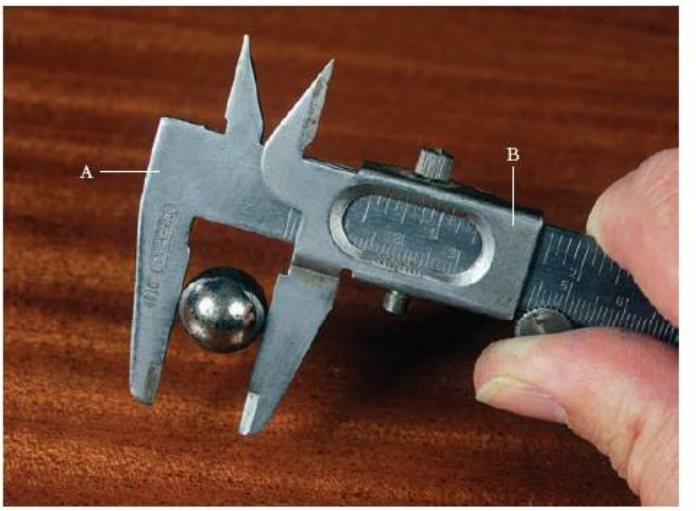


Figure 2.7 Vernier caliper

A sliding part B moves along the scale. The slider has the vernier scale engraved on it. The zero of the vernier corresponds with reference posts on the sliding part. One set of reference posts, those with the straight parts on the inside, is used like the jaws of a screw gauge: the object to be measured is placed between the jaws or reference posts, and the sliding part B is moved along until the object is gripped tightly. A reading to the nearest mm is taken on the fixed scale, at the zero end of the vernier scale. The reading to a tenth of a mm is obtained by finding where a graduation of the vernier scale coincides with a graduation of the fixed scale. Figure 2.8 shows the scale of a vernier caliper giving a reading of 25.4 mm.

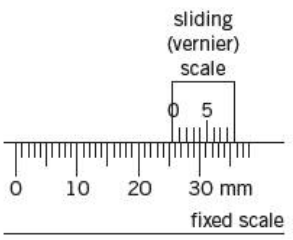


Figure 2.8 Vernier scale with a reading of 25.4 mm

www.iranino.co

IRANINOVIN

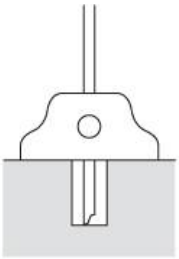


Figure 2.9 Measurement of the depth of a blind hole

The second set of jaws has the straight parts on the outside. These can be used to measure the diameter of a hole. The jaws are placed inside the hole and are moved apart until they are in contact with the edges of the hole. The scale and vernier can then be read.

A pin at the end of the sliding part of the caliper can be used to measure the depth of a blind hole; for example, a hole which has been drilled in, but not right through, a wooden board. The end of the fixed scale is placed on the board, across the hole, and the pin moved into the hole until it reaches the bottom (see Figure 2.9). The reading of the scale and vernier gives the depth of the hole.

As with the micrometer screw gauge, the vernier caliper should be checked for a systematic zero error before taking a reading.

The learning objectives for this course do not require students to have knowledge of the vernier scale for the theory papers. However, some laboratories may be equipped with vernier calipers for practical work.

Examples

1 Figure 2.10a shows the scale of a micrometer screw gauge when the zero error is being checked and Figure 2.10b shows the scale when the gauge is tightened on an object. What is the length of the object?

From Figure 2.10a, the zero error is +0.12 mm. The reading in Figure 2.10b is 15.62 mm. The length of the object is thus (15.62 – 0.12) mm = **15.50 mm**.

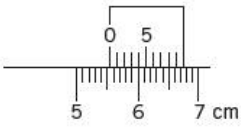


Figure 2.11

a)

b)

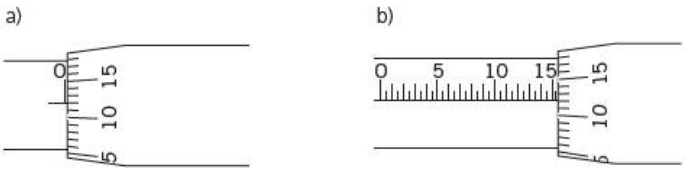


Figure 2.10 a) and b)

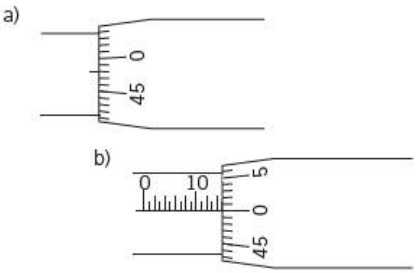


Figure 2.12 a) and b)

2 Figure 2.11 shows the scale of a vernier caliper. What is the reading?

The zero of the vernier scale is between the 5.5 cm and 5.6 cm divisions of the fixed scale. There is coincidence between the third graduation of the vernier scale and one of the graduations of the fixed scale. The reading is thus **5.53 cm or 55.3 mm**.

Now it's your turn

1 Figures 2.12a and 2.12b show the scales of a micrometer screw gauge when the zero is being checked, and again when measuring the diameter of an object. What is the diameter?

Choice of method

A summary of the range and reading uncertainty of length-measuring instruments is given in Table 2.1.

Table 2.1 Length-measuring instruments

Instrument	range	uncertainty in length	notes
metre rule	1 m	1 mm	check zero, calibration errors
micrometer screw gauge	50 mm	0.01 mm	check zero error
vernier caliper	100 mm	0.1 mm	versatile: inside and outside diameters, depth

In deciding which instrument to use in a particular experiment, you should consider first the nature of the length measurement you have to make. For example, if you need to find the diameter of a steel sphere, the screw gauge and caliper techniques are obvious candidates. You should then consider whether you need the greater precision

www.iranino.co

IRANINOVIN

of the micrometer. In a particular experiment, the uncertainty in the diameter of the sphere may be the dominant uncertainty (see Section 2.2 Errors and uncertainties page 31), and in such a case the fact that the precision available with the screw gauge is ten times that for the vernier caliper will decide the argument. In an experiment which may last some time, you should also think about the resources of your laboratory. Is it sensible to use what may be one of only a small number of available screw gauges, when they may also be in demand by other students for other experiments? Sometimes, in design questions, you are asked to think about the cost of setting up an experiment. We have mentioned the difference in cost of a steel metre rule compared with wooden or plastic rules. It would be foolish economics to supply a steel rule for each of a number of students, when it would be perfectly adequate to provide each of them with a wooden rule and have one steel rule available in the laboratory for calibration purposes.

Application: measurement of pressure difference

A difference in gas pressure may be measured by comparing the heights of liquid in the two arms of a U-tube. Figure 2.13 shows a U-tube connected to a container of gas. The pressure above the liquid in tube A is atmospheric pressure p_{atm}. The pressure above the liquid in tube B, and hence the pressure of gas in the container, is p. The relation between the pressures is

$$p = p_{atm} + \Delta h \rho g$$

where Δh is the difference in vertical height between the levels of the liquid in the two arms of the tube, ρ is the density of the liquid, and g is the acceleration of free fall. To find the pressure of the gas, all we need to do is to measure the difference Δh between the liquid levels, assuming that p_{atm}, ρ (and g) are known.

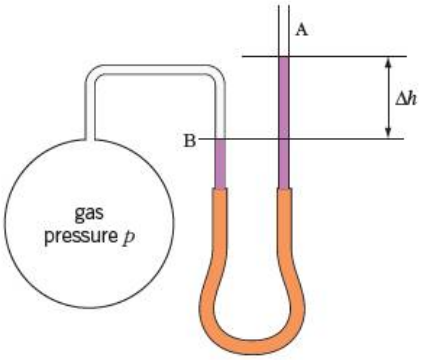


Figure 2.13

In some laboratories, a U-tube mounted on a board to which a vertical millimetre scale is attached may be available. This device is called a manometer. It is then a simple matter to measure Δh. If the manometer contains oil or water, the liquid in the tube will have a curved surface which is concave downwards (Figure 2.14a). This surface is called the **meniscus**. Use a set-square to find the reading on the vertical scale corresponding to the *bottom* of the meniscus in each side of the U-tube. The surface of the liquid in a manometer filled with mercury will be convex (Figure 2.14b), and in this case you should read to the *top* of the meniscus. Again, use a set-square to avoid parallax error. If a manometer is not available, you will have to arrange your own system of U-tube and scale; for example, a half-metre rule. Make sure that the scale is clamped vertically.

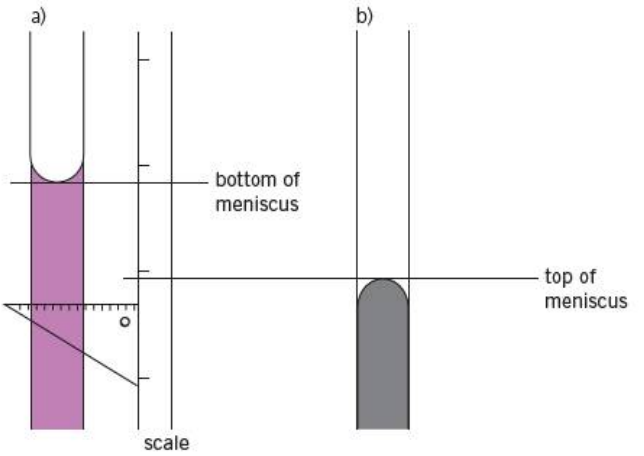


Figure 2.14 a) and b)

www.iranino.co

IRANINOVIN

Methods of measuring mass

The method of measuring mass is with a balance. In fact, balances compare the *weight* of the unknown mass with the weight of a standard mass. But because weight is proportional to mass, equality between the unknown weight and the weight of the standard mass means that the unknown mass is equal to the standard mass.

In your laboratory, you may have access to a number of different types of balance, including the top-pan balance, the lever balance and the spring balance. It is important that you should familiarise yourself with the use of all types that are available to you, so that you do not restrict your choice to one particular type. Note also that some types of spring balance may be calibrated in force units (that is, in newton) rather than in mass units (kilogram).



Figure 2.15 Top-pan balance

The top-pan balance

The top-pan balance (Figure 2.15) is a direct-reading instrument, based on a pressure sensor, or sometimes a spring. The unknown mass is placed on the pan, and its weight applies a force to the sensor. The mass corresponding to this force is displayed on a digital read-out.

When using the balance, ensure that the initial (unloaded) reading is zero. There is a control for adjusting the zero reading. The balance may have a tare facility, for use in backing off the mass of an empty container so that the mass of material added to the container is obtained directly. This works in the same way as adjusting the balance for zero error.

The uncertainty in the reading of a particular top-pan balance will be quoted in the manufacturer's manual. As with other digital instruments, it is likely to be expressed as a percentage uncertainty of the reading shown on the scale, together with the uncertainty in the final figure of the display.

The spring balance and the lever balance

Spring balances (Figure 2.16) are based on Hooke's law (see Topic 9): the extension of a loaded spring is proportional to the load. The extension is measured directly, by a marker moving along a straight scale, or by a pointer moving over a circular scale. As with any instrument using a scale and pointer, you should take care not to introduce a parallax error when you take readings. Position yourself so that your line of sight is perpendicular to the scale. Before placing the object of unknown mass on the pan, check for zero error. There is likely to be a zero-error adjustment screw on the balance.

Lever balances are based on the principle of moments. In one common type (Figure 2.17), the unknown mass is placed on a pan, and balance is achieved by sliding a mass along a bar, calibrated in mass units, until the bar is horizontal. This represents the condition in which the moment of the load is equal and opposite to the moment of the sliding mass and the bar. A reading is taken from the edge of the sliding mass on the divisions marked on the bar. In this case, parallax error is less likely to be serious. Again, check for zero error before taking a reading.



Figure 2.16 Spring balance

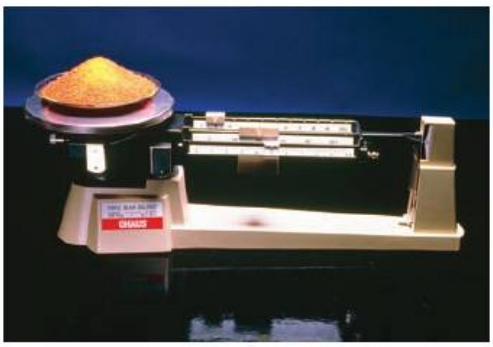


Figure 2.17 Lever balance



Figure 2.18 Lever balance with circular scale

www.iranino.co

IRANINOVIN

Another type of lever balance has a pointer moving along a circular scale (Figure 2.18). A weight on the pointer arm is placed in one of two positions in order to change the range (for example, from 0–100 g to 0–1 kg).

Both of these types of balance are used more for the convenience of obtaining a rapid, approximate reading, rather than for an accurate determination. An indication of the uncertainty involved in readings with a particular balance can be obtained from the smallest division on the scale.

Example

The mass of a quantity of chemical is determined using a lever balance. Over the range of masses involved, the separation between mass graduations on the bar is 2 g. The reading for the mass of the empty container is 56 g, and the reading for the mass of the container plus the chemical is 104 g. Find the mass of the chemical, and the uncertainty in this value.

By subtraction, the mass of the chemical is 104 – 56 = **48 g**.

The uncertainty in each reading is likely to be half of the smallest division of the mass graduations on the beam that is ±1 g. Each of the two readings has an uncertainty of ±1 g: the uncertainty in the mass of the chemical is thus **±2 g**.

Now it's your turn

2 The mass of a chemical used to make up a solution is determined as follows. A dish containing the chemical is placed on the pan of a spring balance. The pointer reading on the scale is shown in Figure 2.19a. The chemical is then tipped into a known volume of water, and the empty dish replaced on the pan, giving the pointer reading shown in Figure 2.19b. What is the mass of chemical, and what is the uncertainty in this value?

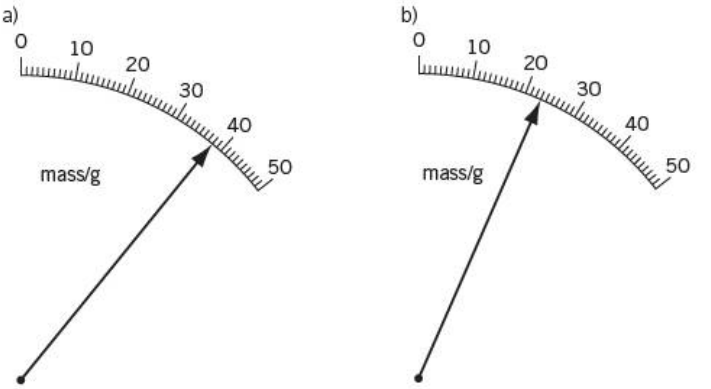


Figure 2.19

Choice of method

As stated above, the top-pan and the spring balance are direct-reading instruments. This means that readings can be obtained quickly and conveniently. The lever balance requires adjustment of the sliding mass, but this takes only a very short time. There may be some rules in your laboratory about which type of balance should be used for which task. In general, the pan of the balance should be kept clean and dry. Do not weigh out loose chemicals on the pan; always use a container, the mass of which you have determined beforehand, or for which you have made allowance using the tare control.

In general, choose a balance of sensitivity appropriate to the experiment you are carrying out.

Application: current balance

A U-shaped magnet is placed on a top-pan balance (Figure 2.20). A wire is clamped so that it runs along the channel of the magnet. The wire is connected in a circuit with a d.c. supply, a rheostat (variable resistor), an ammeter and a switch. When the supply is switched on, the balance reading is seen to change, because a force is exerted on the wire in the magnetic field. By Newton's third law, a force is also exerted on the

www.iranino.co

magnet. This is detected by the change Δm in the mass reading. This change must be converted into force F by multiplying Δm by g. The variation with current I of the magnetic force F may be determined.
The equation

$F = BIl$ (see Topic 22)

where l is the length of the wire in the magnetic field, may be verified. The direction of the force, as predicted theoretically by Fleming's left-hand rule, may also be verified by checking whether the mass reading increases or decreases for a given current direction.

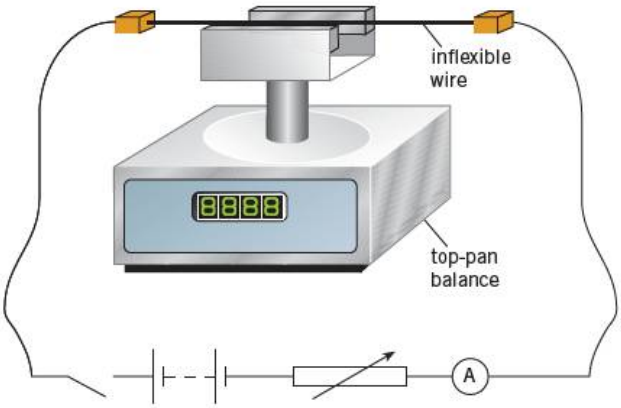


Figure 2.20 Current balance experiment

Measuring an angle

Angles are measured using an instrument called a protractor. This looks like a semi-circular, or sometimes circular, ruler, with its scale marked out in angular measure, invariably degrees rather than radians. The centre of the circle is clearly marked.

To measure the angle between two lines, the centre of the circle of the protractor is placed exactly over the point of intersection of the lines and one line is aligned with the 0° direction of the protractor (Figure 2.21). The angle between the lines is then given by the reading on the scale at which the second line passes through the circumference of the circle.

If the direction of a single line needs to be defined, this is always referred to the direction of the Ox axis, the horizontal axis pointing towards the right, as zero.

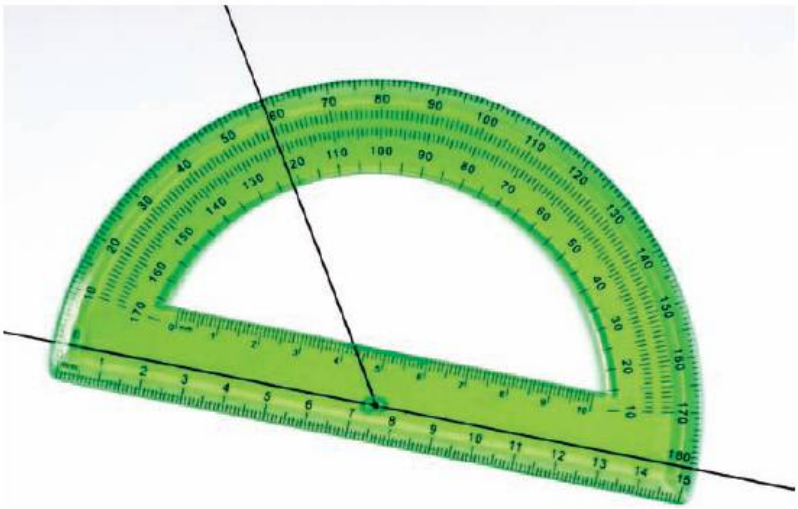


Figure 2.21 Using a protractor

www.iranino.co

IRANINOVIN

Most protractors used in schools and colleges have a diameter of about 10 cm. The size of the scale at the circumference of the circle is then such that the interval between scale divisions is 1°. It is easy to make a reading to the nearest degree, and sometimes to half a degree, if the line being measured is fine enough. Protractors of larger diameter may be marked in half degrees.

Methods of measuring time

The experiments you will meet in your practical physics course deal with the measurement of time intervals, rather than with absolute time. The basic method of measuring a time interval is with a stopclock or stopwatch. In each case, the instrument is started and stopped by pressing a lever or a button, and re-set by pressing another control. You should familiarise yourself with the way of operating the instrument before you start a timing experiment in earnest. Remember that the reaction time of the experimenter (a few tenths of a second) is likely to be much greater than the uncertainty of the instrument itself. If you do not reduce the effects of reaction time, an unacceptable systematic error may be built in to the experiment. As explained in the section on Errors and uncertainties (page 31), one way of reducing the effect of reaction time is to time enough events (for example, the swings of a pendulum) to make the interval being measured very much larger than the experimenter's reaction time. A good technique is to count the events (the swings), commencing by counting down to zero, and starting the timer at the zero count. Wherever possible, work with at least 20 seconds' worth of events (oscillations), and repeat each set of timings three times. (Sometimes, when carrying out experiments on damped oscillations, you will have to be satisfied with fewer swings, but try not to go below intervals of 10 seconds.)



Figure 2.22 Analogue stopclock

The stopclock

A mechanical, spring-powered stopclock will have an analogue display; that is, a hand (or hands) which move round a dial (Figure 2.22). Such an instrument is likely to read to the nearest one-fifth of a second.



Figure 2.23 Digital timer

The stopwatch or digital timer

This instrument has a digital display (Figure 2.23). It is based on the oscillations of a quartz crystal. The read-out will probably be to the nearest one-hundredth of a second. In addition to the start, stop and re-set controls, digital stopwatches often have a 'lap' facility, which allows one reading to be held in the display while the watch is still running. Because of this complexity, it is vital that you know the functions of all the controls.

Your own wristwatch may well have a built-in stopwatch, which may be just as precise as the watches available in the laboratory. However, the start and stop controls on wrist stopwatches are sometimes rather small, and it is important that you should not fumble a start or stop signal.

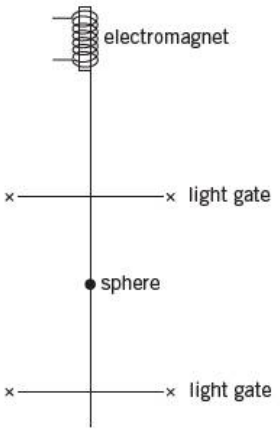


Figure 2.24 Determination of the acceleration of free fall

www.iranino.com

Choice of method

Often, students are attracted to a digital stopwatch because it reads to one-hundredth of a second. However, in all experiments in which the start and stop signals are applied manually, such precision is unnecessary and inappropriate. The reaction time of the experimenter, which is likely to be a few tenths of a second, will cancel out the precision of the watch. It would be misleading, and bad practice, to enter times such as '21.112 s' in a table of results if the systematic error due to reaction time had not been fully accounted for. Thus, if you are doing an experiment on the timing of oscillations and there are no digital timers available, you will be at no disadvantage if you have to use an analogue stopclock.

Application: determination of the acceleration of free fall

A steel sphere is released from an electromagnet and falls under gravity. As it falls, it passes through light gates which switch an electronic timer on and then off (Figure 2.24). The acceleration of free fall can be determined from the values of the time intervals and distances.

IRANINOVIN

This is an experiment in which electronic switching is essential in order to reduce the potentially very large error caused by the reaction time of the experimenter. Here, timing to one-hundredth of a second is essential.

Application: measurement of frequency using a cathode-ray oscilloscope

A cathode-ray oscilloscope (c.r.o.) has a calibrated time-base, so that measurements from the screen of the c.r.o. can be used to give values of time intervals. One application is to measure the frequency of a periodic signal, for example the sine-wave output of a signal generator. The signal is connected to the Y-input of the c.r.o., and the Y-amplifier and time-base controls are adjusted until a trace of at least one, but fewer than about five, complete cycles of the signal is obtained on the screen. The distance L on the graticule (the scale on the screen) corresponding to one complete cycle is measured (Figure 2.25). It is good practice to measure the length of, say, four cycles, and then divide by four so as to obtain an average value of L. The graticule will probably be divided into centimetre and perhaps millimetre or two-millimetre divisions. If the time-base setting is x (which will be in units of seconds, milliseconds or microseconds per centimetre), the time T for one cycle is given by $T = Lx$. The frequency f of the signal is then obtained from $f = 1/T$. The uncertainty of the determination will depend on how well you can estimate the measurement of the length of the cycle from the graticule. Remembering that the trace has a finite width, you can probably measure this length to an uncertainty of about ±2mm.

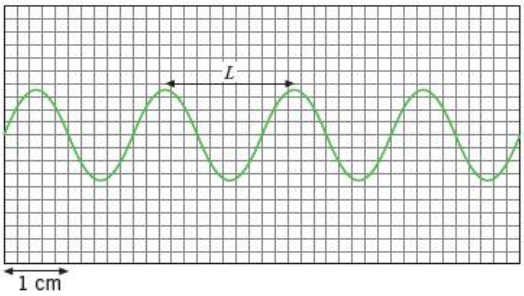


Figure 2.25 The use of a c.r.o. to measure frequency

As with most instruments you will use, your laboratory time will be so limited that you will probably have to take the time-base settings on trust. However, it is worth thinking about possible methods of checking the calibration. You could try checking against a calibrated signal generator: but who is to say which of the signal generator or c.r.o. has the correct calibration? Another method would be to connect a microphone to the Y-input, and sound a tuning fork of known frequency near the microphone.

Example

The output of a signal generator is connected to the Y-input of a c.r.o. When the time-base control is set at 0.50 milliseconds per centimetre, the trace shown in Figure 2.26 is obtained. What is the frequency of the signal?

Two complete cycles of the trace occupy 6.0 cm on the graticule. The length of one cycle is therefore 3.0 cm. The time-base setting is 0.50 ms cm^{-1}, so 3.0 cm is equivalent to $3.0 \times 0.50 = 1.5$ ms. The frequency is thus $1/1.5 \times 10^{-3} =$ **670 Hz**.

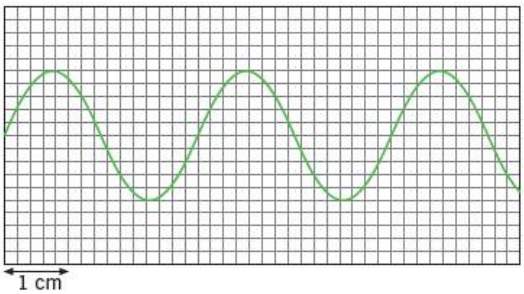


Figure 2.26

www.iranino.co

IRANINOVIN

Now it's your turn

3 The same signal is applied to the Y-input of the c.r.o. as in the example on page 25, but the time-base control is changed to 2.0 milliseconds per centimetre. How many complete cycles of the trace will appear on the screen, which is 8.0 cm wide?

Methods of measuring temperature

The SI unit of temperature, the kelvin (K), is based on the ideal gas (or thermodynamic) scale of temperature. The scale may be arrived at using an instrument called a constant-volume gas thermometer (see pages 212–3). The equation relating the Celsius temperature scale to the thermodynamic scale is

$$\theta = T - 273.15$$

where θ is in degrees Celsius and T is in kelvin.

Fortunately, in your practical course you will come across nothing more complicated than a liquid-in-glass (probably a mercury-in-glass) thermometer. You may, however, do experiments on thermocouple thermometers and resistance thermometers, and assess their suitability for use as a thermometer.

The mercury-in-glass thermometer

Liquid-in-glass thermometers are based on the thermal expansion of a liquid. A quantity of liquid is contained in a bulb at the end of a thin capillary tube. The space above the liquid contains an inert gas at low pressure. If the bulb is placed in a beaker of water which is gradually heated, the liquid expands and the thread of liquid occupies more and more of the capillary tube. The capillary tube is graduated: the position of the end of the thread gives the temperature.

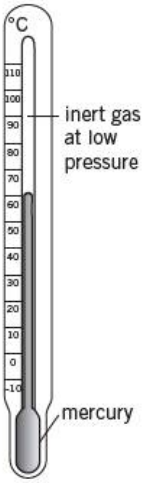


Figure 2.27 Mercury-in-glass thermometer

Most thermal physics experiments which you will carry out will involve the measurement of temperatures between 0°C (the temperature of melting ice) and 100°C (the temperature of steam above boiling water at a pressure of 1 atmosphere). The most useful thermometer covering this range is a mercury-in-glass thermometer (Figure 2.27) with graduations from –10°C to 110°C, in 1°C intervals. You will find it easy to take readings to the nearest half degree, and perhaps to 0.2°C. There are a number of precautions you should take when using the thermometer. Always allow time for the thermometer to reach thermal equilibrium with its surroundings. If you are measuring the temperature of a beaker of liquid which is being heated, the liquid must be thoroughly stirred before taking the reading. (Because of convection currents, there is a temperature difference of several degrees between the top and the bottom of the liquid.) The thermometer is calibrated for use at a standard depth of immersion; this may be stated on the stem. If it is not, try to ensure that the thermometer is always immersed in the liquid to the same depth. The length of the bulb plus about 20 mm is a reasonable guide.

There are some points to be made about safety. Thermometers are relatively fragile instruments. Because of their shape, they have a tendency to roll along the bench-top. Make sure that your thermometer does not roll off and fall to the ground. If a thermometer does break and the mercury in it comes out, do not be tempted to play with it. Mercury is a poison. To reduce the risk of breakage, do not use the thermometer as a stirrer, unless it is of a robust type designated as a 'stirring thermometer'. If you have to fit a thermometer through a rubber bung, make sure that the hole in the bung is large enough, and lubricate the rubber thoroughly with soap. Wear gloves and grip the thermometer so that, if it breaks, your wrist will not be cut.

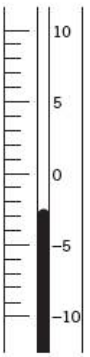


Figure 2.28

Example

The temperature of a mixture of ice, salt and water is measured using a mercury-in-glass thermometer. When thermal equilibrium has been reached, the mercury thread in the thermometer is as shown in Figure 2.28. What is the temperature of the mixture? What is the uncertainty in this value?

By interpolation between the scale divisions, the temperature reading is **–2.5 °C**. The uncertainty is probably about **±0.5 °C**.

www.iranino.co

IRANINOVIN

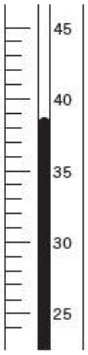


Figure 2.29



Figure 2.30 Thermocouple thermometer

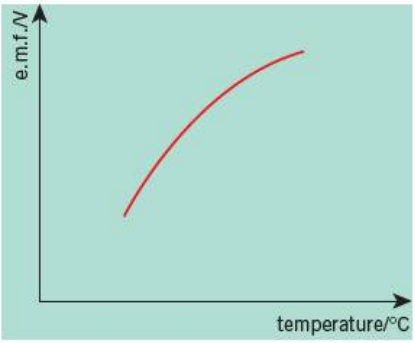


Figure 2.31 Calibration curve for a thermocouple

Now it's your turn

4 The temperature of a solidifying liquid is measured using a liquid-in-glass thermometer. When thermal equilibrium has been reached, the liquid thread in the thermometer is as shown in Figure 2.29. What is the solidification temperature? What is the uncertainty in this value?

The thermocouple thermometer

A thermocouple thermometer consists of two wires made of different metals or alloys, joined at one end. The other ends of the wires are connected to the terminals of a millivoltmeter. This may be a digital instrument, which is calibrated in °C (see Figure 2.30). The thermocouple may also be connected to a datalogger. The junction is placed in thermal contact with the object, the temperature of which is required.

The thermocouple thermometer actually measures the difference in temperature between the junction of the two metals (the hot junction) and a cold junction. In some applications, the cold junction is placed in an ice–water mixture, so as to achieve a known reference (see Figure 11.9, page 216).

The thermocouple may be connected to a millivoltmeter which has not been calibrated in temperature units. In this case, you will have to make use of the known variation of thermoelectric e.m.f. for that particular pair of metals with temperature. You will need to draw a **calibration curve**, a graph of e.m.f., against temperature so that you can read off the temperature corresponding to an e.m.f. reading. Note that this graph is often a curve rather than a straight line, as shown in Figure 2.31. In some cases the curvature is so much that the same e.m.f. can be obtained for two different temperatures. Clearly, this restricts the temperature range over which the thermocouple is useful.

Choice of method

The heat capacity of the bulb of a liquid-in-glass thermometer is much greater than that of the hot junction of a thermocouple. For this reason, the thermocouple is particularly useful when a rapidly varying temperature is to be measured, or when the object, the temperature of which is required, has a small heat capacity.

Mercury-in-glass thermometers are available to cover the temperature range from about −40 °C to 350 °C. Thermocouples using different pairs of metal or alloy wires can cover a much larger range.

The choice of a particular thermometer in a given application will depend on the range of temperatures to be covered, the heat capacity of the object, and whether the temperature is varying rapidly.

Methods of measuring current and potential difference

Your physics laboratory will probably have a selection of instruments for measuring current and potential difference (voltage). The two main types are analogue meters, in which a pointer moves over a scale (Figure 2.32), and digital, in which the value is displayed on a read-out consisting of a series of integers (Figure 2.33).

Analogue meters

The normal analogue meter is restricted to the measurement of the relevant quantity over a single range. For example, a 0–1 A d.c. ammeter will measure direct currents in the range from zero to 1 A. A 0–30 V d.c. voltmeter will measure steady potential differences in the range from zero to 30 V. Some analogue meters have a dual-range facility, with a common negative terminal and two positive terminals, each of which is associated with a separate scale on the instrument. Thus, one scale might be 0–3 A, and the other 0–10 A. Each of the positive terminals is marked with the scale to which it refers. Be careful to take the reading on the scale corresponding to the pair of terminals you have selected.

www.iranino.co

IRANINOVIN

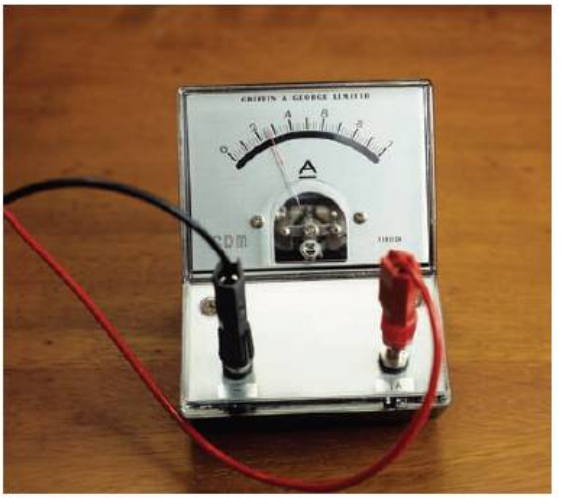


Figure 2.32 Analogue meter



Figure 2.33 Digital meter

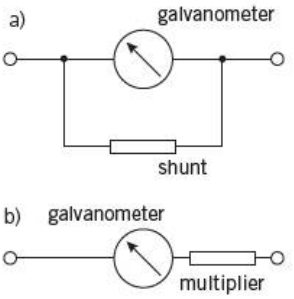


Figure 2.34 a) Galvanometer with shunt, for current measurements b) galvanometer with multiplier, for voltage measurements

Analogue meters are subject to zero error. Before switching on the circuit, check whether the needle is exactly at the zero mark. If it is not, return the needle to zero by adjusting the screw at the needle pivot. There is also the possibility of parallax error. The needle should be read from a position directly above it and the scale, and not from one side. Sometimes a strip of mirror is provided close to the scale so that the experimenter can align the needle with its image in the mirror, ensuring that viewing is vertical. The uncertainty associated with a current or voltage reading from an analogue meter is usually taken to be (±) the smallest scale reading.

A **galvanometer** is a sensitive current-measuring analogue meter. It may be converted into an ammeter by the connection of a suitable resistor in parallel with the meter (Figure 2.34a). Such a resistor is called a *shunt*. The meter may be converted into a voltmeter by the connection of a suitable resistor in series with the meter (Figure 2.34b). Such a resistor is called a *multiplier*.

The manufacturers provide shunts and multipliers which are clearly labelled with the conversion function and full-scale deflection, for attachment to the basic galvanometer (Figure 2.35). All you need do is to select the shunt or multiplier required for your experiment and make sure that you apply the correct factor when reading the scale.

A galvanometer with a centre-zero scale shows negative currents when the needle is to the left-hand side of the zero mark and positive currents when it is to the right. This type of meter is often used as a null indicator; that is, to detect when the current in a part of a circuit is zero.



Figure 2.35 Shunts and multipliers for use with a galvanometer

www.iraninno.co

IRANINOVIN



Figure 2.36 Multimeter

Digital meters

Digital meters may have a zero error. Before switching on the circuit, check whether the reading is zero. If it is not zero, make a note of the reading and take it into account when reading the current or voltage. The use of a digital meter may save you the trouble of selecting an instrument with a suitable range for your application. Most have an auto-ranging function; that is, the instrument selects the most sensitive range for the particular value of current or voltage being measured. All the experimenter has to do is check whether there is a zero error and make an adjustment if necessary, note whether the display indicates 'A' or 'mA', and observe the position of the decimal point!

The uncertainty in the reading of a digital meter is expressed in terms of the overall uncertainty and the uncertainty in the last digit. When in use, you will note that the last digit of the display fluctuates from one figure to another. You can try to estimate the mean of the fluctuations, but if this fluctuation occurs, there is clearly uncertainty in the last digit of the value.

Multimeters

Multimeters, or multifunction instruments, are available in both analogue and digital forms (Figure 2.36). Such meters may include switched options for the measurement of direct and alternating currents and voltages, and of resistance, with several ranges for each quantity being measured. If you use a multimeter, make sure that you are familiar with the controls, so that you can set the instrument to measure the quantity you require.

Choice of method

Much will depend on the selection of meters available in your laboratory. Before you set up your circuit, make a rough calculation to determine the ranges of currents and voltages that you will have to measure. This is a vital part of the planning process, and will help to make sure that you select the appropriate instrument from those that are available.

In some laboratories, multimeters are provided for use primarily as test instruments, to be available to anyone who wishes to make a rapid check on currents, potential differences or resistances in a circuit. If this is the rule in your laboratory, it is bad practice to use a multimeter in a long experiment, when a single-function and single-range instrument would do the job equally successfully.

Remember that, to measure a current in a component in a circuit, an ammeter should be connected in series with the component. To measure the potential difference across the component, a voltmeter should be connected in parallel with the component. The arrangement is shown in Figure 2.37.

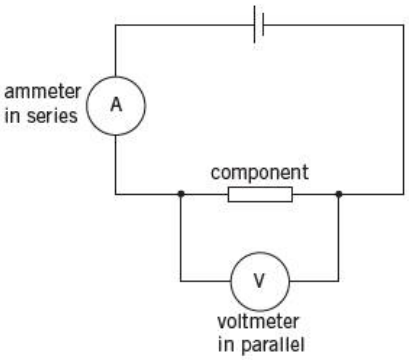


Figure 2.37 An ammeter is connected in series with the component, a voltmeter in parallel

www.iranino.co

IRANINOVIN

Application: measurement of voltage using a cathode-ray oscilloscope

Note, knowledge of this measurement technique is only required in the A Level course but is included here for completeness. The cathode-ray oscilloscope, with its calibrated Y-amplifier, may be used to measure the amplitude of an alternating voltage signal. (We have already seen how the time-base of the c.r.o. may be used to measure time.) The signal is connected to the Y-input, and the Y-amplifier and time-base settings are adjusted until a suitable trace is obtained (Figure 2.38). The amplitude A of the trace is measured. If the Y-amplifier setting is Q (in units of volts per centimetre), the peak value V_0 of the signal is given by $V_0 = AQ$. The peak-to-peak value is $2V_0$, and the r.m.s. (root-mean-square) voltage is $V_0/\sqrt{2}$. (Remember that the reading obtained on an analogue or digital voltmeter is the r.m.s. value.)

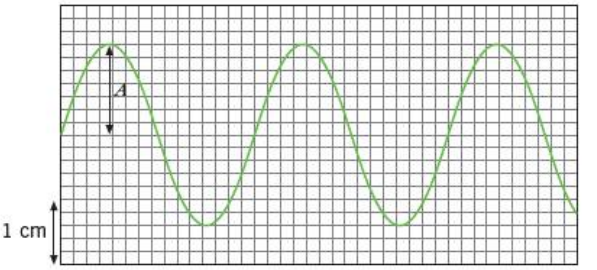


Figure 2.38 Measurement of alternating voltage

Example

The output from a signal generator is connected to the Y-input of a c.r.o. When the Y-amplifier control is set to 5.0 millivolts per centimetre, the trace shown in Figure 2.39 is obtained. Find the peak voltage of the signal, and the r.m.s. voltage.

Measure the amplitude of the trace on the graticule: this is 1.4 cm. The Y-amplfier setting is 5.0 mV cm^{-1}. 1.4 cm is thus equivalent to $1.4 \times 5.0 =$ **7.0 mV**. The peak voltage of the signal is 7.0 mV. The r.m.s. voltage is given by $7.0/\sqrt{2} =$ **4.9 mV r.m.s**.

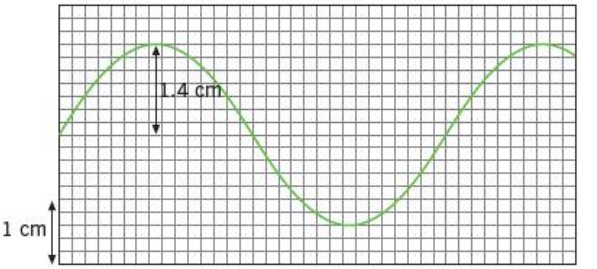


Figure 2.39

Now it's your turn

5 The output from a signal generator is connected to the Y-input of a c.r.o. When the Y-amplifier control is set to 20 millivolts per centimetre, the trace shown in Figure 2.40 is obtained. Find:
   (a) the peak-to-peak voltage of the signal,
   (b) the r.m.s. voltage.

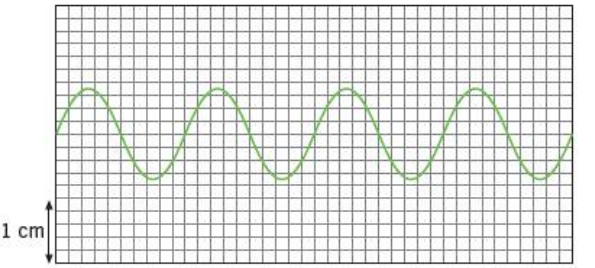


Figure 2.40

www.iranino.co

IRANINOVIN

Measuring magnetic flux density

The flux density of a magnetic field may be measured using a **Hall probe**, a device which makes use of the Hall effect (see page 322). The Hall probe apparatus used in school or college laboratories consists of a thin slice of a semiconductor material which is placed with its plane at right angles to the direction of the magnetic field. The control unit is arranged to pass a certain current through the semiconductor slice; the Hall voltage, which is proportional to the magnetic flux density, is read off on an analogue or digital meter, which is already calibrated in units of magnetic flux density (tesla). The arrangement is illustrated in Figure 2.41.

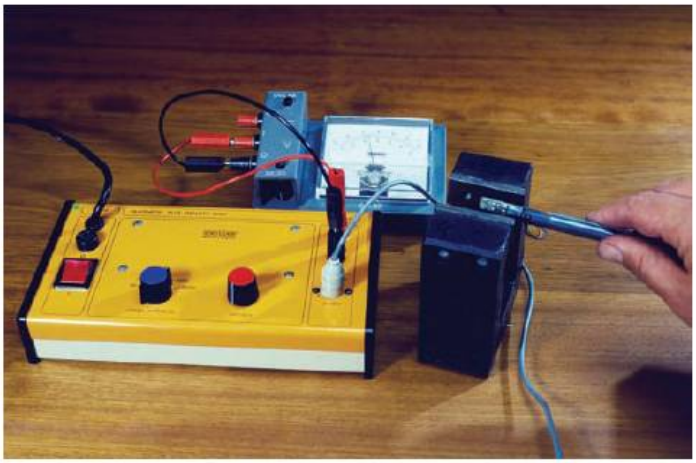


Figure 2.41 Hall probe apparatus

The use of the Hall probe to measure magnetic flux density is only required for the A level syllabus but is included here for completeness.

2.2 Errors and uncertainties

Accuracy and precision

Accuracy is the degree to which a measurement approaches the 'true value'.

Accuracy depends on the equipment used, the skill of the experimenter and the techniques used. Reducing systematic error or uncertainty (see page 35) in a measurement improves its accuracy.

Precision is the determined by the size of the random error (see page 35) in the measurements.

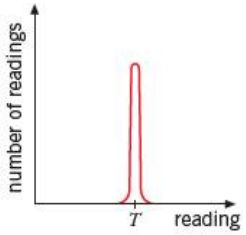
Precision is that part of accuracy which is within the control of the experimenter. The experimenter may choose different measuring instruments and may use them with different levels of skill, thus affecting the precision of measurement.

If we want to measure the diameter of a steel sphere or a marble, we could use a metre rule, or a vernier caliper, or a micrometer screw gauge. The choice of measuring instrument would depend on the precision with which we want the measurement to be made. For example, the metre rule could be used to measure to the nearest millimetre, the vernier caliper to the nearest tenth of a millimetre, and the micrometer screw gauge to the nearest one-hundredth of a millimetre. We could show the readings as follows:

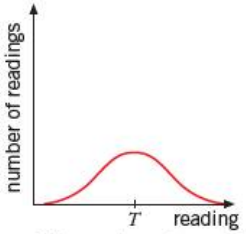
metre rule:	1.2 ± 0.1 cm
vernier caliper:	1.21 ± 0.01 cm
micrometer screw gauge:	1.212 ± 0.001 cm

www.iranino.co

IRANINOVIN

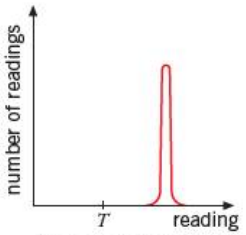


a) precise and accurate

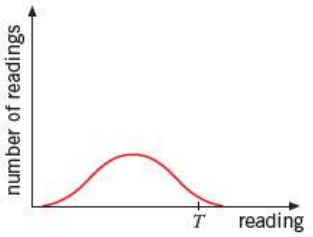


b) imprecise but accurate

Figure 2.42



a) precise but not accurate



b) imprecise and not accurate

Figure 2.43

The degree of precision to which the measurement is made increases as we move from the metre rule to the vernier caliper and finally to the micrometer screw gauge. Note that the number of significant figures quoted for the measurement increases as the precision increases. In fact, the number of significant figures in a measurement gives an indication of the precision of the measurement.

When a measurement is repeated many times with a precise instrument, the readings are all close together, as shown in Figure 2.42a. Using a measuring instrument with less precision means that there would be a greater spread of readings, as shown in Figure 2.42b, resulting in greater uncertainty.

Precision is partly to do with the accuracy of an observation or measurement. A reading may be very precise but it need not be accurate. Accuracy is concerned with how close a reading is to its true value. For example, a micrometer screw gauge may be precise to ±0.001 cm but, if there is a large zero error, then the readings from the scale for the diameter of a sphere or marble would not be accurate. The distinction between precision and accuracy is illustrated in Figure 2.43. On each of the graphs the value T is the true value of the quantity.

Uncertainty

In the list on page 31, each of the measurements is shown with its precision. For example, using the metre rule, the measurement of the diameter is 1.2 cm with a precision of 0.1 cm. In reality, precision is not the only factor affecting the accuracy of the measurement.

The total range of values within which the measurement is likely to lie is known as its **uncertainty**.

For example, a measurement of 46.0 ± 0.5 cm implies that the most likely value is 46.0 cm, but it could be as low as 45.5 cm or as high as 46.5 cm. The uncertainty in the measurement is ±0.5 cm or ±(0.5/46) × 100% = ±1%.

It is important to understand that, when writing down measurements, the number of significant figures of the measurement indicates its uncertainty. Some examples of uncertainty are given in Table 2.2.

Table 2.2 Examples of uncertainty

Instrument	uncertainty	typical reading
stopwatch with 0.1 s divisions	±0.1 s	16.2 s
thermometer with 1 °C intervals	±0.5 °C	22.5 °C
ammeter with 0.1 A divisions	±0.1 A	2.1 A

Note that while a particular temperature is shown as a number with the unit °C, a temperature *interval* is correctly shown as a number with the unit deg C. However, most people use the unit °C for both a particular temperature and a temperature interval.

It should be remembered that the uncertainty in a reading is not wholly confined to the reading of its scale or to the skill of the experimenter. Any measuring instrument has a built-in uncertainty. For example, a metal metre rule expands as its temperature rises. At only one temperature will readings of the scale be precise. At all other temperatures, there will be an uncertainty due to the expansion of the scale. Knowing by how much the rule expands would enable this uncertainty to be removed and hence improve precision.

Manufacturers of digital meters quote the uncertainty for each meter. For example, a digital voltmeter may be quoted as ±1% ±2 digits. The ±1% applies to the total reading shown on the scale and the ±2 digits is the uncertainty in the final display figure. This means that the uncertainty in a reading of 4.00 V would be (±4.00 × 1/100) ± 0.02 = ±0.06 V. This uncertainty would be added to any further uncertainty due to a fluctuating reading.

The uncertainty in a measurement is sometimes referred to as being its *error*. This is not strictly true. Error would imply that a mistake has been made. There is no mistake in taking the measurement, but there is always some doubt or some uncertainty as to its value.

www.iranino.co

IRANINOVIN

Example

A student takes a large number of imprecise readings for the current in a wire. He uses an ammeter with a zero error of $-\Delta I$, meaning that all scale readings are too small by ΔI. The true value of the current is I. Sketch a distribution curve of the number of readings plotted against the measured value of the current. Label any relevant values.

This is the case illustrated in Figure 2.43b. The peak of the curve is centred on a value of $I - \Delta I$.

Now it's your turn

6 A large number of precise readings for the diameter D of a wire is made using a micrometer screw gauge. The gauge has a zero error $+E$, which means that all readings are too large. Sketch a distribution curve of the number of readings plotted against the measured value of the diameter.

7 The manufacturer of a digital ammeter quotes its uncertainty as ±1.5% ±2 digits.
   (a) Determine the uncertainty in a constant reading of 2.64 A.
   (b) The meter is used to measure the current from a d.c. power supply. The current is found to fluctuate randomly between 1.98 A and 2.04 A. Determine the most likely value of the current, with its uncertainty.

Choice of instruments

The precision of an instrument required for a particular measurement is related to the measurement being made. Obviously, if the diameter of a hair is being measured, a high-precision micrometer screw gauge is required, rather than a metre rule. Similarly, a galvanometer should be used to measure currents of the order of a few milliamperes, rather than an ammeter. Choice is often fairly obvious where single measurements are being made, but care has to be taken where two readings are subtracted. Consider the following example.

The distance of a lens from a fixed point is measured using a metre rule. The distance is 95.2 cm (see Figure 2.44). The lens is now moved closer to the fixed point and the new distance is 93.7 cm. How far has the lens moved? The answer is obvious: (95.2 – 93.7) = 1.5 cm. But how precise is the measurement?

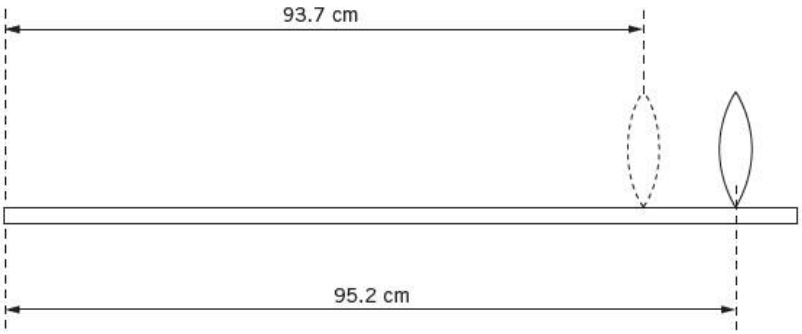


Figure 2.44

We have seen that the uncertainty in each measurement using a metre rule is, optimistically, ±1 mm ($\frac{1}{2}$mm at the zero end of the rule plus $\frac{1}{2}$mm when finding the position of the centre of the lens). This means that each separate measurement of length has an uncertainty of about (1/940 × 100)%, i.e. about 0.1%. That appears to be good! However, the uncertainty in the distance moved is ±2 mm, because both distances have an uncertainty, and these uncertainties add up, so the uncertainty is ±(2/15 × 100)% = ±13%. This uncertainty is, quite clearly, unacceptable. Another means by which the distance moved could be measured must be devised to reduce the uncertainty.

During your A/AS course, you will meet with many different measuring instruments. You must learn to recognise which instrument is most appropriate for a particular

www.iranino.co

IRANINOVIN

measurement. A stopwatch may be suitable for measuring the period of oscillation of a pendulum but you would have difficulty using it to find the time taken for a stone to fall vertically from rest through a distance of 1 m. Choice of appropriate instruments is likely to be examined when you are planning experiments.

Example

Suggest appropriate instruments for the measurement of the dimensions of a single page of this book.

The obvious instrument to measure the height and width of a page is a 30 cm ruler, which can be read to ±1 mm. The width, the smaller dimension, is about 210 mm, so the actual uncertainty is 210 mm ± 1 mm and the percentage uncertainty is about ±0.5%. It is not sensible to try to measure the thickness of a single page, even with a micrometer screw gauge, as the percentage uncertainty will be very high. Instead, use the screw gauge to measure the thickness of a large number of pages (but don't include the covers!). Four hundred pages are about 18 mm thick. The uncertainty in this measurement, using a screw gauge, is ±0.01 mm, giving a percentage uncertainty of about ±0.05% in the thickness of all 400 pages. This is also the percentage uncertainty in the thickness of a single page. If an uncertainty of ±0.5% is acceptable, a vernier caliper should be used instead of the screw gauge.

Now it's your turn

8 Suggest appropriate instruments for the measurement of:
- (a) the discharge current of a capacitor (of the order of 10^{-6} A),
- (b) the time for a feather to fall in air through a distance of about 40 cm,
- (c) the time for a ball to fall vertically through a distance of about 40 cm,
- (d) the length of a pendulum having a period of about 1 s,
- (e) the temperature of some water as it cools to room temperature,
- (f) the temperature of a roaring Bunsen flame,
- (g) the weight of 20 small glass beads,
- (h) the weight of a house brick.

9 The diameter of a ball is measured using a metre rule and a two set-squares, as illustrated in Figure 2.45. The readings on the rule are 16.8 cm and 20.4 cm. Each reading has an uncertainty of ±1 mm. Calculate, for the diameter of the ball:
- (a) its actual uncertainty,
- (b) its percentage uncertainty.

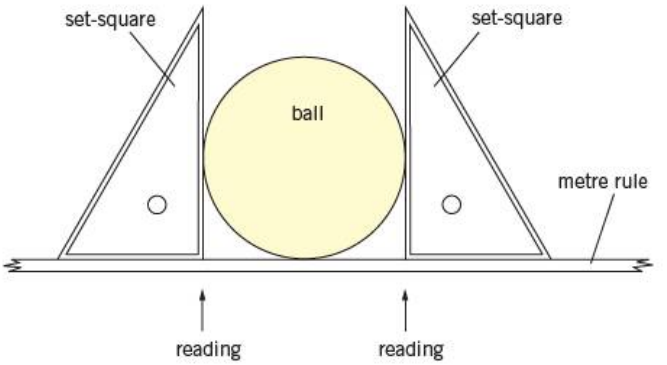


Figure 2.45

Suggest an alternative, but more precise, method by which the diameter could be measured.

Systematic and random uncertainty (error)

Not only is the choice of instrument important, so that any measurement is made with acceptable precision, but also the techniques of measurement must optimise accuracy. That is, your experimental technique must reduce as far as possible any uncertainties in readings. These uncertainties may be classed as either **systematic** or **random**.

www.iranino.co

IRANINOVIN

Systematic uncertainty (error)

A systematic uncertainty will result in all readings being either above or below the accepted value. This uncertainty cannot be eliminated by repeating readings and then averaging. Instead systematic uncertainty can be reduced only by improving experimental techniques. Examples of systematic uncertainty are:

- **zero error on an instrument**
  The scale reading is not zero before measurements are taken – see Figure 2.46. Check before starting the experiment.
- **wrongly calibrated scale**
  In school laboratories we assume that measuring devices are correctly calibrated, and would not be expected to check the calibration in an experiment. However, if you have doubts, you can check the calibration of an ammeter by connecting several in series in the circuit, or of a voltmeter by connecting several in parallel. Rules can be checked by laying several of them alongside each other. Thermometers can be checked by placing several in well-stirred water. These checks will not enable you to say which of the instruments are calibrated correctly, but they will show you if there is a discrepancy.
- **reaction time of experimenter**
  When timings are carried out manually, it must be accepted that there will be a delay between the experimenter observing the event and starting the timing device. This delay, called the reaction time, may be as much as a few tenths of a second. To reduce the effect, you should arrange that the intervals you are timing are much greater than the reaction time. For example, you should time sufficient swings of a pendulum for the total time to be of the order of at least ten seconds, so that a reaction time of a few tenths of a second is less important.

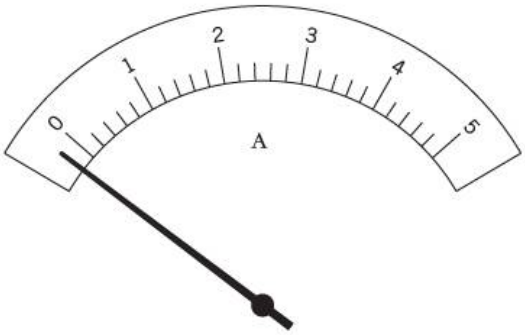


Figure 2.46 This ammeter has a zero error of about –0.2 A.

Random uncertainty (error)

Random uncertainty results in readings being scattered around the accepted value. Random uncertainty may be reduced by repeating a reading and averaging, and by plotting a graph and drawing a best-fit line. Examples of random errors are:

- reading a scale, particularly if this involves the experimenter's judgement about interpolation between scale readings
- timing oscillations without the use of a reference marker, so that timings may not always be made to the same point of the swing
- taking readings of a quantity that varies with time, involving the difficulty of reading both a timer scale and another meter simultaneously
- reading a scale from different angles introduces a variable parallax error. (In contrast, if a scale reading is always made from the same non-normal angle, this will introduce a systematic error.)

www.iranino.co

IRANINOVIN

Example

The current in a resistor is to be measured using an analogue ammeter. State one source of **(a)** a systematic uncertainty, **(b)** a random uncertainty. In both cases, suggest how the uncertainty may be reduced.

(a) Systematic uncertainty could be a zero error on the meter, or a wrongly calibrated scale. This can be reduced by checking for a zero reading before starting the experiment, or using two ammeters in series to check that the readings agree.

(b) Random uncertainty could be a parallax error caused by taking readings from different angles. This can be reduced by the use of a mirror behind the scale and viewing normally.

Now it's your turn

10 The length of a pencil is measured with a 30 cm rule. Suggest one possible source of **(a)** a systematic uncertainty, **(b)** a random uncertainty. In each case, suggest how the uncertainty may be reduced.

11 The diameter of a wire is to be measured to a precision of ±0.01 mm.
   (a) Name a suitable instrument.
   (b) Suggest a source of systematic uncertainty.
   (c) Explain why it is good practice to average a set of diameter readings, taken spirally along the length of the wire.

Combining uncertainties

There are two simple rules for obtaining an estimate of the overall uncertainty in a final result. The rules are:

1 For quantities which are added or subtracted to give a final result, add the actual uncertainties.
2 For quantities which are multiplied together or divided to give a final result, add the fractional uncertainties.

Suppose that we wish to obtain the value of a physical quantity x by measuring two other quantities, y and z. The relation between x, y and z is known, and is

$$x = y + z$$

If the uncertainties in y and z are Δy and Δz respectively, the uncertainty Δx in x is given by

$$\Delta x = \Delta y + \Delta z$$

If the quantity x is given by

$$x = y - z$$

the uncertainty in x is again given by

$$\Delta x = \Delta y + \Delta z$$

Examples

1 I_1 and I_2 are two currents coming into a junction in a circuit. The current I going out of the junction is given by

$$I = I_1 + I_2$$

In an experiment, the values of I_1 and I_2 are determined as 2.0 ± 0.1 A and 1.5 ± 0.2 A respectively. What is the value of I? What is the uncertainty in this value?

Using the given equation, the value of I is given by $I = 2.0 + 1.5 = \mathbf{3.5\,A}$. The rule for combining the uncertainties gives $\Delta I = 0.1 + 0.2 = 0.3$ A. The result for I is thus **(3.5 ± 0.3) A**.

2 In an experiment, a liquid is heated electrically, causing the temperature to change from 20.0 ± 0.2 °C to 21.5 ± 0.5 °C. Find the change of temperature, with its associated uncertainty.

www.iranino.co

IRANINOVIN

The change of temperature is 21.5 – 20.0 = **1.5 °C**. The rule for combining the uncertainties gives the uncertainty in the temperature change as 0.2 + 0.5 = 0.7 °C. The result for the temperature change is thus **(1.5 ± 0.7) °C**.

Note that this second example shows that a small difference between two quantities may have a large uncertainty, even if the uncertainty in measuring each of the quantities is small. This is an important factor in considering the design of experiments, where the difference between two quantities may introduce an unacceptably large error.

Now it's your turn

12 Two set-squares and a ruler are used to measure the diameter of a cylinder. The cylinder is placed between the set-squares, and the set-squares are aligned with the ruler, in the manner of the jaws of calipers. The readings on the ruler at opposite ends of a diameter are 4.15 cm and 2.95 cm. Each reading has an uncertainty of ±0.05 cm.

(a) What is the diameter of the cylinder?

(b) What is the uncertainty in the diameter?

Now suppose that we wish to find the uncertainty in a quantity x, whose relation to two measured quantities, y and z, is

$$x = Ayz$$

where A is a constant. The uncertainty in the measurement of y is $\pm\Delta y$, and that in z is $\pm\Delta z$. The **fractional uncertainty** in x is given by

$$\frac{\Delta x}{x} = \frac{\Delta y}{y} + \frac{\Delta z}{z}$$

To combine the uncertainties when the quantities are raised to a power, for example

$$x = Ay^a z^b$$

where A is a constant, the rule is

$$\frac{\Delta x}{x} = a\left(\frac{\Delta y}{y}\right) + b\left(\frac{\Delta z}{z}\right)$$

Example

A value of the acceleration of free fall g was determined by measuring the period of oscillation T of a simple pendulum of length l. The relation between g, T and l is

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

In the experiment, l was measured as 0.55 ± 0.02 m and T was measured as 1.50 ± 0.02 s. Find the value of g, and the uncertainty in this value.

Substituting in the equation, $g = 4\pi^2(0.55/1.50^2) = \mathbf{9.7\ m\ s^{-2}}$. The fractional uncertainties are $\Delta l/l = 0.02/0.55 = 0.036$ and $\Delta T/T = 0.02/1.50 = 0.013$.

Applying the rule to find the fractional uncertainty in g

$$\frac{\Delta g}{g} = \frac{\Delta l}{l} + \frac{2\Delta l}{T} = 0.036 + 2 \times 0.013 = 0.062$$

The actual uncertainty in g is given by (value of g) × (fractional uncertainty in g) = 9.7 × 0.062 = 0.60 m s^{-2}. The experimental value of g, with its uncertainty, is thus **(9.7 ± 0.6) m s^{-2}**.

Note that it is not good practice to determine g from the measurement of the period of a pendulum of fixed length. It would be much better to take values of T for a number of different lengths l, and to draw a graph of T^2 against l. The gradient of this graph is $4\pi^2/g$.

Now it's your turn

13 A value of the volume V of a cylinder is determined by measuring the radius r and the length L. The relation between V, r and L is

$$V = \pi r^2 L$$

In an experiment, r was measured as 3.30 ± 0.05 cm, and L was measured as 25.4 ± 0.4 cm. Find the value of V, and the uncertainty in this value.

www.iranino.co

IRANINOVIN

If you find it difficult to deal with the fractional uncertainty rule, you can easily estimate the uncertainty by substituting extreme values into the equation.
For $x = Ay^a z^b$, taking account of the uncertainties in y and z, the lowest value of x is given by

$x_{low} = A(y - \Delta y)^a(z - \Delta z)^b$

and the highest by

$x_{high} = A(y + \Delta y)^a(z + \Delta z)^b$

If x_{low} and x_{high} are worked out, the uncertainty in the value of x is given by $(x_{high} - x_{low})/2$.

Example

Apply the extreme value method to the data for the simple pendulum experiment in the Example on page 37.

Because of the form of the equation for g, the lowest value for g will be obtained if the lowest value of l and the highest value for T are substituted. This gives

$$g_{low} = 4\pi^2(0.53/1.52^2) = 9.1\,\text{m s}^{-2}$$

The highest value for g is obtained by substituting the highest value for l and the lowest value for T. This gives

$$g_{high} = 4\pi^2(0.57/1.48^2) = 10.3\,\text{m s}^{-2}$$

The uncertainty in the value of g is thus $(g_{high} - g_{low})/2 = (10.3 - 9.1)/2 = \mathbf{0.6\,m\,s^{-2}}$, as before.

Now it's your turn

14 Apply the extreme value method to the data for the volume of the cylinder, on page 37.

If the expression for the quantity under consideration involves combinations of products (or quotients) and sums (or differences), then the best approach is the extreme value method.

Summary

- Methods available for the measurement of length include:
  metre rule (range 1 m, reading uncertainty 1 mm)
  micrometer screw gauge (range 50 mm, reading uncertainty 0.01 mm)
  vernier caliper (range 100 mm, reading uncertainty 0.1 mm).
- Methods available for the measurement of mass include:
  top-pan balance
  spring balance
  lever balance.
- Methods available for the measurement of time include:
  stopclock (reading uncertainty 0.2 s)
  stopwatch (reading uncertainty 0.01 s)
  cathode-ray oscilloscope.
- Methods available for the measurement of temperature include:
  liquid-in-glass thermometer
  thermocouple thermometer.
- Methods available for the measurement of current and potential difference include:
  analogue meter
  digital meter
  multimeter
  cathode-ray oscilloscope.
- Methods available for the measurement of magnetic flux density include the Hall probe.
- Accuracy is concerned with how close a reading is to its true value.
- Precision is determined by the size of the random error and can be controlled by the experimenter.

www.iranino.co

IRANINOVIN

- Uncertainty indicates the range of values within which a measurement is likely to lie.
- A systematic uncertainty (or systematic error) is often due to instrumental causes, and results in all readings being above or below the true value. It cannot be eliminated by averaging.
- A random uncertainty (or random error) is due to the scatter of readings around the true value. It may be reduced by repeating a reading and averaging, or by plotting a graph and taking a best-fit line.
- Combining uncertainties:
  for expressions of the form $x = y + z$ or $x = y - z$, the overall uncertainty is $\Delta x = \Delta y + \Delta z$
  for expressions of the form $x = Ay^a z^b$, the overall fractional uncertainty is $\Delta x/x = a(\Delta y/y) + b(\Delta z/z)$

Examination style questions

1 You are asked to measure the internal diameter of a glass capillary tube (diameter about 2 mm). You are also to investigate the uniformity of the tube along its length. Suggest suitable methods.

2 The value of the acceleration of free fall varies slightly at different places on the Earth's surface. Discuss whether this means that

a a top-pan balance,
b a spring balance,
c a lever balance,

should be re-calibrated when they are moved to different locations.

If you needed to, how would you calibrate a balance?

3 The shutter on a particular camera has settings which allow it to be open for (nominally) 1 s, 0.5 s, 0.25 s, 0.125 s, 0.067 s, 0.033 s, 0.017 s, 0.008 s, 0.004 s, 0.002 s and 0.001 s. Suggest a method (or methods) of calibrating the exposure times over this range.

4 Explain the factors you would consider when deciding whether to use a liquid-in-glass or a thermocouple thermometer in particular experimental situations.

5 Summarise the advantages and disadvantages of analogue and digital ammeters.

6 Explain how to use a cathode-ray oscilloscope to measure the characteristics of the sinusoidal output from a signal generator.

7 A metal wire has a cross-section of diameter of approximately 0.8 mm.

a State what instrument should be used to measure the diameter of the wire. *[1]*

b State how the instrument in **a** is
i checked so as to avoid a systematic error in the measurements, *[1]*
ii used so as to reduce random errors. *[2]*

Cambridge International AS and A Level Physics, 9702/22 May/June 2010 Q 1

8 a State the most appropriate instrument, or instruments, for the measurement of the following.
i the diameter of a wire of diameter about 1 mm *[1]*
ii the resistance of a filament lamp *[1]*
iii the peak value of an alternating voltage *[1]*

b The mass of a cube of aluminium is found to be 580 g with an uncertainty in the measurement of 10 g. Each side of the cube has a length of (6.0 ± 0.1) cm.
Calculate the density of aluminium with its uncertainty. Express your answer to an appropriate number of significant figures. *[5]*

Cambridge International AS and A level Physics, 9702/21 May/June 2009 Q 1

9 A simple pendulum may be used to determine a value for the acceleration of free fall g. Measurements are made of the length L of the pendulum and the period T of oscillation. The values obtained, with their uncertainties, are as shown.

$T = (1.93 \pm 0.03)$ s

$L = (92 \pm 1)$ cm

a Calculate the percentage uncertainty in the measurement of
i the period T, *[1]*
ii the length L. *[1]*

b The relationship between T, L and g is given by

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

Using your answers in **a**, calculate the percentage uncertainty in the value of g. *[1]*

c The values of L and T are used to calculate a value of g as 9.751 m s^{-2}.
i By reference to the measurements of L and T, suggest why it would not be correct to quote the value of g as 9.751 m s^{-2}. *[1]*
ii Use your answer in **b** to determine the absolute uncertainty in g.
Hence state the value of g, with its uncertainty, to an appropriate number of significant figures. *[2]*

Cambridge International AS and A level Physics, 9702/22 Oct/Nov 2009 Q 1

www.iranino.co

IRANINOVIN

3 Kinematics

By the end of this topic, you will be able to:

3.1
- **(a)** define and use distance, displacement, speed, velocity and acceleration
- **(b)** use graphical methods to represent distance, displacement, speed, velocity and acceleration
- **(c)** determine displacement from the area under a velocity–time graph
- **(d)** determine velocity using the gradient of a displacement–time graph
- **(e)** determine acceleration using the gradient of a velocity–time graph
- **(f)** derive, from the definitions of velocity and acceleration, equations which represent uniformly accelerated motion in a straight line
- **(g)** solve problems using equations which represent uniformly accelerated motion in a straight line, including the motion of bodies falling in a uniform gravitational field without air resistance
- **(h)** describe an experiment to determine the acceleration of free fall using a falling body
- **(i)** describe and explain motion due to a uniform velocity in one direction and a uniform acceleration in a perpendicular direction

Starting points

- Kinematics is a description of how objects move.
- The motion of objects can be described in terms of quantities such as position, speed, velocity and acceleration.

3.1 Speed, displacement, velocity and acceleration

Average speed

When talking about motion, we shall discuss the way in which the position of a particle varies with time. Think about a particle moving along a straight line. In a certain time, the particle will cover a certain distance. The **average speed** of the particle is defined as the distance moved divided by the time taken. Written as a word equation, this is

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

The unit of speed is the metre per second (m s^{-1}).

One of the most fundamental of physical constants is the speed of light in a vacuum. It is important because it is used in the definition of the metre, and because, according to the theory of relativity, it defines an upper limit to attainable speeds. The range of speeds that you are likely to come across is enormous; some are summarised in Table 3.1.

It is important to recognise that speed has a meaning only if it is quoted relative to a fixed reference. In most cases, speeds are quoted relative to the surface of the Earth, which – although it is moving relative to the Solar System – is often taken to be fixed.

Table 3.1 Examples of speeds

	speed/m s^{-1}
light	3.0×10^8
electron around nucleus	2.2×10^6
Earth around Sun	3.0×10^4
jet airliner	2.5×10^2
typical car speed (80 km per hour)	22
sprinter	10
walking speed	1.5
snail	1×10^{-3}

www.iranino.co

IRANINOVIN

Thus, when we say that a bird can fly at a certain average speed, we are relating its speed to the Earth. However, a passenger on a ferry may see that a seagull, flying parallel to the boat, appears to be practically stationary. If this is the case, the seagull's speed relative to the boat is zero. However, if the speed of the boat through the water is 8 m s^{-1}, then the speed of the seagull relative to Earth is also 8 m s^{-1}. When talking about relative speeds we must also be careful about directions. It is easy if the motions are in the same direction, as in the example of the ferry and the seagull. The addition of velocity vectors is considered in Topic 1 (page 8).

Examples

1 The radius of the Earth is 6.4×10^6 m; one revolution about its axis takes 24 hours (8.6×10^4 s). Calculate the average speed of a point on the Equator relative to the centre of the Earth.

In 24 hours, the point on the equator completes one revolution and travels a distance of $2\pi \times$ the Earth's radius, that is $2\pi \times 6.4 \times 10^6 = 4.0 \times 10^7$ m.
The average speed is (distance moved)/(time taken),
or $4.0 \times 10^7 / 8.6 \times 10^4 =$ **4.7×10^2 m s^{-1}**.

2 How far does a cyclist travel in 11 minutes if his average speed is 22 km h^{-1}?

First convert the average speed in km h^{-1} to a value in m s^{-1}.
22 km (2.2×10^4 m) in 1 hour (3.6×10^3 s) is an average speed of 6.1 m s^{-1}. 11 minutes is 660 s. Since average speed is (distance moved)/(time taken), the distance moved is given by (average speed) × (time taken), or $6.1 \times 660 =$ **4000 m**.

Note the importance of working in consistent units: this is why the average speed and the time were converted to m s^{-1} and s respectively.

3 A train is travelling at a speed of 25 m s^{-1} along a straight track. A boy walks along the corridor in a carriage towards the rear of the train, at a speed of 1 m s^{-1} relative to the train. What is his speed relative to Earth?

In one second, the train travels 25 m forwards along the track. In the same time the boy moves 1 m towards the rear of the train, so he has moved 24 m along the track. His speed relative to Earth is thus $25 - 1 =$ **24 m s^{-1}**.

Now it's your turn

1 The speed of an electron in orbit about the nucleus of a hydrogen atom is 2.2×10^6 m s^{-1}. It takes 1.5×10^{-16} s for the electron to complete one orbit. Calculate the radius of the orbit.
2 The average speed of an airliner on a domestic flight is 220 m s^{-1}. How long will it take to fly between two airports on a flight path 700 km long?
3 Two cars are travelling in the same direction on a long, straight road. The one in front has an average speed of 25 m s^{-1} relative to Earth; the other's is 31 m s^{-1}, also relative to Earth. What is the speed of the second car relative to the first when it is overtaking?

Speed and velocity

In ordinary language, there is no difference between the terms *speed* and *velocity*. However, in physics there is an important distinction between the two. **Velocity** is used to represent a vector quantity: the magnitude of how fast a particle is moving *and* the direction in which it is moving. **Speed** does not have an associated direction. It is a scalar quantity (see Topic 1 page 7).

So far, we have talked about the total distance travelled by a body along its actual path. Like speed, distance is a scalar quantity, because we do not have to specify the direction in which the distance is travelled. However, in defining velocity we introduce a quantity called **displacement**. Displacement of a particle is its change of position. The displacement is the distance travelled in a straight line in a specified direction from the starting point to the finishing point. Consider a cyclist travelling 500 m due east along a straight road, and then turning round and coming back 300 m. The total distance travelled is 800 m, but the displacement is only 200 m due east, since the cyclist has ended up 200 m from the starting point.

www.iranino.co

IRANINOVIN

The **average velocity** is defined as the displacement divided by the time taken.

$$\text{average velocity} = \frac{\text{displacement}}{\text{time taken}}$$

Because distance and displacement are different quantities, the average speed of motion will sometimes be different from the magnitude of the average velocity. If the time taken for the cyclist's trip in the example above is 120 s, the average speed is $800/120 = 6.7\,\text{m s}^{-1}$, whereas the magnitude of the average velocity is $200/120 = 1.7\,\text{m s}^{-1}$. This may seem confusing, but the difficulty arises only when the motion involves a change of direction and we take an average value. If we are interested in describing the motion of a particle at a particular moment in time, the speed at that moment is the same as the magnitude of the velocity at that moment.

We now need to define average velocity more precisely, in terms of a mathematical equation, instead of our previous word equation. Suppose that at time t_1 a particle is at a point x_1 on the x-axis (Figure 3.1). At a later time t_2, the particle has moved to x_2. The displacement (the change in position) is $(x_2 - x_1)$, and the time taken is $(t_2 - t_1)$. The average velocity $\bar{v}$ is then



Figure 3.1

$$\bar{v} = \frac{x_2 - x_1}{t_2 - t_1}$$

The bar over v is the symbol meaning 'average'. As a shorthand, we can write $(x_2 - x_1)$ as Δx, where Δ (the Greek capital letter delta) means 'the change in'. Similarly, $t_2 - t_1$ is written as Δt. This gives us

$$\bar{v} = \frac{\Delta x}{\Delta t}$$

If x_2 were less than x_1, $(x_2 - x_1)$ and Δx would be negative. This would mean that the particle had moved to the left, instead of to the right as in Figure 3.1. The sign of the displacement gives the direction of particle motion. If Δx is negative, then the average velocity v is also negative. The sign of the velocity, as well as the sign of the displacement, indicates the direction of the particle's motion. This is because both displacement and velocity are vector quantities.

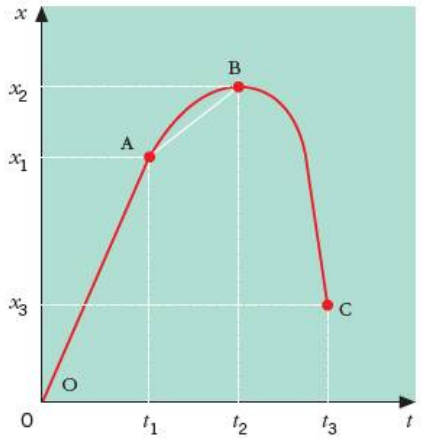


Figure 3.2

Describing motion by graphs

Position–time graphs

Figure 3.2 is a graph of position x against time t for a particle moving in a straight line. This curve gives a complete description of the motion of the particle. We can see from the graph that the particle starts at the origin O (at which $x = 0$) at time $t = 0$. From O to A the graph is a straight line: the particle is covering equal distances in equal periods of time. This represents a period of *uniform velocity*. The average velocity during this time is $(x_1 - 0)/(t_1 - 0)$. Clearly, this is the gradient of the straight-line part of the graph between O and A. Between A and B the particle is slowing down, because the distances travelled in equal periods of time are getting smaller. The average velocity during this period is $(x_2 - x_1)/(t_2 - t_1)$. On the graph, this is represented by the gradient of the straight line joining A and B. At B, for a moment, the particle is at rest, and after B it has reversed its direction and is heading back towards the origin. Between B and C the average velocity is $(x_3 - x_2)/(t_3 - t_2)$. Because x_3 is less than x_2, this is a negative quantity, indicating the reversal of direction.

Calculating the average velocity of the particle over the relatively long intervals t_1, $(t_2 - t_1)$ and $(t_3 - t_2)$ will not, however, give us the complete description of the motion. To describe the motion exactly, we need to know the particle's velocity at every instant. We introduce the idea of **instantaneous velocity**. To define instantaneous velocity we make the intervals of time over which we measure the average velocity shorter and shorter. This has the effect of approximating the curved displacement–time graph by a series of short straight-line segments.

www.iranino.co

IRANINOVIN

The approximation becomes better the shorter the time interval, as illustrated in Figure 3.3. Eventually, in the case of extremely small time intervals (mathematically we would say 'infinitesimally small'), the straight-line segment has the same direction as the tangent to the curve. This limiting case gives the instantaneous velocity as the gradient of the tangent to the displacement–time curve.

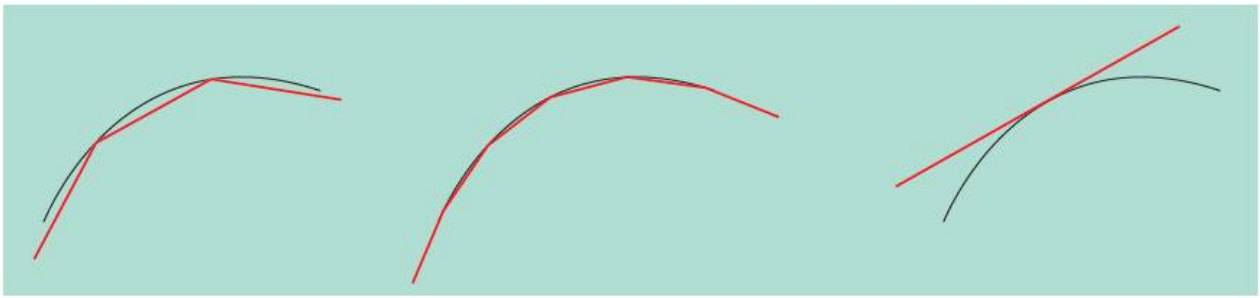


Figure 3.3

Displacement–time and velocity–time graphs

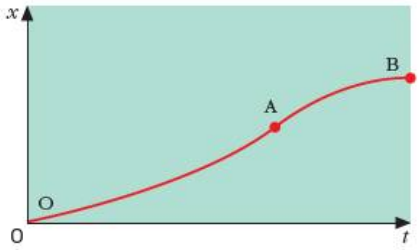


Figure 3.4

Figure 3.4 is a sketch graph showing how the displacement of a car, travelling along a straight test track, varies with time. We interpret this graph in a descriptive way by noting that between O and A the distances travelled in equal intervals of time are progressively increasing: that is, the velocity is increasing as the car is accelerating. Between A and B the distances for equal time intervals are decreasing; the car is slowing down. Finally, there is no change in position, even though time passes, so the car must be at rest. We can use Figure 3.4 to deduce the details of the way in which the car's instantaneous velocity v varies with time. To do this, we draw tangents to the curve in Figure 3.4 at regular intervals of time, and measure the slope of each tangent to obtain values of v. The plot of v against t gives the graph in Figure 3.5. This confirms our descriptive interpretation: the velocity increases from zero to a maximum value, and then decreases to zero again. We will look at this example in more detail on pages 47–48, where we shall see that the area under the velocity–time graph in Figure 3.5 gives the displacement x.

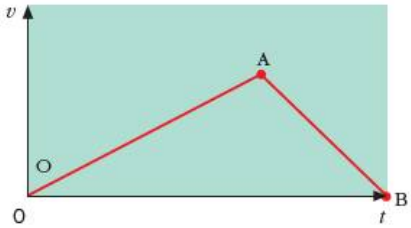


Figure 3.5

Acceleration

We have used the word *accelerating* in describing the increase in velocity of the car in the previous section. Acceleration is a measure of the rate at which the velocity of the particle is changing. **Average acceleration** is defined by the word equation

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

The unit of acceleration is the unit of velocity (the metre per second) divided by the unit of time (the second), giving the metre per (second)^2 which is represented as m s^{-2}. In symbols, this equation is

$$\bar{a} = \frac{v_2 - v_1}{t_2 - t_1} = \frac{\Delta v}{\Delta t}$$

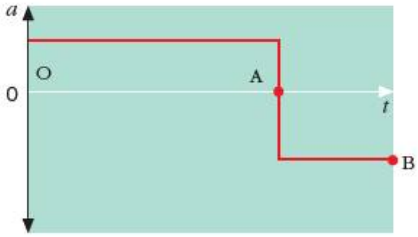


Figure 3.6

where v_1 and v_2 are the velocities at time t_1 and t_2 respectively. To obtain the **instantaneous acceleration**, we take extremely small time intervals, just as we did when defining instantaneous velocity. Because it involves a change in velocity (a vector quantity), acceleration is also a vector quantity: we need to specify both its magnitude and its direction.

We can deduce the acceleration of a particle from its velocity–time graph by drawing a tangent to the curve and finding the slope of the tangent. Figure 3.6 shows the result of doing this for the car's motion described by Figure 3.4 (the displacement–time graph) and Figure 3.5 (the velocity–time graph). The car accelerates at a constant rate between O and A, and then decelerates (that is, slows down) uniformly between A and B.

www.iranino.co

IRANINOVIN

Table 3.2 Examples of accelerations

	acceleration/m s^{-2}
due to circular motion of electron around nucleus	9×10^{26}
car crash	1×10^{3}
free fall on Earth	10
family car	2
free fall on Moon	2
at Equator, due to rotation of Earth	3×10^{-2}
due to circular motion of Earth around Sun	6×10^{-5}

An acceleration with a very familiar value is the acceleration of free fall near the Earth's surface (see page 45): this is 9.81 m s^{-2}, often approximated to 10 m s^{-2}. To illustrate the range of values you may come across, some accelerations are summarised in Table 3.2.

Examples

1 A sports car accelerates along a straight test track from rest to 70 km h^{-1} in 6.3 s. What is its average acceleration?

First convert the data into consistent units. 70 km (7.0×10^4 m) in 1 hour (3.6×10^3 s) is 19 m s^{-1}. Since average acceleration is (change of velocity)/(time taken), the acceleration is 19/6.3 = **3.0 m s^{-2}**.

2 A railway train, travelling along a straight track, takes 1.5 minutes to come to rest from a speed of 115 km h^{-1}. What is its average acceleration while braking?

115 km h^{-1} is 31.9 m s^{-1}, and 1.5 minutes is 90 s. The average acceleration is (change of velocity)/(time taken) = −31.9/90 = **−0.35 m s^{-2}**.

Note that the acceleration is a negative quantity because the change of velocity is negative: the final velocity is less than the initial. A negative acceleration is often called a deceleration.

Now it's your turn

4 A sprinter, starting from the blocks, reaches his full speed of 9.0 m s^{-1} in 1.5 s. What is his average acceleration?

5 A car is travelling at a speed of 25 m s^{-1}. At this speed, it is capable of accelerating at 1.8 m s^{-2}. How long would it take to accelerate from 25 m s^{-1} to the speed limit of 31 m s^{-1}?

6 At an average speed of 24 km h^{-1}, how many kilometres will a cyclist travel in 75 minutes?

7 An aircraft travels 1600 km in 2.5 hours. What is its average speed, in m s^{-1}?

8 Does a car speedometer register speed or velocity? Explain.

9 An aircraft travels 1400 km at a speed of 700 km h^{-1}, and then runs into a headwind that reduces its speed over the ground to 500 km h^{-1} for the next 800 km. What is the total time for the flight? What is the average speed of the aircraft?

10 A sports car can stop in 6.1 s from a speed of 110 km h^{-1}. What is its acceleration?

11 Can the velocity of a particle change if its speed is constant? Can the speed of a particle change if its velocity is constant? If the answer to either question is 'yes', give examples.

Uniformly accelerated motion

Having defined displacement, velocity and acceleration, we shall use the definitions to derive a series of equations, called the *kinematic equations*, which can be used to give a complete description of the motion of a particle in a straight line. The mathematics will be simplified if we deal with situations in which the acceleration does not vary with time; that is, the acceleration is uniform (or constant). This approximation applies for many practical cases. However, there are two important types of motion for which the kinematic equations do not apply: circular motion and the oscillatory motion called simple harmonic motion. We shall deal with these separately in Topic 7 and Topic 13.

Think about a particle moving along a straight line with constant acceleration a. Suppose that its initial velocity, at time $t = 0$, is u. After a further time t its velocity has increased to v. From the definition of acceleration as (change in velocity)/(time taken), we have $a = (v - u)/t$ or, re-arranging,

$$v = u + at$$

From the definition of average velocity $\bar{v}$ as (distance travelled)/(time taken), over the time t the distance travelled s will be given by the average velocity multiplied by the time taken, or

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

www.iranino.co

IRANINOVIN

The average velocity $\bar{v}$ is written in terms of the initial velocity u and final velocity v as

$$\bar{v} = \frac{u+v}{2}$$

and, using the previous equation for v,

$$\bar{v} = \frac{(u+u+at)}{2} = u + \frac{at}{2}$$

Substituting this we have

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

The right-hand side of this equation is the sum of two terms. The ut term is the distance the particle would have travelled in time t if it had been travelling with a constant speed u, and the $\frac{1}{2}at^2$ term is the additional distance travelled as a result of the acceleration.

The equation relating the final velocity v, the initial velocity u, the acceleration a and the distance travelled s is

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

If you wish to see how this is obtained from previous equations, see the Maths Note below.

Maths Note

From $v = u + at$,
$t = (v - u)/a$
Substitute this in $s = ut + \frac{1}{2}at^2$
$s = u(v - u)/a + \frac{1}{2}a(v - u)^2/a^2$

Multiplying both sides by $2a$ and expanding the terms,
$2as = 2uv - 2u^2 + v^2 - 2uv + u^2$
or $v^2 = u^2 + 2as$

The five equations relating the various quantities which define the motion of the particle in a straight line in uniformly accelerated motion are

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

In these equations u is the initial velocity, v is the final velocity, a is the acceleration, s is the distance travelled, and t is the time taken. The average velocity $\bar{v}$ is given by $\frac{(u+v)}{2}$.

In solving problems involving kinematics, it is important to understand the situation before you try to substitute numerical values into an equation. Identify the quantity you want to know, and then make a list of the quantities you know already. This should make it obvious which equation is to be used.

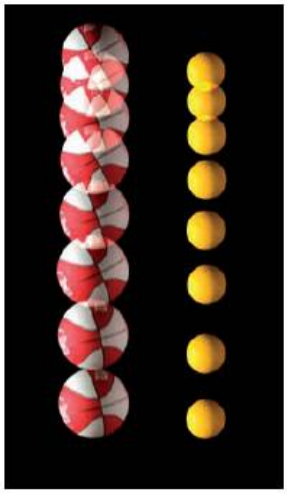


Figure 3.7 Strobo-flash photograph of objects in free fall

Free fall acceleration

A very common example of uniformly accelerated motion is when a body falls freely near the Earth's surface. Because of the gravitational attraction of the Earth, all objects fall with the same uniform acceleration. This acceleration is called the **acceleration of free fall**, and is represented by the symbol g. It has a value of 9.81 m s^{-2}, and is directed downwards. For completeness, we ought to qualify this statement by saying that the fall must be in the absence of air resistance, but in most situations this can be assumed to be true.

The acceleration of free fall may be determined by an experiment in which the time of fall t of a body between two points a distance s apart is measured.

www.iranino.co

IRANINOVIN

If the body falls from rest, we can use the second of the equations for uniformly accelerated motion in the form

$$g = 2s/t^2$$

to calculate the value of g. Note that, because the time of fall is likely to be only a few tenths of a second, precise timing to one-hundredth of a second is required. An experiment involving the switching of light gates by the falling object has been described in Topic 2 (Figure 2.24). The light gates are connected to an electronic timer.

Until the sixteenth century, the idea of the acceleration of a falling body was not fully appreciated. It was commonly thought that heavier bodies fell faster than light ones. This idea was a consequence of the effect of air resistance on light objects with a large surface area, such as feathers. However, Galileo Galilei (1564–1642) suggested that, in the absence of resistance, all bodies would fall with the same constant acceleration. He showed mathematically that, for a body falling from rest, the distance travelled is proportional to the square of the time. Galileo tested the relation experimentally by timing the fall of objects from various levels of the Leaning Tower of Pisa (Figure 3.9). This is the relation we have derived as $s = ut + \frac{1}{2}at^2$. For a body starting from rest, $u = 0$ and $s = \frac{1}{2}at^2$. That is, the distance is proportional to time squared.



Figure 3.9 Leaning Tower of Pisa



Figure 3.8 Galileo in his study

Examples

1 A car increases its speed from 25 m s^{-1} to 31 m s^{-1} with a uniform acceleration of 1.8 m s^{-2}. How far does it travel while accelerating?

In this problem we want to know the distance s. We know the initial speed $u = 25$ m s^{-1}, the final speed $v = 31$ m s^{-1}, and the acceleration $a = 1.8$ m s^{-2}. We need an equation linking s with u, v and a; this is

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

Substituting the values, we have $31^2 = 25^2 + 2 \times 1.8s$.
Re-arranging, $s = (31^2 - 25^2)/(2 \times 1.8) =$ **93 m**.

2 The average acceleration of a sprinter from the time of leaving the blocks to reaching her maximum speed of 9.0 m s^{-1} is 6.0 m s^{-2}. For how long does she accelerate? What distance does she cover in this time?

In the first part of this problem, we want to know the time t. We know the initial speed $u = 0$, the final speed $v = 9.0$ m s^{-1}, and the acceleration $a = 6.0$ m s^{-2}. We need an equation linking t with u, v and a; this is

$$v = u + at$$

Substituting the values, we have $9.0 = 0 + 6.0t$. Re-arranging, $t = 9.0/6.0 =$ **1.5 s**.

www.iranino.co

IRANINOVIN

For the second part of the problem, we want to know the distance s. We know the initial speed $u = 0$, the final speed $v = 9.0\,\text{m s}^{-1}$, and the acceleration $a = 6.0\,\text{m s}^{-2}$; we have also just found the time $t = 1.5\,\text{s}$. There is a choice of equations linking s with u, v, a and t. We can use

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

Substituting the values, $s = 0 + \frac{1}{2} \times 6.0 \times (1.5)^2 = \mathbf{6.8\,m}$.

Another relevant equation is $\bar{v} = \Delta x/\Delta t$. Here the average velocity $\bar{v}$ is given by $\bar{v} = (u + v)/2 = 4.5\,\text{m s}^{-1}$. $\Delta x/\Delta t$ is the same as s/t, so $4.5 = s/1.5$, and $s = 4.5 \times 1.5 = \mathbf{6.8\,m}$ as before.

3 A cricketer throws a ball vertically upward into the air with an initial velocity of $18.0\,\text{m s}^{-1}$. How high does the ball go? How long is it before it returns to the cricketer's hands?

In the first part of the problem, we want to know the distance s. We know the initial velocity $u = 18.0\,\text{m s}^{-1}$ upwards and the acceleration $a = g = 9.81\,\text{m s}^{-2}$ downwards. At the highest point the ball is momentarily at rest, so the final velocity $v = 0$. The equation linking s with u, v and a is

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

Substituting the values, $0 = (18.0)^2 + 2(-9.81)s$. Thus $s = -(18.0)^2/2(-9.81) = \mathbf{16.5\,m}$. Note that here the ball has an upward velocity but a downward acceleration, and that at the highest point the velocity is zero but the acceleration is not zero.

In the second part we want to know the time t for the ball's up-and-down flight. We know u and a, and also the overall displacement $s = 0$, as the ball returns to the same point at which it was thrown. The equation to use is

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

Substituting the values, $0 = 18.0t + \frac{1}{2}(-9.81)t^2$. Doing some algebra, $t(36.0 - 9.81t) = 0$. There are two solutions, $t = 0$ and $t = 36.0/9.81 = 3.7\,\text{s}$. The $t = 0$ value corresponds to the time when the displacement was zero when the ball was on the point of leaving the cricketer's hands. The answer required here is **3.7 s**.

Now it's your turn

12 An airliner must reach a speed of $110\,\text{m s}^{-1}$ to take off. If the available length of the runway is 2.4 km and the aircraft accelerates uniformly from rest at one end, what minimum acceleration must be available if it is to take off?

13 A speeding motorist passes a traffic police officer on a stationary motorcycle. The police officer immediately gives chase: his uniform acceleration is $4.0\,\text{m s}^{-2}$, and by the time he draws level with the motorist he is travelling at $30\,\text{m s}^{-1}$. How long does it take for the police officer to catch the car? If the car continues to travel at a steady speed during the chase, what is that speed?

14 A cricket ball is thrown vertically upwards with a speed of $15.0\,\text{m s}^{-1}$. What is its velocity when it first passes through a point 8.0 m above the cricketer's hands?

Graphs of the kinematic equations

It is often useful to represent the motion of a particle graphically, instead of by means of a series of equations. In this section we bring together the graphs which correspond to the equations we have already derived. We shall see that there are some important links between the graphs.

First, think about a particle moving in a straight line with constant velocity. Constant velocity means that the particle covers equal distances in equal intervals of time. A graph of displacement x against time t is thus a straight line, as in Figure 3.10. Here the particle has started at $x = 0$ and at time $t = 0$. The slope of the graph is equal to the magnitude of the velocity, since, from the definition of average velocity, $\bar{v} = (x_2 - x_1)/(t_2 - t_1) = \Delta x/\Delta t$. Because this graph is a straight line, the average velocity and the instantaneous velocity are the same. The equation describing the graph is $x = vt$.

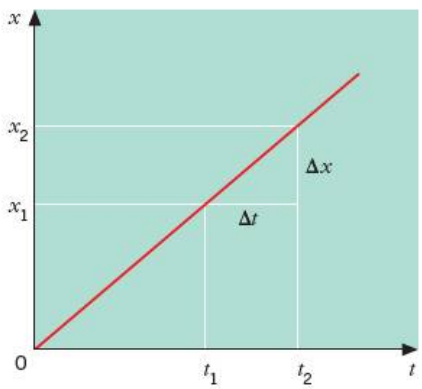


Figure 3.10

www.iranino.co

IRANINOVIN

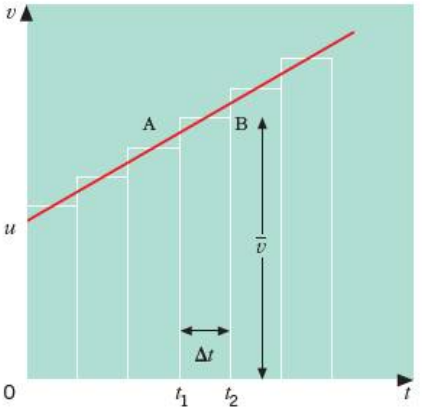


Figure 3.11

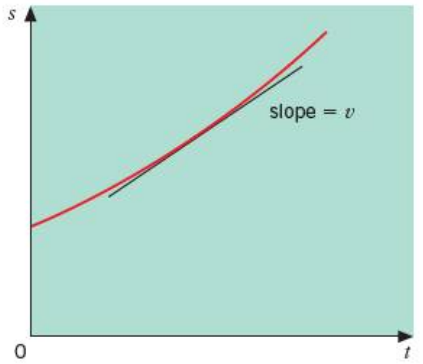


Figure 3.12

Now think about a particle moving in a straight line with constant acceleration. The particle's velocity will change by equal amounts in equal intervals of time. A graph of the magnitude v of the velocity against time t will be a straight line, as in Figure 3.11. Here the particle has started with velocity u at time $t = 0$. The slope of the graph is equal to the magnitude of the acceleration. The graph is a straight line showing that the acceleration is a constant. The equation describing the graph is $v = u + at$.

An important feature of the velocity–time graph is that we can deduce the displacement of the particle by calculating the area between the graph and the t-axis, between appropriate limits of time. Suppose we want to obtain the displacement of the particle between times t_1 and t_2 in Figure 3.11. Between these times the average $\bar{v}$ velocity is represented by the horizontal line AB. The area between the graph and the t-axis is equal to the area of the rectangle whose top edge is AB, or average velocity $\bar{v}$. This area is $\bar{v}\Delta t$. But, by the definition of average velocity ($\bar{v} = \Delta x/\Delta t$), $\bar{v}\Delta t$ is equal to the displacement Δx during the time interval Δt.

We can deduce the graph of displacement s against time t from the velocity–time graph by calculating the area between the graph and the t-axis for a succession of values of t. As shown in Figure 3.11, we can split the area up into a number of rectangles. The displacement at a certain time is then just the sum of the areas of the rectangles up to that time. Figure 3.12 shows the result of plotting the displacement s determined in this way against time t. It is a curve with a slope which increases the higher the value of t, indicating that the particle is accelerating. The slope at a particular time gives the magnitude of the instantaneous velocity. The equation describing Figure 3.12 is $s = ut + \frac{1}{2}at^2$.

Example

The displacement–time graph for a car on a straight test track is shown in Figure 3.13. Use this graph to draw velocity–time and acceleration–time graphs for the test run.

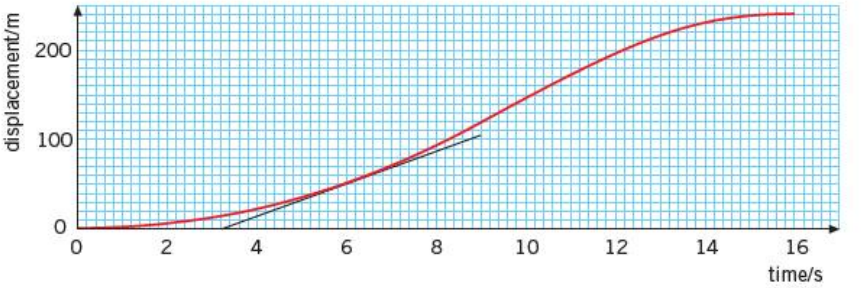


Figure 3.13 Displacement–time graph

We have already met this graph when we discussed the concepts of velocity and acceleration (Figure 3.4, page 43). In Figure 3.13 it has been re-drawn to scale, and figures have been put on the displacement and time axes. We find the magnitude of the velocity by measuring the gradient of the displacement–time graph. As an example, a tangent to the graph has been drawn at $t = 6.0$ s. The slope of this tangent is 18 m s^{-1}. If the process is repeated at different times, the following velocities are determined.

t/s	2	4	6	8	10	12	14	16
v/m s^{-1}	6	12	18	24	30	20	10	0

These values are plotted on the velocity–time graph of Figure 3.14. Check some of the values by drawing tangents yourself. *Hint*: When drawing tangents, use a mirror or a transparent ruler.

Figure 3.14 shows two straight-line portions. Initially, from $t = 0$ to $t = 10$ s, the car is accelerating uniformly, and from $t = 10$ s to $t = 16$ s it is decelerating. The acceleration is given by $a = \Delta v/\Delta t = 30/10 = 3$ m s^{-2} up to $t = 10$ s. Beyond $t = 10$ s the acceleration is $-30/6 = -5$ m s^{-2}. (The minus sign shows that the car is decelerating.)

www.iranino.co

IRANINOVIN



Figure 3.14 Velocity–time graph

The acceleration–time graph is plotted in Figure 3.15.

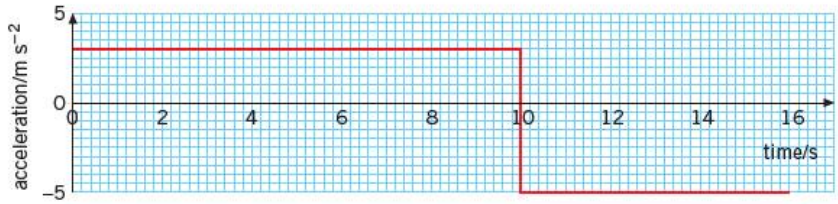


Figure 3.15 Acceleration–time graph

Finally, we can confirm that the area under a velocity–time graph gives the displacement. The area under the line in Figure 3.14 is

$$(\tfrac{1}{2} \times 10 \times 30) + (\tfrac{1}{2} \times 6 \times 30) = 240\,\text{m}$$

the value of s at $t = 16$ s on Figure 3.13.



Figure 3.16 Cricketer bowling the ball

Now it's your turn

15 In a test of a sports car on a straight track, the following readings of velocity v were obtained at the times t stated.

t/s	0	5	10	15	20	25	30	35
v/m s^{-1}	0	15	23	28	32	35	37	38

(a) On graph paper, draw a velocity–time graph and use it to determine the acceleration of the car at time $t = 5$ s.

(b) Find also the total distance travelled between $t = 0$ and $t = 30$ s.

Note: These figures refer to a case of non-uniform acceleration, which is more realistic than the previous example. However, the same rules apply: the acceleration is given by the slope of the velocity–time graph at the relevant time, and the distance travelled can be found from the area under the graph.

Two-dimensional motion under a constant force

So far we have been dealing with motion along a straight line; that is, one-dimensional motion. We will now think about the motion of particles moving in paths in two dimensions. We shall need to make use of ideas we have already learnt regarding vectors in Topic 1. The particular example we shall take is where a particle moves in a plane under the action of a constant force. An example is the motion of a ball thrown at an angle to the vertical (Figure 3.16), or an electron moving at an angle to an electric field. In the case of the ball, the constant force acting on it is its weight. For the electron, the constant force is the force provided by the electric field.

www.iranino.co

IRANINOVIN

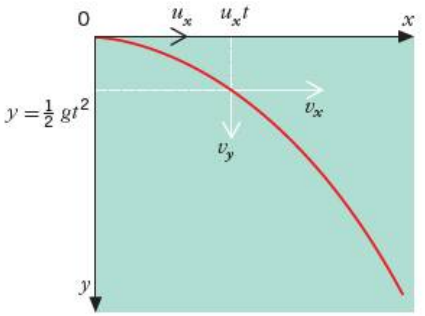


Figure 3.17

This topic is often called **projectile motion**. Galileo first gave an accurate analysis of this motion. He did so by splitting the motion up into its vertical and horizontal components, and considering these separately. The key is that the two components can be considered independently.

As an example, think about a particle sent off in a horizontal direction and subject to a vertical gravitational force (its weight). As before, air resistance will be neglected. We will analyse the motion in terms of the horizontal and vertical components of velocity. The particle is projected at time $t = 0$ at the origin of a system of x, y co-ordinates (Figure 3.17) with velocity u_x in the x-direction. Think first about the particle's vertical motion (in the y-direction). Throughout the motion, it has an acceleration of g (the acceleration of free fall) in the y-direction. The initial value of the vertical component of velocity is $u_y = 0$. The vertical component increases continuously under the uniform acceleration g. Using $v = u + at$, its value v_y at time t is given by $v_y = gt$. Also at time t, the vertical displacement y downwards is given by $y = \frac{1}{2}gt^2$. Now for the horizontal motion (in the x-direction): here the acceleration is zero, so the horizontal component of velocity remains constant at u_x. At time t the horizontal displacement x is given by $x = u_x t$. To find the velocity of the particle at any time t, the two components v_x and v_y must be added vectorially. The direction of the resultant vector is the direction of motion of the particle. The curve traced out by a particle subject to a constant force in one direction is a **parabola**.



Figure 3.18 Water jets from a garden sprinkler showing a parabola-shaped spray

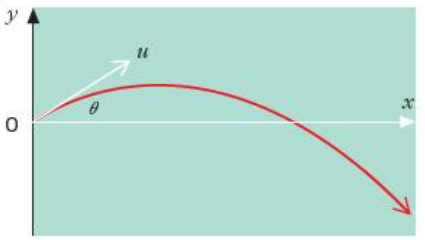


Figure 3.19

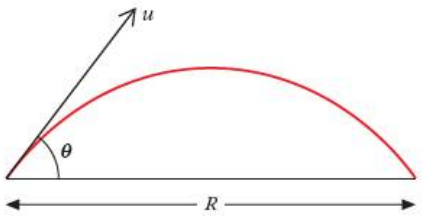


Figure 3.20

If the particle had been sent off with velocity u at an angle θ to the horizontal, as in Figure 3.19, the only difference to the analysis of the motion is that the initial y-component of velocity is $u \sin \theta$. In the example illustrated in Figure 3.19, this is upwards. Because of the downwards acceleration g, the y-component of velocity decreases to zero, at which time the particle is at the crest of its path, and then increases in magnitude again but this time in the opposite direction. The path is again a parabola.

For the particular case of a particle projected with velocity u at an angle θ to the horizontal from a point on level ground (Figure 3.20), the range R is defined as the distance from the point of projection to the point at which the particle reaches the ground again. We can show that R is given by

$$R = \frac{(u^2 \sin 2\theta)}{g}$$

For details, see the Maths Note opposite.

www.iranino.co

IRANINOVIN

Maths Note

Suppose that the particle is projected from the origin ($x = 0$, $y = 0$). We can interpret the range R as being the horizontal distance x travelled at the time t when the value of y is again zero. The equation which links displacement, initial speed, acceleration and time is $s = ut + \frac{1}{2}at^2$. Adapting this for the vertical component of the motion, we have

$$0 = (u \sin \theta)t - \tfrac{1}{2}gt^2$$

The two solutions of this equation are $t = 0$ and $t = (2u \sin \theta)/g$. The $t = 0$ case is when the particle was projected; the second is when it returns to the ground at $y = 0$. We use this second value of t with the horizontal component of velocity $u \cos \theta$ to find the distance x travelled (the range R). This is

$$x = R = (u \cos \theta)t = (2u^2 \sin \theta \cos \theta)/g$$

There is a trigonometric relationship $\sin 2\theta = 2 \sin \theta \cos \theta$, use of which puts the range expression in the required form

$$R = (u^2 \sin 2\theta)/g$$

We can see that R will have its maximum value for a given speed of projection u when $\sin 2\theta = 1$, that is when $2\theta = 90°$, or $\theta = 45°$. The value of this maximum range is $R_{max} = u^2/g$.

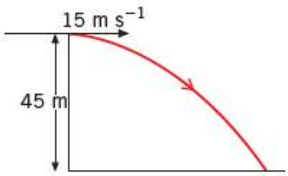


Figure 3.21

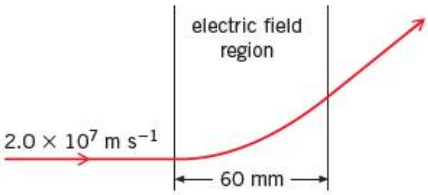


Figure 3.22

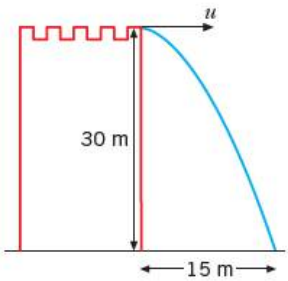


Figure 3.23

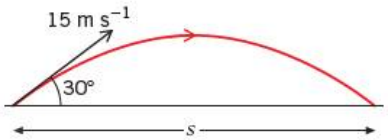


Figure 3.24

Examples

1 A stone is thrown from the top of a vertical cliff, 45 m high above level ground, with an initial velocity of 15 m s^{-1} in a horizontal direction (Figure 3.21). How long does it take to reach the ground? How far from the base of the cliff is it when it reaches the ground?

To find the time t for which the stone is in the air, work with the vertical component of the motion, for which we know that the initial component of velocity is zero, the displacement $y = 45$ m, and the acceleration a is 9.81 m s^{-2}. The equation linking these is $y = \frac{1}{2}gt^2$. Substituting the values, we have $45 = \frac{1}{2} \times 9.81t^2$. This gives

$t = \sqrt{(2 \times 45 / 9.81)} = \mathbf{3.0\,s}$.

For the second part of the question, we need to find the horizontal distance x travelled in the time t. Because the horizontal component of the motion is not accelerating, x is given simply by $x = u_x t$. Substituting the values, we have $x = 15 \times 3.0 = \mathbf{45\,m}$.

2 An electron, travelling with a velocity of 2.0×10^7 m s^{-1} in a horizontal direction, enters a uniform electric field. This field gives the electron a constant acceleration of 5.0×10^{15} m s^{-2} in a direction perpendicular to its original velocity (Figure 3.22). The field extends for a horizontal distance of 60 mm. What is the magnitude and direction of the velocity of the electron when it leaves the field?

The horizontal motion of the electron is not accelerated. The time t spent by the electron in the field is given by $t = x/u_x = 60 \times 10^{-3}/2.0 \times 10^7 = 3.0 \times 10^{-9}$ s. When the electron enters the field, its vertical component of velocity is zero; in time t, it has been accelerated to $v_y = at = 5.0 \times 10^{15} \times 3.0 \times 10^{-9} = 1.5 \times 10^7$ m s^{-1}. When the electron leaves the field, it has a horizontal component of velocity $v_x = 2.0 \times 10^7$ m s^{-1}, unchanged from the initial value u_x. The vertical component is $v_y = 1.5 \times 10^7$ m s^{-1}. The resultant velocity v is given by:

$$v = \sqrt{(v_x^2 + v_y^2)} = \sqrt{[(2.0 \times 10^7)^2 + (1.5 \times 10^7)^2]} = \mathbf{2.5 \times 10^7\,m\,s^{-1}}.$$

The direction of this resultant velocity makes an angle θ to the horizontal, where θ is given by $\tan \theta = v_y/v_x = 1.5 \times 10^7/2.0 \times 10^7$. The angle θ is **37°**.

Now it's your turn

16 A ball is thrown horizontally from the top of a tower 30 m high and lands 15 m from its base (Figure 3.23). What is the ball's initial speed?

17 A football is kicked on level ground at a velocity of 15 m s^{-1} at an angle of 30° to the horizontal (Figure 3.24). How far away is the first bounce?

18 A car accelerates from 5.0 m s^{-1} to 20 m s^{-1} in 6.0 s. Assuming uniform acceleration, how far does it travel in this time?

19 If a raindrop were to fall from a height of 1 km, with what velocity would it hit the ground if there were no air resistance?

20 Traffic police can estimate the speed of vehicles involved in accidents by the length of the marks made by skidding tyres on the road surface. It is known that the maximum deceleration that a car can attain when braking on a normal road surface is about 9 m s^{-2}. In one accident, the tyre-marks were found to be 125 m long. Estimate the speed of the vehicle before braking.

www.iranino.co

IRANINOVIN

21 On a theme park ride, a cage is travelling upwards at constant speed. As it passes a platform alongside, a passenger drops coin A through the cage floor. At exactly the same time, a person standing on the platform drops coin B from the platform.
(a) Which coin, A or B (if either), reaches the ground first?
(b) Which (if either) has the greater speed on impact?

22 William Tell was faced with the agonising task of shooting an apple placed on his son Jemmy's head.

Assume that William is placed 25 m from Jemmy; his crossbow fires a bolt with an initial speed of 45 m s^{-1}. The crossbow and apple are on the same horizontal line. At what angle to the horizontal should William aim so that the bolt hits the apple?

23 The position of a sports car on a straight test track is monitored by taking a series of photographs at fixed time intervals. The following record of position x was obtained at the stated times t.

t/s	0	0.5	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
x/m	0	0.4	1.8	4.2	7.7	12.4	18.3	25.5	33.9	43.5	54.3

On graph paper, draw a graph of x against t. Use your graph to obtain values for the velocity v of the car at a number of values of t. Draw a second graph of v against t. From this graph, what can you deduce about the acceleration of the car?

Summary

- Speed is a scalar quantity and is described by magnitude only. Velocity is a vector quantity and requires magnitude and direction.
- Displacement is the distance travelled in a straight line in a specified direction and is a vector quantity.
- Average speed is defined by: (distance travelled)/(time taken)
- Average velocity is defined by: (displacement)/(time taken) or $\Delta x/\Delta t$
- The instantaneous velocity is the average velocity measured over an infinitesimally short time interval.
- Average acceleration is defined by: (change in velocity)/(time taken) or $\Delta v/\Delta t$
- Acceleration is a vector. Instantaneous acceleration is the average acceleration measured over an infinitesimally short time interval.
- The gradient of a displacement–time graph gives the velocity.
- The gradient of a velocity–time graph gives the acceleration.
- The area between a velocity–time graph and the time axis gives the displacement.
- The equations for a body moving in a straight line with uniform acceleration are:

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

- Objects falling freely near the surface of the Earth in the absence of air resistance, experience the same acceleration, the acceleration of free fall g, which has the value $g = 9.81$ m s^{-2}.
- The motion of projectiles is analysed in terms of two independent motions at right angles. The horizontal component of the motion is at a constant velocity, while the vertical motion is subject to a constant acceleration g.

www.iranino.co

IRANINOVIN

Examination style questions

1 In a driving manual, it is suggested that, when driving at 13 m s^{-1} (about 45 km per hour), a driver should always keep a minimum of two car lengths between the driver's car and the one in front.
   a Suggest a scientific justification for this safety tip, making reasonable assumptions about the magnitudes of any quantities you need.
   b How would you expect the length of this 'exclusion zone' to depend on speed for speeds higher than 13 m s^{-1}?

2 A student, standing on the platform at a railway station, notices that the first two carriages of an arriving train pass her in 2.0 s, and the next two in 2.4 s. The train is decelerating uniformly. Each carriage is 20 m long. When the train stops, the student is opposite the last carriage. How many carriages are there in the train?

3 A ball is to be kicked so that, at the highest point of its path, it just clears a horizontal cross-bar on a pair of goal-posts. The ground is level and the cross-bar is 2.5 m high. The ball is kicked from ground level with an initial speed of 8.0 m s^{-1}.
   a Calculate the angle of projection of the ball and the distance of the point where the ball was kicked from the goal-line.
   b Also calculate the horizontal velocity of the ball as it passes over the cross-bar.
   c For how long is the ball in the air before it reaches the ground on the far side of the cross-bar?

4 An athlete competing in the long jump leaves the ground at an angle of 28° and makes a jump of 7.40 m.
   a Calculate the speed at which the athlete took off.
   b If the athlete had been able to increase this speed by 5%, what percentage difference would this have made to the length of the jump?

5 A hunter, armed with a bow and arrow, takes direct aim at a monkey hanging from the branch of a tree. At the instant that the hunter releases the arrow, the monkey takes avoiding action by releasing its hold on the branch. By setting up the relevant equations for the motion of the monkey and the motion of the arrow, show that the monkey was mistaken in its strategy.

6 A ball is thrown horizontally from the top of a building, as shown in Fig. 3.25.

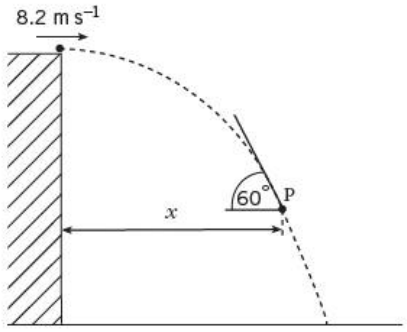


Fig. 3.25

The ball is thrown with a horizontal speed of 8.2 m s^{-1}. The side of the building is vertical. At point P on the path of the ball, the ball is distance x from the building and is moving at an angle of 60° to the horizontal. Air resistance is negligible.

   a For the ball at point P,
      i show that the vertical component of its velocity is 14.2 m s^{-1}, [2]
      ii determine the vertical distance through which the ball has fallen, [2]
      iii determine the horizontal distance x. [2]
   b The path of the ball in **a**, with an initial horizontal speed of 8.2 m s^{-1}, is shown again in Fig. 3.26.

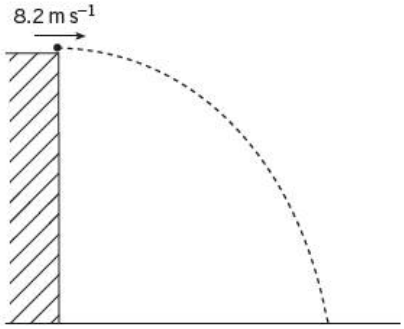


Fig. 3.26

On a copy of Fig. 3.26, sketch the new path of the ball for the ball having an initial horizontal speed
      i greater than 8.2 m s^{-1} and with negligible air resistance (label this path G) [2]
      ii equal to 8.2 m s^{-1} but with air resistance (label this path A). [2]

Cambridge International AS and A Level Physics, 9702/21 Oct/Nov 2010 Q 2

7 A student has been asked to determine the linear acceleration of a toy car as it moves down a slope. He sets up the apparatus as shown in Fig. 3.27.

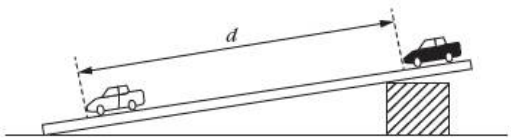


Fig. 3.27

The time t to move from rest through a distance d is found for different values of d. A graph of d (y-axis) is plotted against t^2 (x-axis) as shown in Fig. 3.28.

www.iranino.co

IRANINOVIN

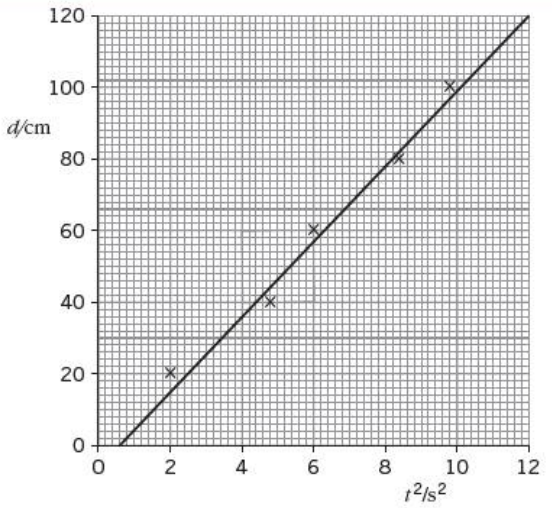


Fig. 3.28

a Theory suggests that the graph is a straight line through the origin.

Name the feature on Fig. 3.28 that indicates the presence of

i random error,

ii systematic error. *[2]*

b i Determine the gradient of the line of the graph in Fig. 3.28. *[2]*

ii Use your answer to **i** to calculate the acceleration of the toy down the slope. Explain your working. *[3]*

Cambridge International AS and A level Physics, 9702/02 May/June 2004 Q 3

8 A car is travelling along a straight road at speed v. A hazard suddenly appears in front of the car. In the time interval between the hazard appearing and the brakes on the car coming into operation, the car moves forward a distance of 29.3 m. With the brakes applied, the front wheels of the car leave skid marks on the road that are 12.8 m long, as illustrated in Fig. 3.29.

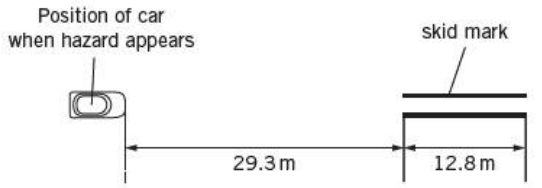


Fig. 3.29

It is estimated that, during the skid, the magnitude of the deceleration of the car is 0.85 g, where g is the acceleration of free fall.

a Determine

i the speed v of the car before the brakes are applied, *[2]*

ii the time interval between the hazard appearing and the brakes being applied. *[2]*

b The legal speed limit on the road is 60 km per hour.

Use both of your answers in **a** to comment on the standard of the driving of the car. *[3]*

Cambridge International AS and A level Physics, 9702/02 Oct/Nov 2008 Q 2

9 a Define

i *velocity*, *[1]*

ii *acceleration*. *[1]*

b A car of mass 1500 kg travels along a straight horizontal road.

The variation with time t of the displacement x of the car is shown in Fig. 3.30.

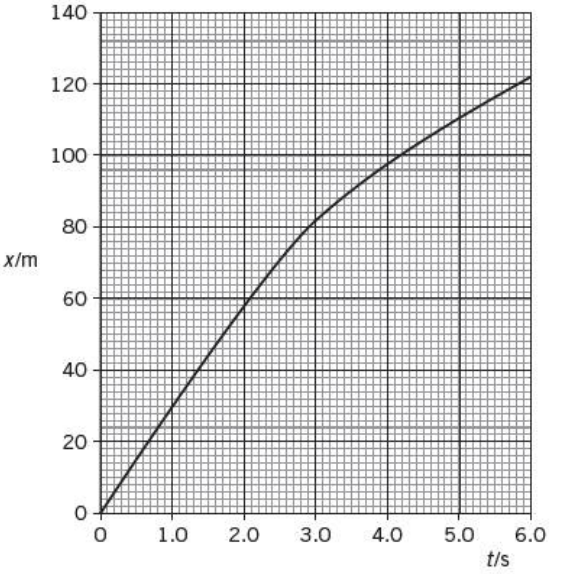


Fig. 3.30

i Use Fig. 3.30 to describe qualitatively the velocity of the car during the first six seconds of the motion shown. Give reasons for your answers. *[3]*

ii Calculate the average velocity during the time interval $t = 0$ to $t = 1.5$ s. *[1]*

iii Show that the average acceleration between $t = 1.5$ s and $t = 4.0$ s is $-7.2\ \text{m s}^{-2}$. *[2]*

iv Calculate the average force acting on the car between $t = 1.5$ s and $t = 4.0$ s. *[2]*

Cambridge International AS and A level Physics, 9702/23 Oct/Nov 2013 Q 3

www.iranino.co

IRANINOVIN

4 Dynamics

By the end of this topic, you will be able to:

4.1 (a) understand that mass is the property of a body that resists change in motion
(b) recall the relationship $F = ma$, and solve problems using it, appreciating that acceleration and force are always in the same direction
(c) define and use linear momentum as the product of mass and velocity
(d) define and use force as the rate of change of momentum
(e) state and apply each of Newton's laws of motion

4.2 (a) describe and use the concept of weight as the effect of a gravitational field on a mass and recall that the weight of a body is equal to the product of its mass and the acceleration of free fall
(b) describe qualitatively the motion of bodies falling in a uniform gravitational field with air resistance

4.3 (a) state the principle of conservation of momentum
(b) apply the principle of conservation of momentum to solve simple problems including elastic and inelastic interactions between bodies in both one and two dimensions
(c) recognise that, for a perfectly elastic collision, the relative speed of approach is equal to the relative speed of separation
(d) understand that, while momentum of a system is always conserved in interactions between bodies, some change in kinetic energy may take place

Starting points

- Motion of an object can be described in terms of displacement, velocity and acceleration.
- A force is required to make a body accelerate.

4.1 Relationships involving force and mass

When you push a trolley in a supermarket or pull a case behind you at an airport, you are exerting a force. When you hammer in a nail, a force is being exerted. When you drop a book and it falls to the floor, the book is falling because of the force of gravity. When you lean against a wall or sit on a chair, you are exerting a force. Forces can change the shape or dimensions of bodies. You can crush a drinks can by squeezing it and applying a force; you can stretch a rubber band by pulling it. In everyday life, we have a good understanding of what is meant by force and the situations in which forces are involved. In physics the idea of force is used to add detail to the descriptions of moving objects.

As with all physical quantities, a method of measuring force must be established. One way of doing this is to make use of the fact that forces can change the dimensions of bodies in a reproducible way. It takes the same force to stretch a spring by the same change in length (provided the spring is not overstretched by applying a very large force). This principle is used in the spring balance. A scale shows how much the spring has been extended, and the scale can be calibrated in terms of force. Laboratory spring balances are often called newton balances, because the newton is the SI unit of force.

Forces are vector quantities: they have magnitude as well as direction. A number of forces acting on a body are often shown by means of a force diagram drawn to scale, in which the forces are represented by lines of length proportional to the magnitude of the force, and in the appropriate direction (see Topic 1). The combined effect of several forces acting on a body is known as the **resultant** force.

www.iranino.co

IRANINOVIN



Figure 4.1

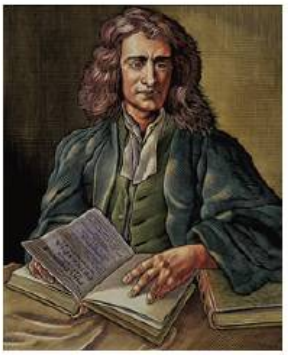


Figure 4.2 Isaac Newton

Force and motion

The Greek philosopher Aristotle believed that the natural state of a body was a state of rest, and that a force was necessary to make it move and to keep it moving. This argument requires that the greater the force, the greater the speed of the body.

Nearly two thousand years later, Galileo questioned this idea. He suggested that motion at a constant speed could be just as natural a state as the state of rest. He introduced an understanding of the effect of **friction** on motion.

Imagine a heavy box being pushed along a rough floor at constant speed (Figure 4.1). This may take a considerable force. The force required can be reduced if the floor is made smooth and polished, and reduced even more if a lubricant, for example grease, is applied between the box and the floor. We can imagine a situation where, when friction is reduced to a vanishingly small value, the force required to push the box at constant speed is also vanishingly small.

Galileo realised that the force of friction was a force that opposed the pushing force. When the box is moving at constant speed, the pushing force is exactly equal to the frictional force, but in the opposite direction, so that there is a net force of zero acting on the box. In the situation of vanishingly small friction, the box will continue to move with constant speed, because there is no force to slow it down.

Newton's laws of motion

Isaac Newton (1642–1727) used Galileo's ideas to produce a theory of motion, expressed in his three laws of motion. The **first law of motion** re-states Galileo's deduction about the natural state of a body.

> Every body continues in its state of rest, or with uniform velocity, unless acted on by a resultant force.

This law tells us what a force does: it disturbs the state of rest or velocity of a body. The property of a body to stay in a state of rest or uniform velocity is called **inertia**.

Newton's second law tells us what happens if a force is exerted on a body. It causes the velocity to change. A force exerted on a body at rest makes it move – it gives it a velocity. A force exerted on a moving body may make its speed increase or decrease, or change its direction of motion. A change in speed or velocity is acceleration. Newton's second law relates the magnitude of this acceleration to the force applied. It also introduces the idea of the mass of a body. Mass is a measure of the inertia of a body to change in velocity. The bigger the mass, the more difficult it is to change its state of rest or velocity. A simplified form of Newton's **second law** is

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

The direction of the acceleration is in the direction of the resultant force. In a word equation the relation between force and acceleration is

> *force = mass × acceleration*

and in symbols

> $F = ma$

where F is the resultant force, m is the mass and a is the acceleration. Here we have made the constant of proportionality equal to unity (that is, we use an equals sign rather than a proportionality sign) by choosing quantities with units which will give us this simple relation. In SI units, the force F is in newtons, the mass m in kilograms and the acceleration a in metres (seconds)$^{-2}$.

> One newton is defined as the force which will give a 1 kg mass an acceleration of 1 m s^{-2} in the direction of the force.

www.iranino.co

IRANINOVIN



Figure 4.3



Figure 4.4 Space rocket launch

When you push a supermarket trolley, the trolley experiences a force (Figure 4.3). The trolley applies an equal and opposite force on another body – you. Newton understood that the body on which the force is exerted applies another force back on the body which is applying the force. When body A applies a force on body B then body B applies an equal and opposite force on body A. Newton's **third law** relates these two forces.

Whenever one body exerts a force on another, the second body exerts an equal and opposite force on the first.

Very often this law is stated as:

To every action, there is an equal and opposite reaction.

But this statement does not highlight the very important point that the action force and the reaction force act on *different objects*. To take the example of the supermarket trolley, the action force exerted by you on the trolley is equal and opposite to the reaction force exerted by the trolley on you.

Newton's third law has applications in every branch of everyday life. We walk because of this law. When you take a step forward, your foot presses against the ground. The ground then exerts an equal and opposite force on you. This is the force, on you, which propels you in your path. Space rockets work because of the law (Figure 4.4). To expel the exhaust gases from the rocket, the rocket exerts a force on the gases. The gases exert an equal and opposite force on the rocket, accelerating it forward.

Examples

1 An object of mass 1.5 kg is to be accelerated at 2.2 m s^{-2}. What force is required?

From Newton's second law, $F = ma = 1.5 \times 2.2 =$ **3.3 N**.

2 A car of mass 1.5 tonnes (1.5×10^3 kg), travelling at 80 km h^{-1}, is to be stopped in 11 s. What force is required?

The acceleration of the car can be obtained from $v = u + at$ (see Topic 3). The initial speed u is 80 km h^{-1}, or 22 m s^{-1}. The final speed v is 0. Then $a = -22/11 = -2.0$ m s^{-2}. This is negative because the car is decelerating.

By Newton's second law, $F = ma = 1.5 \times 10^3 \times 2.0 =$ **3.0×10^3 N**.

Now it's your turn

1 A force of 5.0 N is applied to a body of mass 3.0 kg. What is the acceleration of the body?

2 A stone of mass 50 g is accelerated from a catapult to a speed of 8.0 m s^{-1} from rest over a distance of 30 cm. What average force is applied by the rubber of the catapult?

Momentum

We shall now introduce a quantity called **momentum**, and see how Newton's laws are related to it.

The momentum of a particle is defined as the product of its mass and its velocity.

In words

momentum = mass × velocity

and in symbols

$p = mv$

The unit of momentum is the unit of mass times the unit of velocity; that is, kg m s^{-1}. An alternative unit is the newton second (Ns). Momentum, like velocity, is a vector quantity. Its complete name is **linear momentum**, to distinguish it from angular momentum, which does not concern us here.

www.iranino.co

IRANINOVIN

Newton's first law states that every body continues in a state of rest, or with uniform velocity, unless acted on by a resultant force. We can express this law in terms of momentum. If a body maintains its uniform velocity, its momentum is unchanged. If a body remains at rest, again its momentum (zero) does not change. Thus, an alternative statement of the first law is that the momentum of a particle remains constant unless an external resultant force acts on the particle. As an equation

$$p = \text{constant}$$

This is a special case, for a single particle, of a very important conservation law: the principle of conservation of momentum. This word 'conservation' here means that the quantity remains constant.

Newton's second law is expressed in terms of momentum. We already have it in a form which relates the force acting on a body to the product of the mass and the acceleration of the body. Remember that the acceleration of a body is the rate of change of its velocity. The product of mass and acceleration then is just the mass times the rate of change of velocity. For a body of constant mass, this is just the same as the rate of change of (mass × velocity). But (mass × velocity) is momentum, so the product of mass and acceleration is identical to the rate of change of momentum. Thus, Newton's second law is stated as

The resultant force acting on a body is proportional to the rate of change of its momentum.

The constant of proportionality is made equal to unity as described on page 56. Hence the second law used in problem solving is

The resultant force acting on a body is equal to the rate of change of momentum.

Expressed in terms of symbols

$$F = \Delta p/\Delta t$$

for constant mass m

$$F = \Delta(mv)/\Delta t = m(\Delta v/\Delta t) = ma$$

Note that F represents the **resultant force** acting on the body.

Continuing with the idea of force being equal to rate of change of momentum, the third law relating to action and reaction forces becomes: the rate of change of momentum due to the action force on one body is equal and opposite to the rate of change of momentum due to the reaction force on the other body. The action and reaction forces act on each body for the same time (Δt). Hence $F\Delta t$ is equal and opposite for each body. Therefore when two bodies exert action and reaction forces on each other, their changes of momentum are equal and opposite.

4.2 Weight

We saw in Topic 3 that all objects released near the surface of the Earth fall with the same acceleration (the acceleration of free fall) if air resistance can be neglected. The force causing this acceleration is the gravitational attraction of the Earth on the object, or the force of gravity. The force of gravity which acts on an object is called the **weight** of the object. We can apply Newton's second law to the weight. For a body of mass m falling with the acceleration of free fall g, the weight W is given by

$$W = mg$$

The SI unit of force is the newton (N). This is also the unit of weight. The weight of an object is obtained by multiplying its mass in kilograms by the acceleration of free

www.iranino.co

IRANINOVIN



Figure 4.5 A newton balance

fall, 9.81 m s^{-2}. Thus a mass of one kilogram has a weight of 9.81 N. Because weight is a force and force is a vector, we ought to be aware of the direction of the weight of an object. It is towards the centre of the Earth. Because weight always has this direction, we do not need to specify direction every time we give the magnitude of the weight of objects.

How do we measure mass and weight? If you hang an object from a newton balance, you are measuring its weight (Figure 4.5). The unknown weight of the object is balanced by a force provided by the spring in the balance. From a previous calibration, this force is related to the extension of the spring. There is the possibility of confusion here. Laboratory newton balances may, indeed, be calibrated in newtons. But all commercial spring balances – for example, the balances at fruit and vegetable counters in supermarkets – are calibrated in kilograms. Such balances are really measuring the weight of the fruit and vegetables, but the scale reading is in mass units, because there is no distinction between mass and weight in everyday life. The average shopper thinks of 5 kg of mangoes as having a weight of 5 kg. In fact, the mass of 5 kg has a weight of 49 N.

The word 'balance' in the spring balance and in the laboratory top-pan balance relates to the balance of forces. In each case, the unknown force (the weight) is equalled by a force which is known through calibration.

A way of comparing masses is to use a beam balance, or lever balance (see Figure 2.17). Here the weight of the object is balanced against the weight of some masses, which have previously been calibrated in mass units. The word 'balance' here refers to the equilibrium of the beam: when the beam is horizontal, the moment of the weight on one side of the pivot is equal and opposite to the moment on the other side of the pivot. Because weight is given by the product of mass and the acceleration of free fall, the equality of the weights means that the masses are also equal.

We have introduced the idea of weight by thinking about an object in free fall. But objects at rest also have weight: the gravitational attraction on a book is the same whether it is falling or whether it is resting on a table. The fact that the book is at rest tells us, by Newton's first law, that the resultant force acting on it is zero. So there must be another force acting on the book which exactly balances its weight. In Figure 4.6 the table exerts an upwards force on the book. This force is equal in magnitude to the weight but opposite in direction. It is a **normal contact force**: 'contact' because it occurs due to the contact between book and table, and 'normal' because it acts perpendicularly to the plane of contact.

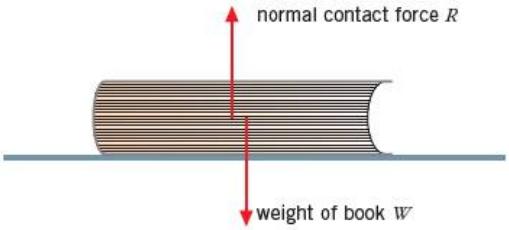


Figure 4.6 A book resting on a table: forces on the book. (The forces act in the same vertical line, but are separated slightly for clarity.)

The book remains at rest on the table because the weight W of the book downwards is exactly balanced by the normal contact force R exerted by the table on the book. The vector sum of these forces is zero, so the book is in equilibrium. A very common mistake is to state that 'By Newton's third law, W is equal to R. But these two forces are both acting on the book, and cannot be related by the third law. Third-law forces always act on *different* bodies.

To see the application of the third law, think about the normal contact force R. This is an upwards force exerted by the *table*. The reaction to this is a downwards force R' exerted by the *book*. By Newton's third law, these forces are equal and opposite. This situation is illustrated in Figure 4.7.

www.iranino.co

IRANINOVIN

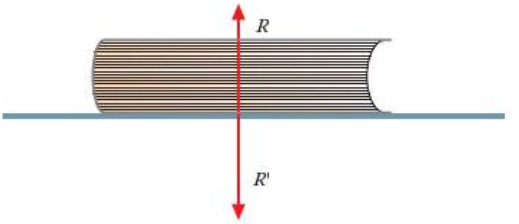


Figure 4.7 A book resting on a table: action and reaction forces, acting at the point of contact

Having considered the action and reaction forces between book and table, we ought to think about the reaction force to the weight of the book, regarded as an action force, even when the book is not on the table. This is not so easy, because there does not seem to be an obvious second force. But remember that the weight is due to the gravitational attraction of the Earth on the book. If the Earth attracts the book, the book also attracts the Earth. This gravitational force of the book on the Earth is the reaction force. We can test whether the two forces do, indeed, act on different bodies. The action force (the weight of the book) acts on the book. The reaction force (the attraction of the Earth to the book) acts on the Earth. Thus, the condition that action and reaction forces should act on different bodies is satisfied.

Non-uniform motion



Figure 4.8 A parachutist about to land

We have mentioned that, in most situations, air resistance can be neglected. In fact, there are some applications in which this resistance is most important. One such case is the fall of a parachutist, where air resistance plays a vital part. The velocity of a body falling through a resistive fluid (a liquid or a gas) does not increase indefinitely, but eventually reaches a maximum velocity, called the terminal velocity. The force due to air resistance increases with speed. When this resistive force has reached a value equal and opposite to the weight of the falling body, the body no longer accelerates and continues at uniform velocity. This is a case of motion with non-uniform acceleration. The acceleration starts off with a value of g, but decreases to zero at the time when the **terminal velocity** is achieved. Thus, raindrops and parachutists are normally travelling at a constant speed by the time they approach the ground (Figure 4.8).

Problem solving

In dealing with problems involving Newton's laws, start by drawing a general sketch of the situation. Then consider each body in your sketch. Show all the forces acting on that body, both known forces and unknown forces you may be trying to find. Here it is a real help to try to draw the arrows which represent the forces in approximately the correct direction and approximately to scale. Label each force with its magnitude or with a symbol if you do not know the magnitude. For each force, you must know the cause of the force (gravity, friction, and so on), and you must also know *on* what object that force acts and *by* what object it is exerted. This labelled diagram is referred to as a **free-body diagram**, because it detaches the body from the others in the situation. Having established all the forces acting on the body, you can use Newton's second law to find unknown quantities. This procedure is illustrated in the example which follows on page 61.

Newton's second law equates the resultant force acting on a body to the product of its mass and its acceleration. In some problems, the system of bodies is in equilibrium. They are at rest, or are moving in a straight line with uniform speed. In this case, the acceleration is zero, so the resultant force is also zero. In other cases, the resultant force is not zero and the objects in the system are accelerating.

Whichever case applies, you should remember that forces are vectors. You will probably have to resolve the forces into two components at right angles, and then apply the second law to each set of components separately. Problems can often be simplified by making a good choice of directions for resolution. You will end up with a set of equations, based on the application of Newton's second law, which must be solved to determine the unknown quantity.

www.iranino.co

IRANINOVIN

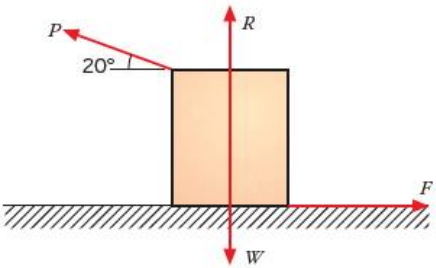


Figure 4.9

Examples

1 A box of mass 5.0 kg is pulled along a horizontal floor by a force P of 25 N, applied at an angle of 20° to the horizontal (Figure 4.9). A frictional force F of 20 N acts parallel to the floor.

Calculate the acceleration of the box.

The free-body diagram is shown in Figure 4.9. Resolving the forces parallel to the floor, the component of the pulling force, acting to the left, is 25 cos 20 = 23.5 N.
The frictional force, acting to the right, is 20 N.
The resultant force to the left is thus 23.5 – 20.0 = 3.5 N.

From Newton's second law, $a = F/m = 3.5/5.0 =$ **0.70 m s^{-2}**.

2 What is the magnitude of the momentum of an α-particle of mass 6.6×10^{-27} kg travelling with a speed of 2.0×10^{7} m s^{-1}?

$p = mv = 6.6 \times 10^{-27} \times 2.0 \times 10^{7} =$ **1.3 × 10^{-19} kg m s^{-1}**

Now it's your turn

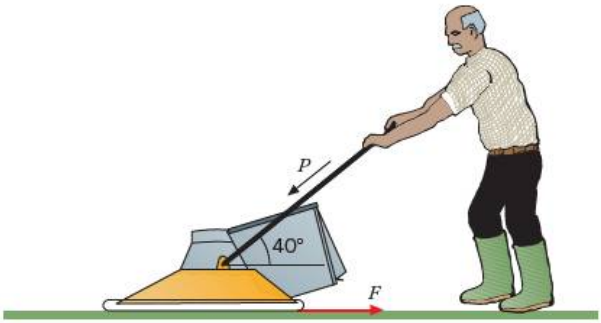


Figure 4.10

3 A person gardening pushes a lawnmower of mass 18 kg at constant speed. To do this requires a force P of 80 N directed along the handle, which is at an angle of 40° to the horizontal (Figure 4.10).
(a) Calculate the horizontal retarding force F on the mower.
(b) If this retarding force were constant, what force, applied along the handle, would accelerate the mower from rest to 1.2 m s^{-1} in 2.0 s?

4 What is the magnitude of the momentum of an electron of mass 9.1×10^{-31} kg travelling with a speed of 7.5×10^{6} m s^{-1}?

4.3 The principle of conservation of momentum



Figure 4.11 System of two particles

We have already seen that Newton's first law states that the momentum of a single particle is constant, if no external force acts on the particle. Now think about a system of two particles (Figure 4.11). We allow these particles to exert some sort of force on each other: it could be gravitational attraction or, if the particles were charged, it could be electrostatic attraction or repulsion.

These two particles are isolated from the rest of the Universe, and experience no outside forces at all. If the first particle exerts a force F on the second, Newton's third law tells us that the second exerts a force $-F$ on the first. The minus sign indicates that the forces are in opposite directions. As we saw in the last section, we can express this law in terms of change of momentum. The change of momentum of the second particle as a result of the force exerted on it by the first is equal and opposite to the change of momentum of the first particle as a result of the force exerted on it by the second. Thus, the changes of momentum of the individual particles cancel out, and the momentum of the system of two particles remains constant. The particles have merely exchanged some momentum.

www.iranino.co

IRANINOVIN

The situation is expressed by the equation

$$p = p_1 + p_2 = \text{constant}$$

where p is the total momentum, and p_1 and p_2 are the individual momenta.

We could extend this idea to a system of three, four, or finally any number n of particles.

If no external force acts on a system, the total momentum of the system remains constant, or is conserved.

A system on which no external force acts is often called an *isolated system*. The fact that the total momentum of an isolated system is constant is the **principle of conservation of momentum**. It is a direct consequence of Newton's third law of motion.

Collisions

We now use the principle of conservation of momentum to analyse a system consisting of two colliding particles. (If you want a real example to think about, try snooker balls.)

Consider two particles A and B making a direct, head-on collision. Particle A has mass m_1 and is moving with velocity u_1 in the direction from left to right; B has mass m_2 and has velocity u_2 in the direction from right to left (Figure 4.12). As velocity is a vector quantity, this is the same as saying that the velocity is $-u_2$ from left to right. The particles collide. After the collision they have velocities $-v_1$ and v_2 respectively in the direction from left to right. That is, both particles are moving back along their directions of approach.

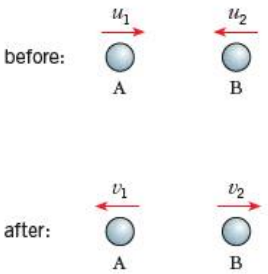


Figure 4.12 Collision between two particles

According to the principle of conservation of momentum, the total momentum of this isolated system remains constant, whatever happens as a result of the interaction of the particles. Thus, the total momentum before the collision must be equal to the total momentum after the collision. The momentum before the collision is

$$m_1u_1 - m_2u_2$$

and the momentum after is

$$-m_1v_1 + m_2v_2$$

Because total momentum is conserved

$$m_1u_1 - m_2u_2 = -m_1v_1 + m_2v_2$$

Knowing the masses of the particles and the velocities before collision, this equation would allow us to calculate the relation between the velocities after the collision.

The way to approach collision problems is as follows.

- Draw a labelled diagram showing the colliding bodies before collision. Draw a separate diagram showing the situation after the collision. Take care to define the directions of all the velocities.
- Obtain an expression for the total momentum before the collision, remembering that momentum is a vector quantity. Similarly, find the total momentum after the collision, taking the same reference direction.
- Then equate the momentum before the collision to the momentum afterwards.

Example

A cannon of mass 1.5 tonnes (1.5×10^3 kg) fires a cannon-ball of mass 5.0 kg (Figure 4.13). The speed with which the ball leaves the cannon is 70 m s^{-1} relative to the Earth. What is the initial speed of recoil of the cannon?

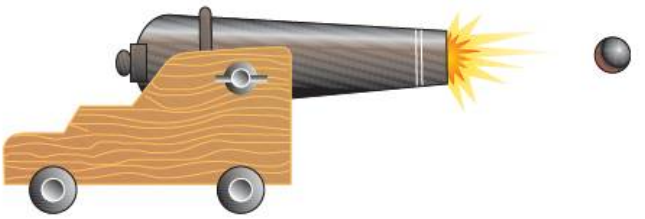


Figure 4.13

www.iranino.co

IRANINOVIN

The system under consideration is the cannon and the cannon-ball. The total momentum of the system before firing is zero. Because the total momentum of an isolated system is constant, the total momentum after firing must also be zero. That is, the momentum of the cannon-ball, which is $5.0 \times 70 = 350\,\text{kg m s}^{-1}$ to the right, must be exactly balanced by the momentum of the cannon. If the initial speed of recoil is v, the momentum of the cannon is $1500v$ to the left. Thus, $1500v = 350$ and $v = \mathbf{0.23\,m\,s^{-1}}$.

Now it's your turn

5 An ice-skater of mass 80 kg, initially at rest, pushes his partner, of mass 65 kg, away from him so that she moves with an initial speed of $1.5\,\text{m s}^{-1}$. What is the initial speed of the first skater after this manoeuvre?

Momentum and impulse

It is now useful to introduce a quantity called **impulse** and relate it to a change in momentum.

If a constant force F acts on a body for a time Δt, the impulse of the force is given by $F\Delta t$.

The unit of impulse is given by the unit of force, the newton, multiplied by the unit of time, the second: it is the newton second (N s).

We know from Newton's second law that the force acting on a body is equal to the rate of change of momentum of the body. We have already expressed this as the equation

$$F = \Delta p/\Delta t$$

If we multiply both sides of this equation by Δt, we obtain

$$F\Delta t = \Delta p$$

We have already defined $F\Delta t$ as the impulse of the force. The right-hand side of the equation (Δp) is the change in the momentum of the body. So, Newton's second law tells us that the **impulse of a force is equal to the change in momentum**. It is useful for dealing with forces that act over a short interval of time, as in a collision. The forces between colliding bodies are seldom constant throughout the collision, but the equation can be applied to obtain information about the average force acting.

Note that the idea of impulse explains why there is an alternative unit for momentum. On page 57 we introduced the kg m s^{-1} and the N s as possible units for momentum. The kg m s^{-1} is the logical unit, the one you arrive at if you take momentum as being the product of mass and velocity. The N s comes from the impulse–momentum equation: it is the unit of impulse, and because impulse is equal to change of momentum, it is also a unit for momentum.

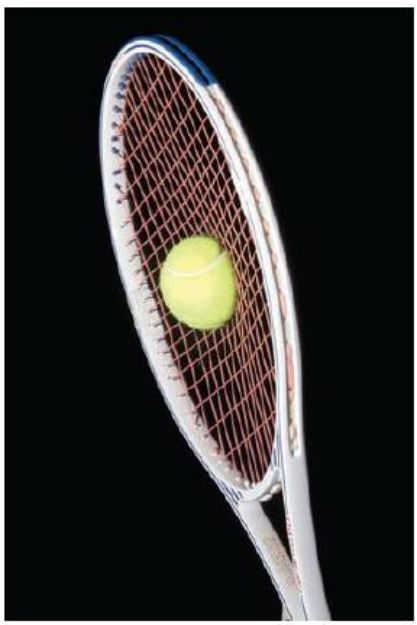


Figure 4.14

Example

Some tennis players can serve the ball at a speed of $55\,\text{m s}^{-1}$. The tennis ball has a mass of 60 g. In an experiment, it is determined that the ball is in contact with the racket for 25 ms during the serve (Figure 4.14). Calculate the average force exerted by the racket on the ball.

The change in momentum of the ball as a result of the serve is $0.060 \times 55 = 3.3\,\text{kg m s}^{-1}$. By the impulse–momentum equation, the change in momentum is equal to the impulse of the force. Since impulse is the product of force and time, $Ft = 3.3$ N s.

Here t is 0.025 s; thus $F = 3.3/0.025 = \mathbf{130\,N}$.

Now it's your turn

6 A golfer hits a ball of mass 45 g at a speed of $40\,\text{m s}^{-1}$ (Figure 4.15). The golf club is in contact with the ball for 3.0 ms. Calculate the average force exerted by the club on the ball.

www.iranino.co

IRANINOVIN



Figure 4.15

Elastic and inelastic collisions

In some collisions, kinetic energy is conserved as well as momentum. By the **conservation of kinetic energy**, we mean that the total kinetic energy of the colliding bodies before collision is the same as the total kinetic energy afterwards. This means that no energy is lost in the permanent deformation of the colliding bodies, or as heat and sound. There is a transformation of energy during the collision: while the colliding bodies are in contact, some of the kinetic energy is transformed into elastic potential energy, but as the bodies separate, it is transformed into kinetic energy again.

Using the same notation for the masses and speeds of the colliding particles as in the section on Collisions on page 62 (see Figure 4.16), the total kinetic energy of the particles before collision is

$$\tfrac{1}{2}m_1u_1^2 + \tfrac{1}{2}m_2u_2^2$$

The total kinetic energy afterwards is

$$\tfrac{1}{2}m_1v_1^2 + \tfrac{1}{2}m_2v_2^2$$

If the collision is **elastic**, the kinetic energy before collision is equal to the kinetic energy after collision.

$$\tfrac{1}{2}m_1u_1^2 + \tfrac{1}{2}m_2u_2^2 = \tfrac{1}{2}m_1v_1^2 + \tfrac{1}{2}m_2v_2^2$$

Note that because energy is a scalar, the directions of motion of the particles are not indicated by the signs of the various terms.

In solving problems about elastic collisions, this equation is useful because it gives another relation between masses and velocities, in addition to that obtained from the principle of conservation of momentum.

When the velocity directions are as defined in Figure 4.16, application of the two conservation conditions shows that

$$u_1 + u_2 = v_1 + v_2$$

for a perfectly elastic collision. That is, the relative speed of approach $(u_1 + u_2)$ is equal to the relative speed of separation $(v_1 + v_2)$. Note that this useful relation applies *only for a perfectly elastic collision.*

Elastic collisions occur in the collisions of atoms and molecules. We shall see in Topic 10 that one of the most important assumptions in the kinetic theory of gases is

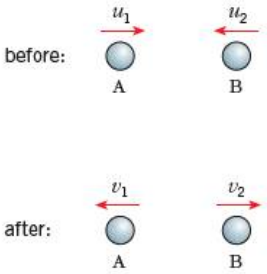


Figure 4.16 Collision between two particles

www.iranino.co

that the collisions of the gas molecules with the walls of the container are perfectly elastic. However, in larger scale collisions, such as those of snooker balls, collisions cannot be perfectly elastic. (The 'click' of snooker balls on impact indicates that a very small fraction of the total energy of the system has been transformed into sound.) Nevertheless, we often make the assumption that such a collision is perfectly elastic.

Collisions in which the total kinetic energy is not the same before and after the event are called **inelastic**.

Total energy must, of course, be conserved. But in an inelastic collision the kinetic energy that does not re-appear in the same form is transformed into heat, sound and other forms of energy. In an extreme case, all the kinetic energy may be lost. A lump of modelling clay dropped on to the floor does not bounce. All the kinetic energy it possessed just before hitting the floor has been transformed into the work done in flattening the lump, and (a much smaller amount) into the sound energy emitted as a 'squelch'.

Although kinetic energy may or may not be conserved in a collision, momentum is always conserved, and so is total energy.

The truth of this statement may not be entirely obvious, especially when considering examples such as the lump of modelling clay which was dropped on to the floor. Surely the clay had momentum just before the collision with the floor, but had no momentum afterwards? True! But for the system of the lump of clay alone, external forces (the attraction of the Earth on the clay, and the force exerted by the floor on the clay on impact) were acting. When external forces act, the conservation principle does not apply. We need to consider a system in which no external forces act. Such a system is the lump of modelling clay and the Earth. While the clay falls towards the floor, gravitational attraction will also pull the Earth towards the clay. Conservation of momentum can be applied in that the total momentum of clay and Earth remains constant throughout the process: before the collision, and after it. The effects of the transfer of the clay's momentum to the Earth are not noticeable due to the difference in mass of the two objects.

Examples

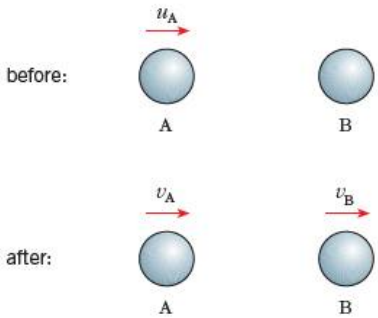


Figure 4.17

1 A snooker ball A moves with speed u_A directly towards a similar ball B which is at rest (Figure 4.17). The collision is elastic. What are the speeds v_A and v_B after the collision?

It is convenient to take the direction from left to right as the direction of positive momentum. If the mass of a billiard ball is m, the total momentum of the system before the collision is mu_A. By the principle of conservation of momentum, the total momentum after collision is the same as that before, or

$$mu_A = mv_A + mv_B$$

The collision is perfectly elastic, so the total kinetic energy before the collision is the same as that afterwards, or

$$\tfrac{1}{2}mu_A^2 = \tfrac{1}{2}mv_A^2 + \tfrac{1}{2}mv_B^2$$

Solving these equations gives $v_A = 0$ and $v_B = u_A$. That is, ball A comes to a complete standstill, and ball B moves off with the same speed as that with which ball A struck it. (Another solution is possible algebraically: $v_A = u_A$ and $v_B = 0$. This corresponds to a non-collision. Ball A is still moving with its initial speed, and ball B is still at rest. In cases where algebra gives us two possible solutions, we need to decide which one is physically appropriate.)

2 A particle of mass m makes a glancing collision with a similar particle, also of mass m, which is at rest (Figure 4.18). The collision is elastic. After the collision the particles move off at angles ϕ and β. State the equations that relate

(a) the x components of the momentum of the particles,

(b) the y components of the momentum of the particles,

(c) the kinetic energy of the particles.

www.iranino.co

IRANINOVIN

From the conservation of momentum:

(a) $mu = mv_1\cos\phi + mv_2\cos\beta$

(b) $0 = mv_1\sin\phi - mv_2\sin\beta$

(c) total kinetic energy is constant as the collision is elastic. Hence, $\frac{1}{2}mu^2 = \frac{1}{2}mv_1^2 + \frac{1}{2}mv_2^2$

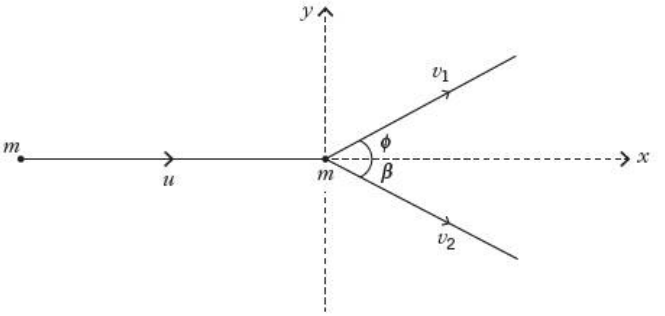


Figure 4.18

Now it's your turn

7 A trolley A moves with speed u_A towards a trolley B of equal mass which is at rest (Figure 4.19). The trolleys stick together and move off as one with speed $v_{A,B}$.

(a) Determine $v_{A,B}$.

(b) What fraction of the initial kinetic energy of trolley A is converted into other forms in this inelastic collision?

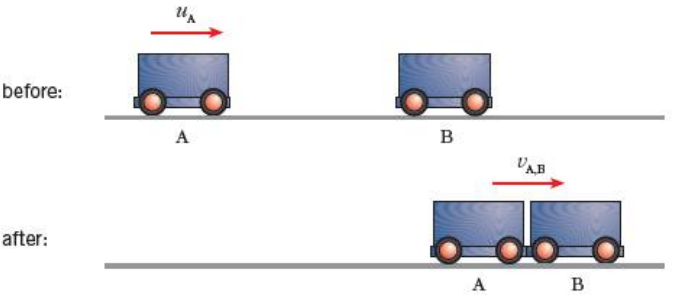


Figure 4.19

Summary

- The force of friction opposes motion.
- The linear momentum p of a body is defined as the product of its mass m and its velocity v. In symbols: $p = mv$. Momentum has units kg m s^{-1} or N s. It is a vector quantity.
- Newton's laws of motion are:
  - — First law: Every body continues in its state of rest, or with uniform velocity, unless acted upon by a resultant force.
  - — Second law: The resultant force acting on a body is proportional to the rate of change of its momentum (this is used to define force). In symbols: $F \propto \Delta p/\Delta t$. If SI units are used $F = \Delta p/\Delta t$
  - — Third law: When one body exerts a force on another body, the second body exerts an equal and opposite force on the first body.
- Newton's first and third laws of motion can also be stated in terms of momentum:
  - — First law: The momentum of a body remains constant unless an external force acts on the body: p = constant
  - — Third law: When two bodies exert action and reaction forces on each other, their changes of momentum are equal and opposite.
- If the mass is constant, the resultant force is equal to mass × acceleration or $F = ma$, where force F is in newtons, mass m is in kilograms and acceleration a is in m s^{-2}.

www.iranino.co

IRANINOVIN

- The acceleration of free fall g provides the link between the mass m and the weight W of a body: $W = mg$
- The principle of conservation of momentum states that the total momentum of an isolated system is constant. An isolated system is one on which no external resultant force acts.
- In collisions between bodies, application of the principle of conservation of momentum shows that the total momentum of the system before the collision is equal to the total momentum after the collision.
- An elastic collision is one in which the total kinetic energy remains the constant. In this situation, the relative speed of approach is equal to the relative speed of separation.
- An inelastic collision is one in which the total kinetic energy is not the same before and after the event.
- Although kinetic energy may or may not be conserved in a collision, momentum is always conserved, and so is total energy.
- The impulse of a force F is the product of the force and the time Δt for which it acts: impulse = $F\Delta t$
- The impulse of a force acting on a body is equal to the change of momentum of the body: $F\Delta t = \Delta p$
- The unit of impulse is N s.

Examination style questions

1 A net force of 95 N accelerates an object at 1.9 m s^{-2}. Calculate the mass of the object.

2 A parachute trainee jumps from a platform 3.0 m high. When he reaches the ground, he bends his knees to cushion the fall. His torso decelerates over a distance of 0.65 m. Calculate:
- a the speed of the trainee just before he reaches the ground,
- b the deceleration of his torso,
- c the average force exerted on his torso (of mass 45 kg) by his legs during the deceleration.

3 If the acceleration of a body is zero, does this mean that no forces act on it?

4 A railway engine pulls two carriages of equal mass with uniform acceleration. The tension in the coupling between the engine and the first carriage is T. Deduce the tension in the coupling between the first and second carriages.

5 Calculate the magnitude of the momentum of a car of mass 1.5 tonnes (1.5×10^3 kg) travelling at a speed of 22 m s^{-1}.

6 When a certain space rocket is taking off, the propellant gases are expelled at a rate of 900 kg s^{-1} and speed of 40 km s^{-1}. Calculate the thrust on the rocket.

7 An insect of mass 4.5 mg, flying with a speed of 0.12 m s^{-1}, encounters a spider's web, which brings it to rest in 2.0 ms. Calculate the average force exerted by the insect on the web.

8 What is your mass? What is your weight?

9 An atomic nucleus at rest emits an α-particle of mass 4 u. The speed of the α-particle is found to be 5.6×10^6 m s^{-1}. Calculate the speed with which the daughter nucleus, of mass 218 u, recoils.

10 A heavy particle of mass m_1, moving with speed u, makes a head-on collision with a light particle of mass m_2, which is initially at rest. The collision is perfectly elastic, and m_2 is very much less than m_1. Describe the motion of the particles after the collision.

11 A light body and a heavy body have the same momentum. Which has the greater kinetic energy?

12 A 45 g ball with speed of 12 m s^{-1} hits a wall at an angle of 30° (see Fig. 4.20). The ball rebounds with the same speed and angle. The contact time of the ball with the wall is 15 ms. Calculate:
- a the change in momentum of the ball,
- b the impulse of the ball,
- c the force exerted on the ball by the wall.

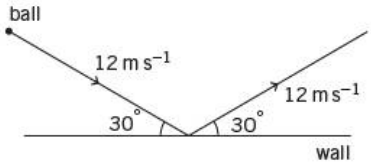


Fig. 4.20

www.iranino.co

IRANINOVIN

13 A bullet of mass 12 g is fired horizontally from a gun with a velocity of 180 m s^{-1}. It hits, and becomes embedded in, a block of wood of mass 2000 g, which is freely suspended by long strings, as shown in Fig. 4.21.

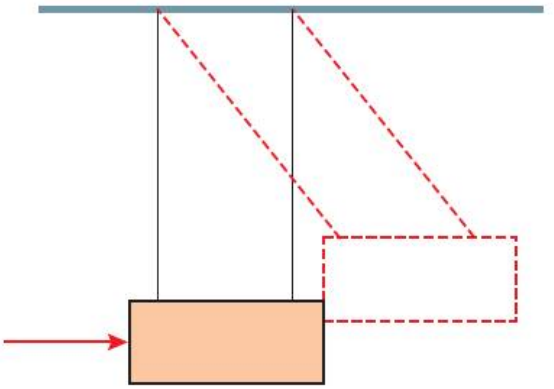


Fig. 4.21

Calculate:

a **i** the magnitude of the momentum of the bullet just before it enters the block,

ii the magnitude of the initial velocity of the block and bullet after impact,

iii the kinetic energy of the block and embedded bullet immediately after the impact.

b Deduce the maximum height above the equilibrium position to which the block and embedded bullet rise after impact.

14 A nucleus A of mass 222 u is moving at a speed of 350 m s^{-1}. While moving, it emits an α-particle of mass 4 u. After the emission, it is determined that the daughter nucleus, of mass 218 u, is moving with speed 300 m s^{-1} in the original direction of the parent nucleus. Calculate the speed of the α-particle.

15 A safety feature of modern cars is the air-bag, which, in the event of a collision, inflates and is intended to decrease the risk of serious injury. Use the concept of impulse to explain why an air-bag might have this effect.

16 Two frictionless trolleys A and B, of mass m and $3m$ respectively, are on a horizontal track (Fig. 4.22). Initially they are clipped together by a device which incorporates a spring, compressed between the trolleys. At time $t = 0$ the clip is released. The velocity of trolley B is u to the right.



Fig. 4.22

a Calculate the velocity of trolley A as the trolleys move apart.

b At time $t = t_1$, trolley A collides elastically with a fixed spring and rebounds. (Compression and expansion of the spring take a negligibly short time.) Trolley A catches up with trolley B at time $t = t_2$.

i Calculate the velocity of trolley A between $t = t_1$ and $t = t_2$.

ii Find an expression for t_2 in terms of t_1.

c When trolley A catches up with trolley B at time t_2 the clip operates so as to link them again, this time without the spring between them, so that they move together with velocity v. Calculate the common velocity v in terms of u.

d Initially the trolleys were at rest and the total momentum of the system was zero. However, the answer to **c** shows that the total momentum after $t = t_2$ is not zero. Discuss this result with reference to the principle of conservation of momentum.

17 A ball of mass m makes a perfectly elastic head-on collision with a second ball, of mass M, initially at rest. The second ball moves off with half the original speed of the first ball.

a Express M in terms of m.

b Determine the fraction of the original kinetic energy retained by the ball of mass m after the collision.

18 **a** **i** Define *force*. *[1]*

ii State Newton's third law of motion. *[3]*

b Two spheres approach one another along a line joining their centres, as illustrated in Fig. 4.23.

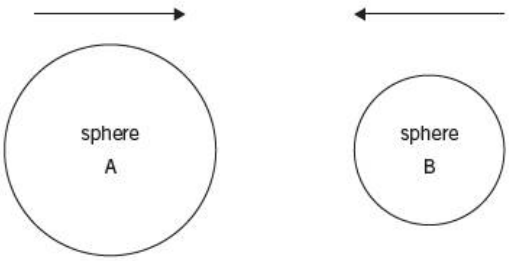


Fig. 4.23

When they collide, the average force acting on sphere A is F_A and the average force acting on sphere B is F_B.

The forces act for time t_A on sphere A and time t_B on sphere B.

i State the relationship between

1 F_A and F_B, *[1]*

2 t_A and t_B *[1]*

ii Use your answers in **i** to show that the change in momentum of sphere A is equal in magnitude and opposite in direction to the change in momentum of sphere B. *[1]*

c For the spheres in **b**, the variation with time of the momentum of sphere A before, during and after the collision with sphere B is shown in Fig. 4.24.

www.iranino.co

IRANINOVIN

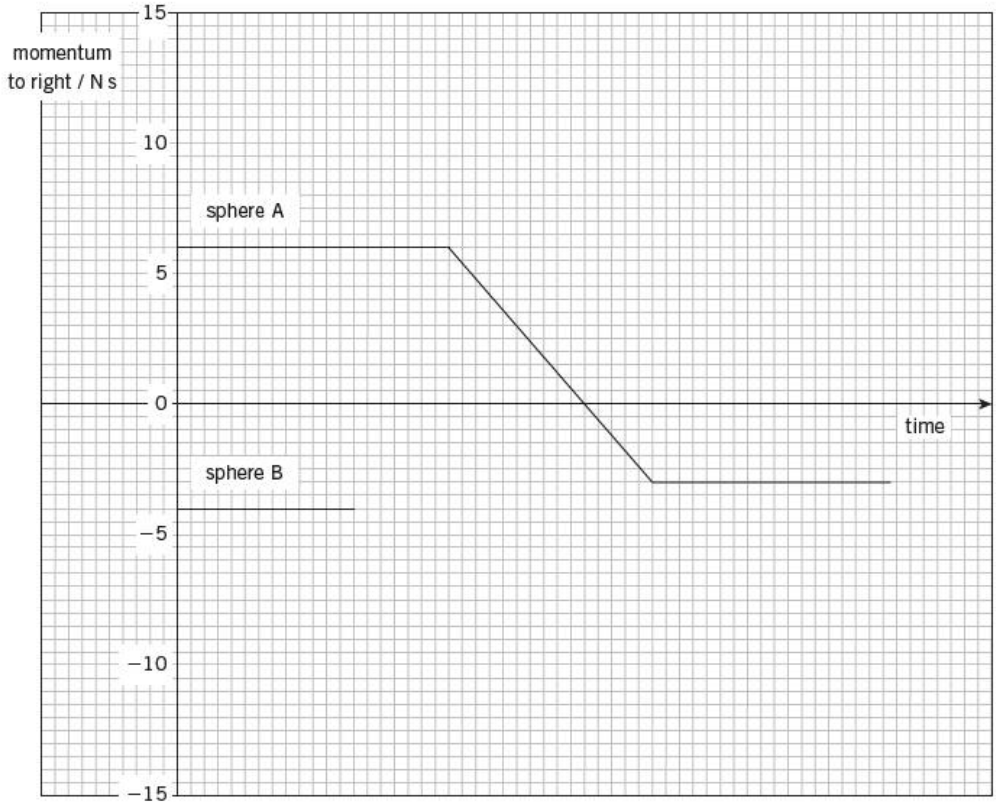


Fig. 4.24

The momentum of sphere B before the collision is also shown on Fig. 4.24.

Copy and complete Fig. 4.24 to show the variation with time of the momentum of sphere B during and after the collision with sphere A. *[3]*

Cambridge International AS and A Level Physics, 9702/22 May/June 2010 Q 3

19 A ball is thrown against a vertical wall. The path of the ball is shown in Fig. 4.25.

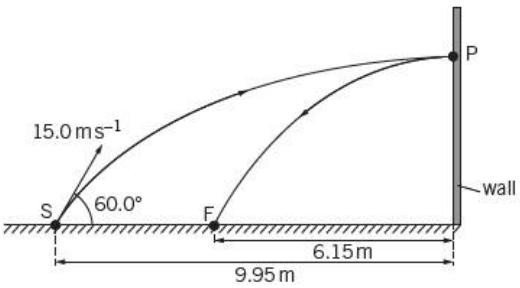


Fig. 4.25

The ball is thrown from S with an initial velocity of 15.0 m s^{-1} at 60.0° to the horizontal. Assume that air resistance is negligible.

a For the ball at S, calculate

i its horizontal component of velocity, *[1]*

ii its vertical component of velocity. *[1]*

b The horizontal distance from S to the wall is 9.95 m. The ball hits the wall at P with a velocity that is at right angles to the wall. The ball rebounds to a point F that is 6.15 m from the wall. Using your answers in **a**,

i calculate the vertical height gained by the ball when it travels from S to P, *[1]*

ii show that the time taken for the ball to travel from S to P is 1.33 s, *[1]*

iii show that the velocity of the ball immediately after rebounding from the wall is about 4.6 m s^{-1}. *[1]*

c The mass of the ball is 60 × 10^{-3} kg.

i Calculate the change in momentum of the ball as it rebounds from the wall. [2]

ii State and explain whether the collision is elastic or inelastic. *[1]*

Cambridge International AS and A level Physics, 9702/21 Oct/Nov 2011 Q 3

20 A small ball is thrown horizontally with a speed of 4.0 m s^{-1}. It falls through a vertical height of 1.96 m before bouncing off a horizontal plate, as illustrated in Fig. 4.26.

www.iranino.co

IRANINOVIN

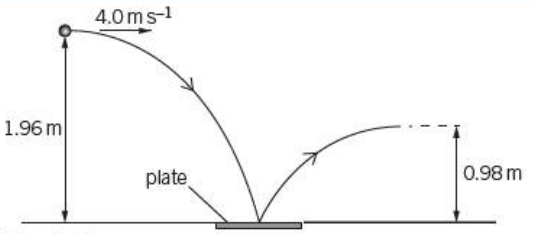


Fig. 4.26

Air resistance is negligible.

a For the ball, as it hits the horizontal plate,

i state the magnitude of the horizontal component of its velocity, *[1]*

ii show that the vertical component of the velocity is 6.2 m s^{-1}. *[1]*

b The components of the velocity in **a** are both vectors. Copy and complete Fig. 4.27 to draw a vector diagram, to scale, to determine the velocity of the ball as it hits the horizontal plate. *[3]*

Fig. 4.27

c After bouncing on the plate, the ball rises to a vertical height of 0.98 m.

i Calculate the vertical component of the velocity of the ball as it leaves the plate. *[2]*

ii The ball of mass 34 g is in contact with the plate for a time of 0.12 s.

Use your answer in **c i** and the data in **a ii** to calculate, for the ball as it bounces on the plate,

1 the change in momentum, *[3]*

2 the magnitude of the average force exerted by the plate on the ball due to this momentum change. *[2]*

Cambridge International AS and A level Physics, 9702/22 Oct/Nov 2009 Q 3

www.iranino.co

IRANINOVIN

5 Forces, density and pressure

By the end of this topic, you will be able to:

5.1 **(a)** describe the force on a mass in a uniform gravitational field and on a charge in a uniform electric field
(b) understand the origin of the upthrust acting on a body in a fluid
(c) show a qualitative understanding of frictional forces and viscous forces including air resistance
(d) understand that the weight of a body may be taken as acting at a single point known as the centre of gravity

5.2 **(a)** define and apply the moment of a force
(b) understand that a couple is a pair of forces that tends to produce rotation only
(c) define and apply the torque of a couple

5.3 **(a)** state and apply the principle of moments
(b) understand that, when there is no resultant force and no resultant torque, a system is in equilibrium
(c) use a vector triangle to represent coplanar forces in equilibrium

5.4 **(a)** define and use density
(b) define and use pressure
(c) derive, from the definitions of pressure and density, the equation $\Delta p = \rho g \Delta h$
(d) use the equation $\Delta p = \rho g \Delta h$

Starting points

- Understand the concept of weight as the effect of a gravitational field.
- The use of vector triangles to add vectors.
- For zero resultant force, the velocity of a body does not change (Newton's first law).

5.1 Types of force

The **weight** of a body is an example of the force acting on a mass in what is called a **field of force**. Near the surface of the Earth, the gravitational field is approximately constant and uniform. This means that in calculations we can take the same value of g, whatever the position on the surface of the Earth or for a short distance (compared with the Earth's radius) above it.

There are other sorts of fields of force. An important example is an electric field. An electric charge experiences a force in an electric field. There are significant similarities between the behaviour of a mass in a gravitational field and an electric charge in an electric field. We shall explore these similarities in Topic 17.

Frictional forces are important in considering the motion of a body (see Topic 4). We use the term **viscous force** to describe the frictional force in a fluid (a liquid or a gas). The property of the fluid determining this force is the **viscosity** of the fluid. An example of such a force is air resistance. In Topic 4 we considered the fact that parachutists eventually fall with a constant, terminal velocity because of air resistance.

When an object is immersed in a fluid, it appears to weigh less than when in a vacuum. It is easier to lift large stones under water than when they are out of the water. The reason for this is that immersion in the fluid provides an **upthrust** or buoyancy force.

5.2 Moment of a force

When a force acts on an object, the force may cause the object to move in a straight line. It could also cause the object to turn or spin (rotate).

www.iranino.co

IRANINOVIN

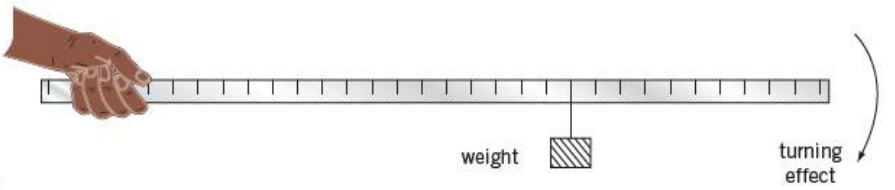


Figure 5.1 Turning effect on a metre rule

Think about a metre rule held in the hand at one end so that the rule is horizontal (Figure 5.1). If a weight is hung from the ruler we can feel a turning effect on the ruler. The turning effect increases if the weight is increased or it is moved further from the hand along the ruler. The turning effect acts at the hand where the metre rule is pivoted. Keeping the weight and its distance along the rule constant, the turning effect can be changed by holding the ruler at an angle to the horizontal. The turning effect becomes smaller as the rule approaches the vertical position.

The turning effect of a force is called the **moment** of the force.

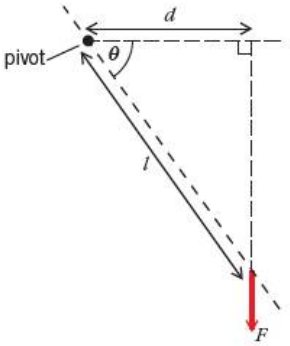


Figure 5.2 Finding the moment of a force

The moment of a force depends on the magnitude of the force and also on the distance of the force from the pivot or fulcrum. This distance must be defined precisely. In the simple experiment above, we saw that the moment of the force depended on the angle of the ruler to the horizontal. Varying this angle meant that the line of action of the force from the pivot varied (see Figure 5.2). The distance required when finding the moment of a force is the perpendicular distance d of the line of action of the force from the pivot.

The moment of a force is defined as the product of the force and the perpendicular distance of the line of action of the force from the pivot.

Referring to Figure 5.2, the force has magnitude F and acts at a point distance l from the pivot. Then, when the ruler is at angle θ to the horizontal,

$$\textit{moment of force} = F \times d$$

$$= F \times l \cos \theta$$

Since force is measured in newtons and distance is measured in metres, the unit of the moment of a force is newton-metre (N m).

Example

In Figure 5.3, a light rod AB of length 45 cm is held at A so that the rod makes an angle of 65° to the vertical. A vertical force of 15 N acts on the rod at B. Calculate the moment of the force about the end A.

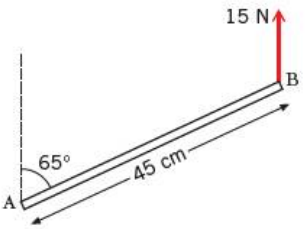


Figure 5.3

moment of force = force × perpendicular distance from pivot

= 15 × 0.45 sin 65

(Remember that the distance must be in metres.)

= 6.1 N m

Now it's your turn

1 Referring to Figure 5.3, calculate the moment of the force about A for a vertical force of 25 N with the rod at an angle of 30° to the vertical.

Couples

When a screwdriver is used, we apply a turning effect to the handle. We do not apply one force to the handle because this would mean the screwdriver would move sideways. Rather, we apply two forces of equal size but opposite direction on opposite sides of the handle (see Figure 5.4).

www.iranino.co

IRANINOVIN

A **couple** consists of two forces, equal in magnitude but opposite in direction whose lines of action do not coincide.

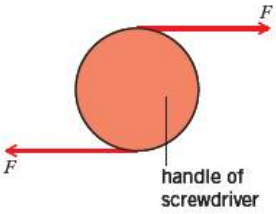


Figure 5.4 Two forces acting as a couple

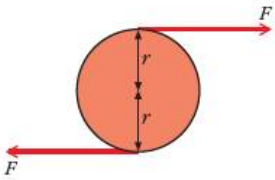


Figure 5.5 Torque of a couple

Consider the two parallel forces, each of magnitude F acting as shown in Figure 5.5 on opposite ends of a diameter of a disc of radius r. Each force produces a moment about the centre of the disc of magnitude Fr in a clockwise direction. The total moment about the centre is $F \times 2r$ or $F \times$ *perpendicular distance between the forces.*

Although a turning effect is produced, this turning effect is not called a moment because it is produced by two forces, not one. Instead, this turning effect is referred to as a **torque**. The unit of torque is the same as that of the moment of a force, i.e. newton-metre.

The torque of a couple is the product of one of the forces and the perpendicular distance between the forces.



Figure 5.6 Tightening a wheel nut requires the application of a torque.

It is interesting to note that, in engineering, the tightness of nuts and bolts is often stated as the maximum torque to be used when screwing up the nut on the bolt. Spanners used for this purpose are called torque wrenches because they have a scale on them to indicate the torque that is being applied.

Example

Calculate the torque produced by two forces, each of magnitude 30 N, acting in opposite directions with their lines of action separated by a distance of 25 cm.

torque = *force* × *separation of forces*

= 30 × 0.25 (distance in metres)

= **7.5 N m**

Now it's your turn

2 The torque produced by a person using a screwdriver is 0.18 N m. This torque is applied to the handle of diameter 4.0 cm. Calculate the force applied to the handle.

5.3 Equilibrium of forces

The principle of moments

A metre rule may be balanced on a pivot so that the rule is horizontal. Hanging a weight on the rule will make the rule rotate about the pivot. Moving the weight to the other side of the pivot will make the rule rotate in the opposite direction. Hanging weights on both sides of the pivot as shown in Figure 5.7 means that the ruler may rotate clockwise, or anticlockwise, or it may remain horizontal. In this horizontal position, there is no resultant turning effect and so the total turning effect of the forces in the clockwise direction equals the total turning effect in the anticlockwise direction. You can check this very easily with the apparatus of Figure 5.7.

www.iranino.co

IRANINOVIN

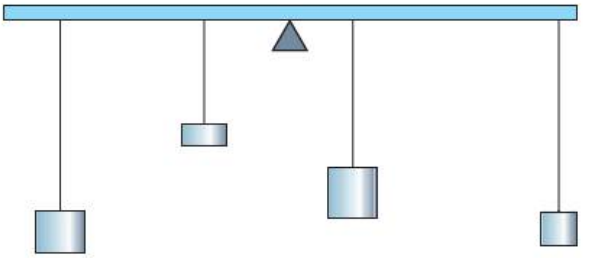


Figure 5.7

When a body has no tendency to change its speed of rotation, it is said to be in **rotational equilibrium**.

The **principle of moments** states that, for a body to be in rotational equilibrium, the sum of the clockwise moments about any point must equal the sum of the anticlockwise moments about that same point.

Example

Some weights are hung from a light rod AB as shown in Figure 5.8. The rod is pivoted. Calculate the magnitude of the force F required to balance the rod horizontally.

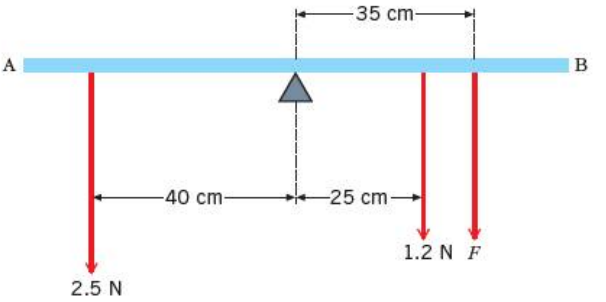


Figure 5.8

Sum of clockwise moments $= (0.25 \times 1.2) + 0.35F$

Sum of anticlockwise moments $= 0.40 \times 2.5$

By the principle of moments

$$(0.25 \times 1.2) + 0.35F = 0.40 \times 2.5$$

$$0.35F = 1.0 - 0.3$$

$$F = \mathbf{2.0N}$$

Now it's your turn

3 Some weights are hung from a light rod AB as shown in Figure 5.9. The rod is pivoted. Calculate the magnitude of the force F required to balance the rod horizontally.

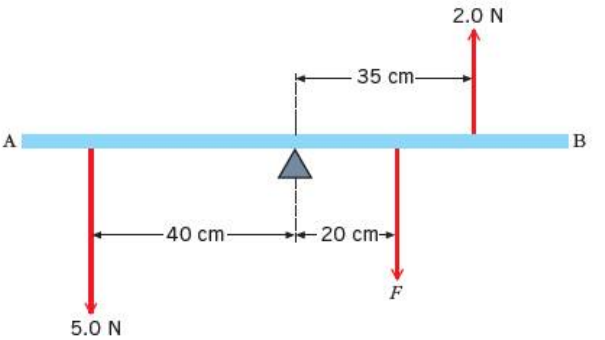


Figure 5.9

www.iranino.co

IRANINOVIN

Centre of gravity

An object may be made to balance at a particular point. When it is balanced at this point, the object does not turn and so all the weight on one side of the pivot is balanced by the weight on the other side. Supporting the object at the pivot means that the only force which has to be applied at the pivot is one to stop the object falling – that is, a force equal to the weight of the object. Although all parts of the object have weight, the whole weight of the object appears to act at this balance point. This point is called the **centre of gravity** (C.G.) of the object.

The centre of gravity of an object is the point at which the whole weight of the object may be considered to act.

The weight of a body can be shown as a force acting vertically downwards at the centre of gravity. For a uniform body such as a ruler, the centre of gravity is at the geometrical centre.

Equilibrium

The principle of moments gives the condition necessary for a body to be in rotational equilibrium. However, the body could still have a resultant force acting on it which would cause it to accelerate linearly. Thus, for complete equilibrium, there cannot be any resultant force in any direction.

In Topic 1 we added forces (vectors) using a vector triangle. When three forces act on an object the condition for equilibrium is that the vector diagram for these forces forms a closed triangle. When four or more forces act on an object the same principles apply. For equilibrium, the closed vector triangle then becomes a closed vector polygon.

For a body to be **in equilibrium**:
1 The sum of the forces in any direction must be zero.
2 The sum of the moments of the forces about any point must be zero.

Example

The uniform rod PQ shown in Figure 5.10 is horizontal and in equilibrium.

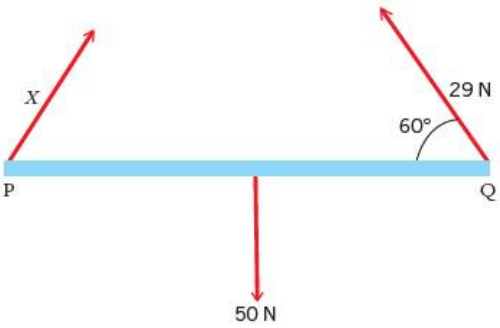


Figure 5.10

The weight of the rod is 50 N. A force of 29 N that acts at end Q is 60° to the horizontal. The force at end P is labelled X. Draw a vector triangle to represent the forces acting on the rod and determine the magnitude and direction of force X.

The forces keep the rod in equilibrium and hence form a closed triangle as shown in Figure 5.11.

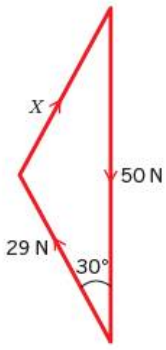


Figure 5.11

A scale diagram can be drawn to show that X is 29 N and acts at 60° to the horizontal.

Now it's your turn

4 The same uniform rod PQ is in equilibrium, as in the above example.
(a) (i) Show that the upward forces equal the downward forces.
(ii) Show that the horizontal force to the left equals the horizontal force to the right.
(b) The length of the rod in Figure 5.10 is 100 cm. Determine the force X by taking moments about Q.

www.iranino.co

IRANINOVIN

5.4 Density and pressure

In this section we will bring together density and pressure to show an important link between them.

The density of a substance is defined as its mass per unit volume.
$\rho = m/V$

The symbol for density is ρ (Greek rho) and its SI unit is kg m^{-3}.

Example

An iron sphere of radius 0.18 m has mass 190 kg. Calculate the density of iron.

First calculate the volume of the sphere from $V = \frac{4}{3}\pi r^3$. This works out at 0.024 m^3.
Application of the formula for density gives $\rho =$ **7800 kg m^{-3}**.

Pressure is defined as force per unit area, where the force F acts perpendicularly to the area A.
$p = F/A$

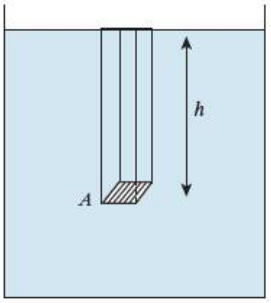


Figure 5.12 Column of liquid above the area A

The symbol for pressure is p and its SI unit is the pascal (Pa), which is equal to one newton per square metre (N m^{-2}).

The link between pressure and density comes when we deal with liquids or with fluids in general. Consider a point at a depth h below the surface of a liquid in a container. What is the pressure due to the liquid? Very simply, the pressure is caused by the weight of the column of liquid above a small area at that depth, as shown in Figure 5.12. The weight of the column is $W = mg = \rho Ahg$, and the pressure is $W/A = \rho gh$.

$p = \rho gh$

The pressure is proportional to the depth below the surface of the liquid. If an external pressure, such as atmospheric pressure, acts on the surface of the liquid, this must be taken into account in calculating the absolute pressure. The absolute pressure is the sum of the external pressure and the pressure due to the depth below the surface of the liquid.

Example

Calculate the excess pressure over atmospheric at a point 1.2 m below the surface of the water in a swimming pool. The density of water is 1.0 × 10^3 kg m^{-3}.

This is a straightforward calculation from $p = \rho gh$.
Substituting, $p = 1.0 \times 10^3 \times 9.8 \times 1.2 =$ **1.2 × 10^4 Pa**.

If the total pressure had been required, this value would be added to atmospheric pressure p_A. Taking p_A to be 1.01 × 10^5 Pa, the total pressure is 1.13 × 10^5 Pa.

Now it's your turn

5 Calculate the difference in blood pressure between the top of the head and the soles of the feet of a student 1.3 m tall, standing upright. Take the density of blood as 1.1 × 10^3 kg m^{-3}.

Upthrust

When an object is immersed in a fluid (a liquid or a gas), it appears to weigh less than when in a vacuum. It is easier to lift large stones under water than when they are out of the water. The reason for this is that immersion in the fluid provides an **upthrust** or buoyancy force.

We can see the reason for the upthrust when we think about an object, such as the cylinder in Figure 5.13, in water. Remember that the pressure in a liquid increases with depth. Thus, the pressure at the bottom of the cylinder is greater than the pressure at the top of the cylinder. This means that there is a bigger force acting upwards on the

www.iranino.co

IRANINOVIN

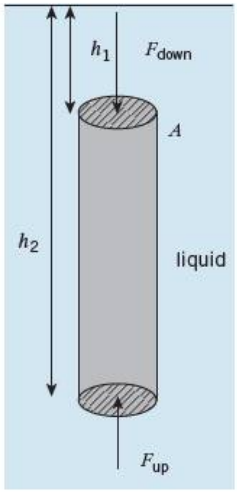


Figure 5.13 Origin of the buoyancy force (upthrust)

base of the cylinder, than there is acting downwards on the top. The difference in these forces is the upthrust or buoyancy force F_b. Looking at Figure 5.13, we can see that

$$F_b = F_{up} - F_{down}$$

and, since

$$p = \rho gh = F/A$$
$$F_b = \rho gA(h_2 - h_1) = \rho gAl$$
$$= \rho gV$$

where l is the length of the cylinder, and V is its volume. The upthrust is simply the **weight of the liquid displaced** by the immersed object. This relation has been derived for a cylinder, but it will also apply to objects of any shape.

The rule that the upthrust acting on a body immersed in a fluid is equal to the weight of the fluid displaced is known as **Archimedes' principle**.

Example

Calculate **(a)** the force needed to lift a metal cylinder when in air and **(b)** the force needed to lift the cylinder when immersed in water.

The density of the metal is 7800 kg m^{-3} and the density of water is 1000 kg m^{-3}. The volume of the cylinder is 0.50 m^3.

(a) force needed in air = weight of cylinder = 0.50 × 7800 × 9.81 = **3.8 × 10^4 N**

(b) force needed in water = weight of cylinder – upthrust
= 0.50 × 7800 × 9.81 – 0.50 × 1000 × 9.81 = **3.3 × 10^4 N**

The difference in the values in **(a)** and **(b)** is the upthrust on the metal cylinder when immersed in water.

[The upthrust of the cylinder in air was neglected as the density of air is very much less than that of the metal.]

Now it's your turn

6 Explain why a boat made of metal is in equilibrium when stationary and floating on water.

Summary

- The moment of a force is a measure of the turning effect of the force.
- The moment of a force is the product of the force and the perpendicular distance of the line of action of the force from the pivot.
- A couple consists of two equal forces acting in opposite directions whose lines of action do not coincide.
- The torque of a couple is a measure of the turning effect of the couple.
- The torque of a couple is the product of one of the forces and the perpendicular distance between the lines of action of the forces.
- The principle of moments states that the sum of the clockwise moments about a point is equal to the sum of the anticlockwise moments about the point.
- The centre of gravity of a body is the point at which the whole weight of the body may be considered to act.
- For a body to be in equilibrium:
  — the sum of the forces in any direction must be zero,
  — the sum of the moments of the forces about any point must be zero.
- Density ρ is defined by the equation $\rho = m/V$, where m is the mass of an object and V is its volume.
- Pressure p is defined by the equation $p = F/A$, where F is the force acting perpendicularly to an area A.
- The total pressure p at a point at a depth h below the surface of a fluid of density ρ is $p = p_A + \rho gh$, p_A being the atmospheric pressure; the difference in pressure between the surface and a point at a depth h is ρgh.
- The upthrust on a body immersed in a fluid is equal to the weight of the fluid displaced.

www.iranino.co

IRANINOVIN

Examination style questions

1 A uniform rod of length 60 cm has a weight of 14 N. It is pivoted at one end and held in a horizontal position by a thread tied to its other end, as shown in Fig. 5.14. The thread makes an angle of 50° with the horizontal. Calculate:

a the moment of the weight of the rod about the pivot,
b the tension T in the thread required to hold the rod horizontally.

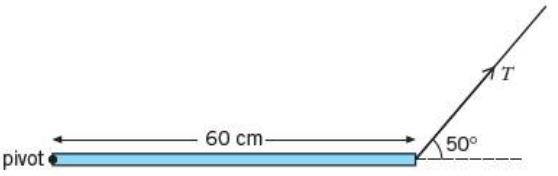


Fig. 5.14

2 A ruler is pivoted at its centre of gravity and weights are hung from the ruler as shown in Fig. 5.15. Calculate:

a the total anticlockwise moment about the pivot,
b the magnitude of the force F.

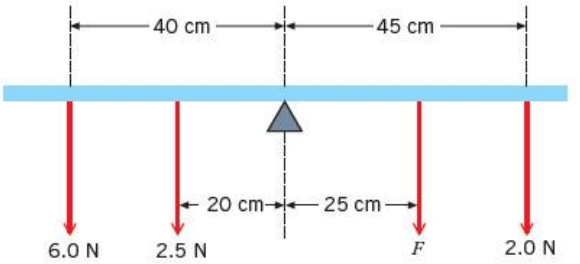


Fig. 5.15

3 A uniform plank of weight 120 N rests on two stools as shown in Fig. 5.16. A weight of 80 N is placed on the plank, midway between the stools. Calculate:

a the force acting on the stool at A,
b the force acting on the stool at B.

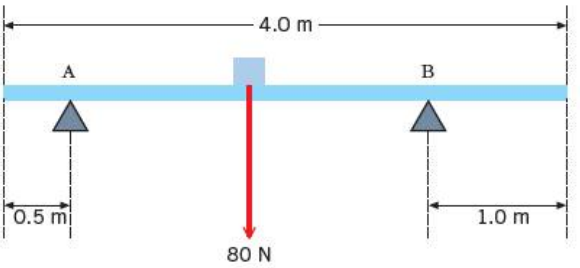


Fig. 5.16

4 A nut is to be tightened to a torque of 16 N m. Calculate the force which must be applied to the end of a spanner of length 24 cm in order to produce this torque.

5 The water in a storage tank is 15 m above a water tap in the kitchen of a house. Calculate the pressure of the water leaving the tap. Density of water = 1.0×10^3 kg m^{-3}.

6 Show that the pressure p due to a liquid of density ρ is proportional to the depth h below the surface of the liquid.

7 a Define *centre of gravity*. [2]
b A uniform rod AB is attached to a vertical wall at A. The rod is held horizontally by a string attached at B and to point C, as shown in Fig. 5.17.

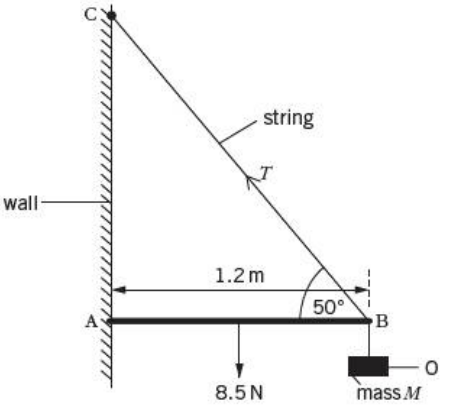


Fig. 5.17

The angle between the rod and the string at B is 50°. The rod has length 1.2 m and weight 8.5 N. An object O of mass M is hung from the rod at B. The tension T in the string is 30 N.

i Use the resolution of forces to calculate the vertical component of T. [1]
ii State the *principle of moments*. [1]
iii Use the principle of moments and take moments about A to show that the weight of the object O is 19 N. [3]
iv Hence determine the mass M of the object O. [1]

c Use the concept of equilibrium to explain why a force must act on the rod at A. [2]

Cambridge International AS and A level Physics, 9702/22 May/June 2013 Q 3

8 a Explain what is meant by *centre of gravity*. [2]
b Define *moment* of a force. [1]
c A student is being weighed. The student, of weight W, stands 0.30 m from end A of a uniform plank AB, as shown in Fig. 5.18.

www.iranino.co

IRANINOVIN

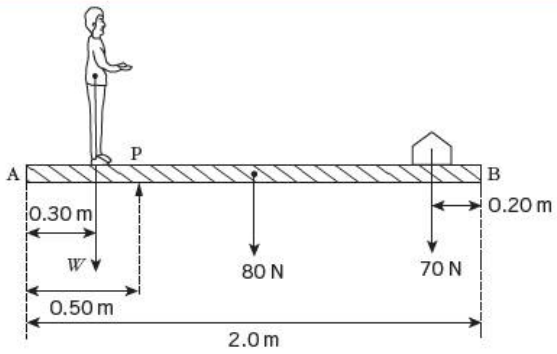


Fig. 5.18

The plank has weight 80 N and length 2.0 m. A pivot P supports the plank and is 0.50 m from end A.

A weight of 70 N is moved to balance the weight of the student. The plank is in equilibrium when the weight is 0.20 m from end B.

i State the two conditions necessary for the plank to be in equilibrium. *[2]*

ii Determine the weight W of the student. *[3]*

iii If only the 70 N weight is moved, there is a maximum weight of student that can be determined using the arrangement shown in Fig. 5.18. State and explain **one** change that can be made to increase this maximum weight. *[2]*

Cambridge International AS and A Level Physics, 9702/21 May/June 2011 Q 3

9 a Distinguish between the moment of a force and the torque of a couple. *[4]*

b One type of weighing machine, known as a steelyard, is illustrated in Fig. 5.19.

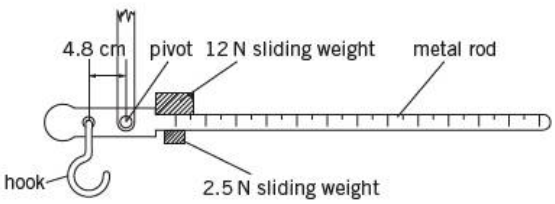


Fig. 5.19

The two sliding weights can be moved independently along the rod.

With no load on the hook and the sliding weights at the zero mark on the metal rod, the metal rod is horizontal. The hook is 4.8 cm from the pivot.

A sack of flour is suspended from the hook. In order to return the metal rod to the horizontal position, the 12 N sliding weight is moved 84 cm along the rod and the 2.5 N weight is moved 72 cm.

i Calculate the weight of the sack of flour. *[2]*

ii Suggest why this steelyard would be imprecise when weighing objects with a weight of about 25 N. *[1]*

Cambridge International AS and A level Physics, 9702/02 Oct/Nov 2008 Q 3

10 a Define the *torque* of a couple. *[2]*

b A uniform rod of length 1.5 m and weight 2.4 N is shown in Fig. 5.20.

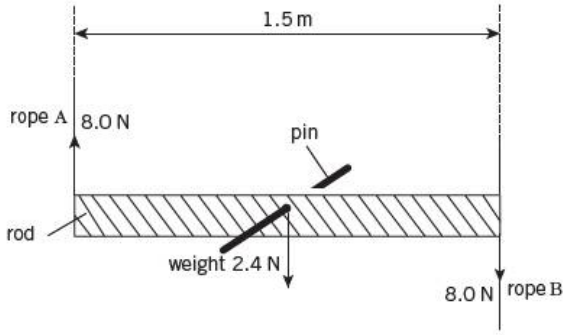


Fig. 5.20

The rod is supported on a pin passing through a hole in its centre. Ropes A and B provide equal and opposite forces of 8.0 N.

i Calculate the torque on the rod produced by ropes A and B. *[1]*

ii Discuss, briefly, whether the rod is in equilibrium. *[2]*

c The rod in **b** is removed from the pin and supported by ropes A and B, as shown in Fig. 5.21.

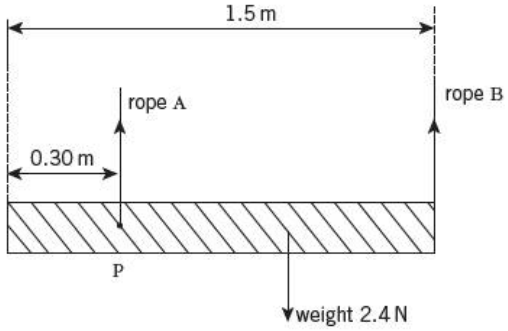


Fig. 5.21

Rope A is now at point P 0.30 m from one end of the rod and rope B is at the other end.

i Calculate the tension in rope B. *[2]*

ii Calculate the tension in rope A. *[1]*

Cambridge International AS and A Level Physics, 9702/21 Oct/Nov 2011 Q 2

11 a Define *density*. *[1]*

c A paving slab has a mass of 68 kg and dimensions 50 mm × 600 mm × 900 mm.

i Calculate the density, in kg m^{-3}, of the material from which the paving slab is made. *[2]*

ii Calculate the maximum pressure a slab could exert on the ground when resting on one of its surfaces. *[3]*

Cambridge International AS and A level Physics, 9702/21 Oct/Nov 2011 Q 1 parts a and c

www.iranino.co

IRANINOVIN

6 Work, energy, power

By the end of this topic, you will be able to:

6.1 (a) give examples of energy in different forms, its conversion and conservation and apply the principle of conservation of energy to simple examples

6.2 (a) understand the concept of work in terms of the product of a force and displacement in the direction of the force

(b) calculate the work done in a number of situations, including the work done by a gas which is expanding against a constant external pressure: $W = p\Delta V$

(c) recall and understand that the efficiency of a system is the ratio of useful energy output from the system to the total energy input

(d) show an appreciation for the implications of energy losses in practical devices and use the concept of efficiency to solve problems

6.3 (a) derive, from the equations of motion, the formula kinetic energy $E_k = \frac{1}{2}mv^2$

(b) recall and apply the formula $E_k = \frac{1}{2}mv^2$

(c) distinguish between gravitational potential energy and elastic potential energy

(d) understand and use the relationship between force and potential energy in a uniform field to solve problems

(e) derive, from the defining equation $W = Fs$, the formula $\Delta E_p = mg\Delta h$ for potential energy changes near the Earth's surface

(f) recall and use the formula $\Delta E_p = mg\Delta h$ for potential energy changes near the Earth's surface

6.4 (a) define power as work done per unit time and derive power as the product of force and velocity

(b) solve problems using the relationships $P = \frac{W}{t}$ and $P = Fv$

Starting points

- Know that there are various forms of energy.
- Understand that energy can be converted from one form to another.
- Machines enable us to do useful work by converting energy from one form to another.

'I'm going to work today.'

'Where do you work?'

'I've done some work in the garden.'

'Lots of work was done lifting the box.'

'I've done my homework.'



Figure 6.1 The weight-lifter uses a lot of energy to lift the weights but they can be rolled along the ground with little effort.

www.iranino.ir

6.2 Work

The words 'work', 'energy' and 'power' are in use in everyday English language but they have a variety of meanings. In physics, they have very precise meanings. The word **work** has a definite interpretation. The vagueness of the term 'work' in everyday speech causes problems for some students when they come to give a precise scientific definition of work.

Work is done when a force moves the point at which it acts (the point of application) in the direction of the force.

work done = force × distance moved by the force in the direction of the force

It is very important to include direction in the definition of work done. A car can be pushed horizontally quite easily but, if the car is to be lifted off its wheels, much more work has to be done and a machine, such as a car-jack, is used.

When a force moves its point of application in the direction of the force, the force does work and the work done by the force is said to be *posttive*. Conversely, if the

IRANINOVIN

direction of the force is *opposite* to the direction of movement, work is done on the force. This work done is then said to be *negative*. This is illustrated in Figure 6.2.

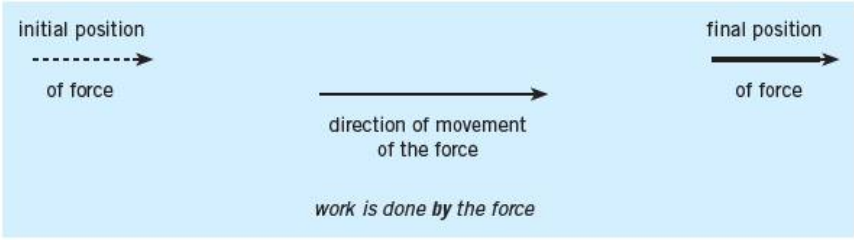


Figure 6.2

An alternative name for distance moved in a particular direction is **displacement**. Displacement is a vector quantity, as is force. However, work done has no direction, only magnitude (size), and is a scalar quantity. It is measured in joules (J).

When a force of one newton moves its point of application by one metre in the direction of the force, one joule of work is done.

work done in joules = force in newtons × distance moved in metres in the direction of the force

It follows that a joule (J) may be said to be a newton-metre (N m). If the force and the displacement are not both in the same direction, then the component of the force in the direction of the displacement must be found. Consider a force F acting along a line at an angle θ to the displacement, as shown in Figure 6.4. The component of the force along the direction of the displacement is $F \cos \theta$.

work done for displacement $x = F \cos \theta \times x$

$= Fx \cos \theta$

Note that the component $F \sin \theta$ of the force is at right angles to the displacement. Since there is no displacement in the direction of this component, no work is done in that direction.



Figure 6.3 The useful work done by the small tug-boat is found using the component of the tension in the rope along the direction of motion of the ship.

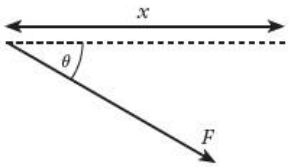


Figure 6.4

Example

A child tows a toy by means of a string as shown in Figure 6.5.

The tension in the string is 1.5 N and the string makes an angle of 25° with the horizontal. Calculate the work done in moving the toy horizontally through a distance of 265 cm.

work done = horizontal component of tension × distance moved

$= 1.5 \cos 25 \times \frac{265}{100}$

$= \mathbf{3.6\,J}$

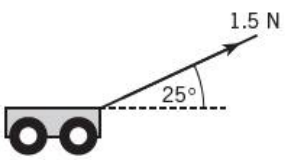


Figure 6.5

Now it's your turn

1 A box weighs 45 N. Calculate the work done in lifting the box through a vertical height of:
   (a) 4.0 m,
   (b) 67 cm.

www.iranino.co

IRANINOVIN

2 A force of 36N acts at an angle of 55° to the vertical. The force moves its point of application by 64cm in the direction of the force. Calculate the work done by:
(a) the horizontal component of the force,
(b) the vertical component of the force.

Work done by an expanding gas

A building can be demolished with explosives (Figure 6.6). When the explosives are detonated, large quantities of gas at high pressure are produced. As the gas expands, it does work by breaking down the masonry. In this section, we will derive an equation for the work done when a gas changes its volume.

Consider a gas contained in a cylinder by means of a frictionless piston of area A, as shown in Figure 6.7. The pressure p of the gas in the cylinder is equal to the atmospheric pressure outside the cylinder. This pressure may be thought to be constant.



Figure 6.6 Explosives produce large quantities of high-pressure gas. When the gas expands, it does work in demolishing the building

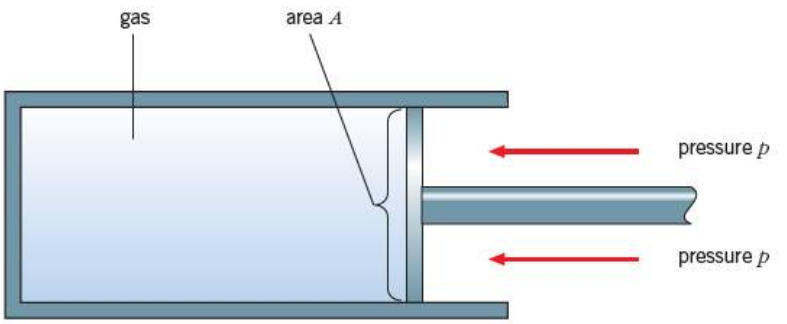


Figure 6.7

Since $pressure = \frac{force}{area}$, the gas produces a force F on the piston given by

$F = pA$

When the gas expands at constant pressure, the piston moves outwards through a distance x. So,

work done by the gas = force × distance moved

$$W = pAx$$

However, Ax is the increase in volume of the gas ΔV. Hence,

$W = p\Delta V$

When the volume of a gas changes at constant pressure,

work done = pressure × change in volume

When the gas *expands*, work is done *by* the gas. If the gas *contracts*, then work is done *on* the gas.

www.iranino.co

IRANINOVIN

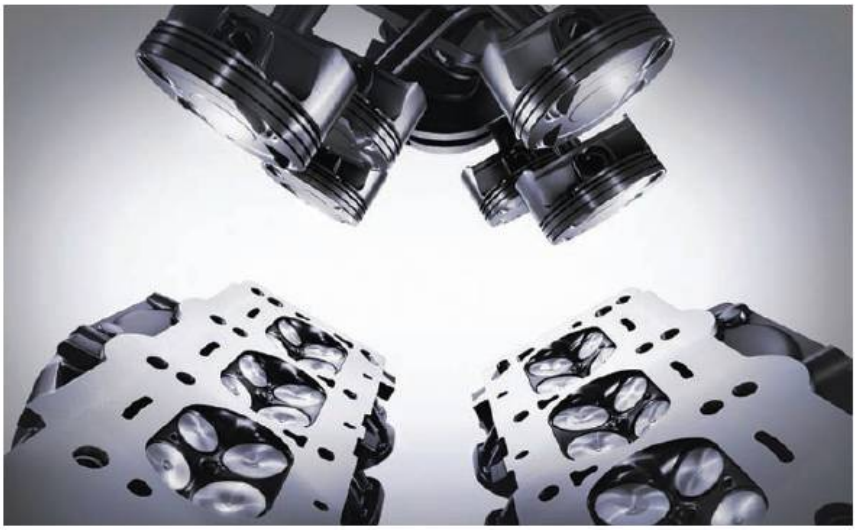


Figure 6.8 It is expanding gases pushing on the pistons which cause work to be done by the engine in a car.

Remember that the unit of work done is the joule (J). The pressure must be in pascals (Pa) or newtons per metre squared (N m^{-2}) and the change in volume in metres cubed (m^3).

Example

A sample of gas has a volume of 750 cm^3. The gas expands at a constant pressure of 1.4×10^5 Pa so that its volume becomes 900 cm^3. Calculate the work done by the gas during the expansion.

$$\begin{aligned} \textit{change in volume } \Delta V &= (900 - 750) \\ &= 150\,cm^3 \\ &= 150 \times 10^{-6}\,m^3 \\ \textit{work done by gas} &= p\Delta V \\ &= (1.4 \times 10^5) \times (150 \times 10^{-6}) \\ &= \mathbf{21\,J} \end{aligned}$$



Figure 6.9 The spring stores energy as it is stretched, releasing the energy as it returns to its original shape.

www.iranino.co

Now it's your turn

3 The volume of air in a tyre is 9.0×10^{-3} m^3. Atmospheric pressure is 1.0×10^5 Pa. Calculate the work done against the atmosphere by the air when the tyre bursts and the air expands to a volume of 2.7×10^{-2} m^3.

4 High-pressure gas in a spray-can has a volume of 250 cm^3. The gas escapes into the atmosphere through a nozzle, so that its final volume is four times the volume of the can. Calculate the work done by the gas, given that atmospheric pressure is 1.0×10^5 Pa.

6.1 Energy

In order to wind up a spring, work has to be done because a force must be moved through a distance. When the spring is released, it can do work; for example, making a child's toy move. When the spring is wound, it stores the ability to do work. Anything that is able to do work is said to have energy.

A body which can do work must have **energy**.

A body with no energy is unable to do work. Energy and work are both scalars. Since work done is measured in joules (J), energy is also measured in joules. Table 6.1 lists some typical values of energy.

IRANINOVIN

Table 6.1 Typical energy values

	order of magnitude of energy/J
radioactive decay of a nucleus	10^{-13}
sound of speech on ear for 1 second	10^{-8}
moonlight on face for 1 second	10^{-3}
beat of the heart	1
burning a match	10^{3}
large cream cake	10^{6}
energy released from 100 kg of coal	10^{10}
Earthquake	10^{19}
energy received on Earth from the Sun in one year	10^{25}
rotational energy of the Milky Way galaxy	10^{50}
estimated energy of formation of the Universe	10^{70}

Energy conversion and conservation

Newspapers sometimes refer to a 'global energy crisis'. In the near future, there may well be a shortage of fossil fuels. Fossil fuels are sources of chemical energy. It would be more accurate to refer to a 'fuel crisis'. When chemical energy is used, the energy is transformed into other forms of energy, some of which are useful and some of which are not. Eventually, all the chemical energy is likely to end up as energy that is no longer useful to us. For example, when petrol is burned in a car engine, some of the chemical energy is converted into the kinetic energy of the car and some is wasted as heat (thermal) energy. When the car stops, its kinetic energy is converted into internal energy in the brakes. The temperature of the brakes increases and heat energy is released. The outcome is that the chemical energy has been converted into heat energy which dissipates in the atmosphere and is of no further use. However, the total energy present in the Universe has remained constant. All energy changes are governed by the **law of conservation of energy**. This law states that

Energy cannot be created or destroyed. It can only be converted from one form to another.

There are many different forms of energy and you will meet a number of these during your A/AS Level Physics studies. Some of the more common forms are listed in Table 6.2.

Table 6.2 Forms of energy

energy	notes
gravitational potential energy	energy due to position of a mass in a gravitational field
kinetic energy	energy due to motion
elastic potential energy	energy stored due to stretching or compressing an object
electrical energy	energy associated with moving charge carriers due to a potential difference
electrostatic potential energy	energy due to the position of a charge in an electric field
sound energy	energy transferred from particle to particle associated with a sound wave
electromagnetic radiation	energy associated with waves in the electromagnetic spectrum
solar energy	electromagnetic radiation from the Sun
internal energy	random kinetic and potential energy of the molecules in an object
chemical energy	energy released during chemical reactions
nuclear energy	energy associated with particles in the nuclei of atoms
thermal energy	energy transferred due to temperature difference (sometimes called heat energy)

www.iranino.co

IRANINOVIN

Example

Map out the energy changes taking place when a battery is connected to a lamp.

Chemical energy in battery → electrical energy in wires → light energy and internal energy of the lamp

Now it's your turn

5 Map out the following energy changes:
(a) a child swinging on a swing,
(b) an aerosol can producing hairspray,
(c) a lump of clay thrown into the air which subsequently hits the ground.



Figure 6.10 When the mass falls, it gains kinetic energy and drives the pile into the ground.

6.3 Kinetic energy

As an object falls, it loses gravitational potential energy and, in so doing, it speeds up. Energy is associated with a moving object. In fact, we know that a moving object can be made to do work as it slows down. For example, a moving hammer hits a nail and, as it stops, does work to drive the nail into a piece of wood.

Kinetic energy is energy due to motion.

Consider an object of mass m moving with a constant acceleration a. In a distance s, the object accelerates from velocity u to velocity v. Then, by referring to the equations of motion (see Topic 3),

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

By Newton's law (see Topic 4), the force F giving rise to the acceleration a is given by

$$F = ma$$

Combining these two equations,

$$v^2 = u^2 + 2\frac{F}{m}s$$

Re-arranging,

$$mv^2 = mu^2 + 2Fs$$

$$2Fs = mv^2 - mu^2$$

$$Fs = \tfrac{1}{2}mv^2 - \tfrac{1}{2}mu^2$$

By definition, the term Fs is the work done by the force moving a distance s. Therefore, since Fs represents work done, then the other terms in the equation, $\frac{1}{2}mv^2$ and $\frac{1}{2}mu^2$, must also have the units of work done, or energy (see Topic 1). The magnitude of each of these terms depends on velocity squared and so $\frac{1}{2}mv^2$ and $\frac{1}{2}mu^2$ are terms representing energy which depends on velocity (or speed). The kinetic energy E_k of an object of mass m moving with speed v is given by

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

For the kinetic energy to be in joules, mass must be in kilograms and speed in metres per second.

The full name for the term $E_k = \frac{1}{2}mv^2$ is *translational kinetic energy* because it is energy due to an object moving in a straight line. It should be remembered that rotating objects also have kinetic energy and this form of energy is known as *rotational kinetic energy*.

www.iranino.co

IRANINOVIN

Example

Calculate the kinetic energy of a car of mass 900kg moving at a speed of 20m s^{-1}. State the form of energy from which the kinetic energy is derived.

$$\text{kinetic energy} = \tfrac{1}{2}mv^2$$
$$= \tfrac{1}{2} \times 900 \times 20^2$$
$$= \mathbf{1.8 \times 10^5\ J}$$

This energy is derived from the chemical energy of the fuel.

Now it's your turn

6 Calculate the kinetic energy of a car of mass 800kg moving at 100 kilometres per hour.

7 A cycle and cyclist have a combined mass of 80kg and are moving at 5.0m s^{-1}. Calculate:
(a) the kinetic energy of the cycle and cyclist,
(b) the increase in kinetic energy for an increase in speed of 5.0m s^{-1}.

Potential energy

Potential energy is the ability of an object to do work as a result of its position or shape.



Figure 6.11 The cars on the rollercoaster have stored gravitational potential energy. This energy is released as the cars fall.

We have already seen that a wound-up spring stores energy. This energy is potential energy because the spring is strained. More specifically, the energy may be called **elastic** (or **strain**) **potential energy**. Elastic potential energy is stored in objects which have had their shape changed elastically. Examples include stretched wires, twisted elastic bands and compressed gases.

Newton's law of gravitation (see Topic 8) tells us that all masses attract one another. We rely on the force of gravity to keep us on Earth! When two masses are pulled apart, work is done on them and so they gain **gravitational potential energy**. If the masses move closer together, they lose gravitational potential energy.

Gravitational potential energy is energy possessed by a mass due to its position in a gravitational field.

Changes in gravitational potential energy are of particular importance for an object near to the Earth's surface because we frequently do work raising masses and, conversely, the energy stored is released when the mass is lowered again. An object of mass m near the Earth's surface has weight mg, where g is the acceleration of free fall. This weight is the force with which the Earth attracts the mass (and the mass attracts the Earth). If the mass moves a *vertical* distance h,

$$\textit{work done} = \textit{force} \times \textit{distance moved}$$
$$= mgh$$

When the mass is raised, the work done is stored as *gravitational potential energy* and this energy can be recovered when the mass falls.

Change in gravitational potential energy $\Delta E_p = mg\Delta h$

It is important to remember that, for the energy to be measured in joules, the mass m must be in kilograms, the acceleration g in metres (second)$^{-2}$ and the change in height Δh in metres.

Notice that a zero point of gravitational potential energy has not been stated. We are concerned with *changes* in potential energy when a mass rises or falls.

www.iranino.co

IRANINOVIN

Examples

1 Map out the energy changes taking place when an object moves from its lowest point to its highest point on the end of a vertical spring after the spring is stretched.

(maximum) elastic potential energy in stretched spring → gravitational potential energy and kinetic energy and (reduced) elastic potential energy of object (as it moves up) → (maximum) gravitational potential energy (zero kinetic energy) and elastic potential energy in the compressed spring at its highest point

2 A shop assistant stacks a shelf with 25 tins of beans, each of mass 460 g (Figure 6.12). Each tin has to be raised through a distance of 1.8 m. Calculate the gravitational potential energy gained by the tins of beans, given that the acceleration of free fall is 9.8 m s^{-2}.

total mass raised = 25 × 460 = 11 500 g

= 11.5 kg

increase in potential energy = $m \times g \times h$

= 11.5 × 9.8 × 1.8

= **200 J**

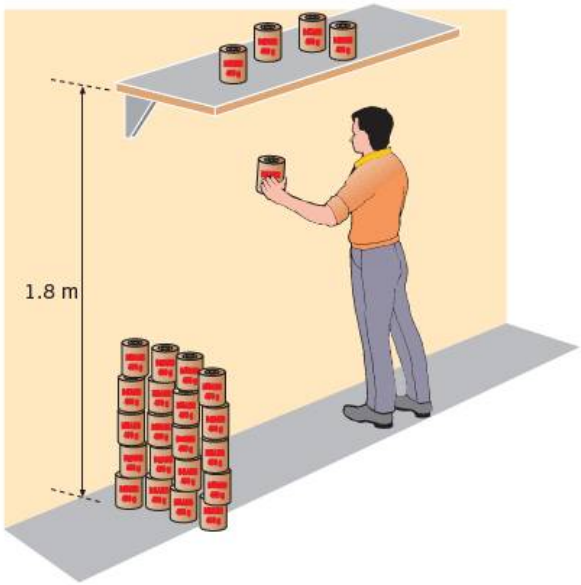


Figure 6.12

Now it's your turn

8 The acceleration of free fall is 9.8 m s^{-2}. Calculate the change in gravitational potential energy when:

(a) a person of mass 70 kg climbs a cliff of height 19 m,

(b) a book of mass 940 g is raised vertically through a distance of 130 cm,

(c) an aircraft of total mass 2.5 × 10^5 kg descends by 980 m.

Efficiency

Machines are used to change energy from one form into some other more useful form. In most energy changes some energy is 'lost' as heat (thermal) energy. For example, when a ball rolls down a slope, the total change in gravitational potential energy is not equal to the gain in kinetic energy because heat (thermal) energy has been produced as a result of frictional forces.

www.iranino.co

IRANINOVIN

Efficiency gives a measure of how much of the total energy may be considered useful and is not 'lost'.

$$efficiency = \frac{useful\ energy\ output}{total\ energy\ input}$$

Efficiency may be given either as a ratio or as a percentage. Since energy cannot be created, efficiency can never be greater than 100% and a 'perpetual motion' machine is not possible (Figure 6.13).

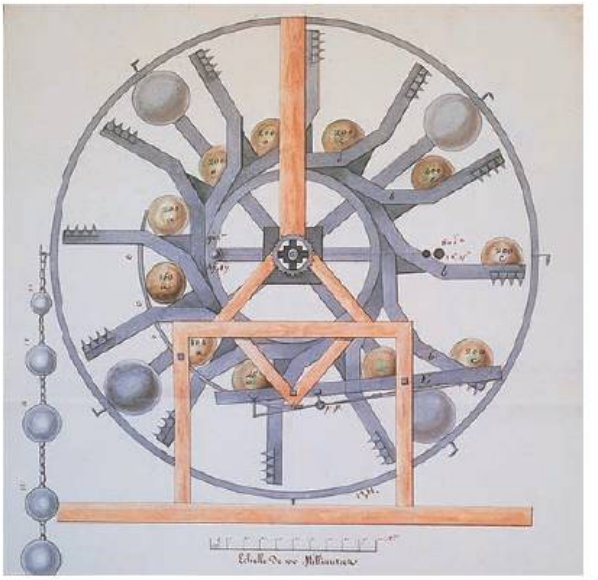


Figure 6.13 An attempt to design a machine to get something for nothing by breaking the law of conservation of energy

Example

A man lifts a weight of 480 N through a vertical distance of 3.5 m using a rope and some pulleys. The man pulls on the rope with a force of 200 N and a length of 10.5 m of rope passes through his hands. Calculate the efficiency of the pulley system.

work done by man = *force* × *distance moved (in direction of the force)*
= 200 × 10.5
= 2100 J

work done lifting load = 480 × 3.5
= 1680 J

as energy is the ability to do work and from the definition of efficiency,

efficiency = *work got out/work put in*
= 1680/2100
= **0.80 or 80%**

Now it's your turn

9 An electric heater converts electrical energy into heat energy. Suggest why this process may be 100% efficient.

10 The electric motor of an elevator (lift) uses 630 kJ of electrical energy when raising the elevator and passengers, of total weight 12 500 N, through a vertical height of 29 m. Calculate the efficiency of the elevator.

www.iranino.co

IRANINOVIN

6.4 Power

Machines such as wind turbines or engines do work for us when they change energy into a useful form. However, not only is the availability of useful forms of energy important, but also the rate at which it can be converted from one form to another. The rate of converting energy or using energy is known as **power**.

We have seen that energy is the ability to do work. Consider a family car and a Grand Prix racing car which both contain the same amount of fuel. They are capable of doing the same amount of work, but the racing car is able to travel much faster. This is because the engine of the racing car can convert the chemical energy of the fuel into useful energy at a much faster rate. The engine is said to be more powerful. Power is the rate of doing work. Power is given by the formula

$$power = \frac{work\ done}{time\ taken}$$

The unit of power is the watt (symbol W) and is equal to a rate of working of 1 joule per second. This means that a light bulb of power 1 W will convert 1 J of electrical energy to other forms of energy (e.g. light and heat) every second. Some typical values of power are shown in Table 6.3.

Table 6.3 Values of power

	power/W
power to operate a small calculator	50×10^{-6}
light power from a torch	4×10^{-3}
loudspeaker output	5
manual labourer working continuously	100
water buffalo working continuously	750
hair dryer	3×10^{3}
motor car engine	80×10^{3}
electric train	5×10^{6}
electricity generating station output	2×10^{9}

Power, like energy, is a scalar quantity.

Care must be taken when referring to power. It is common in everyday language to say that a strong person is 'powerful'. In physics, strength, or force, and power are *not* the same. Large forces may be exerted without any movement and thus no work is done and the power is zero! For example, a large rock resting on the ground is not moving, yet it is exerting a large force.

Consider a force F which moves a distance x at constant velocity v in the direction of the force, in time t. The work done W by the force is given by

$$W = Fx$$

Dividing both sides of this equation by time t gives

$$\frac{W}{t} = F\frac{x}{t}$$

Now, $\frac{W}{t}$ is the rate of doing work, i.e. the power P and $\frac{x}{t} = v$. Hence,

$$P = Fv$$

$$power = force \times velocity$$

www.iranino.co

IRANINOVIN

Example

A small electric motor is used to lift a weight of 1.5 N through a vertical distance of 120 cm in 2.7 s. Calculate the useful power output of the motor.

work done = *force* × *distance moved*

= 1.5 × 1.2 (the distance must be in metres)

= 1.8 J

power = *work done*/*time taken*

= (1.8/2.7)

= **0.67 W**

Now it's your turn

11 Calculate the electrical energy converted into thermal energy when an electric fire, rated at 2.4 kW, is left switched on for a time of 3.0 minutes.

12 The output power of the electric motors of a train is 3.6 MW when the train is travelling at 30 m s^{-1}. Calculate the total force opposing the motion of the train.

13 A boy of mass 60 kg runs up a flight of steps in a time of 1.8 s. There are 22 steps and each one is of height 20 cm. Calculate the useful power developed in the boy's legs. (The acceleration of free fall is 10 m s^{-2}.)



Figure 6.14 A domestic digital electricity meter

The kilowatt-hour

Every household has to pay the 'electricity bill'. Electricity is vital in modern living and this energy does not come free of charge. It is important to realise that what is paid for is electrical energy, not electrical power. Since many electrical appliances in the home have a power of the order of kilowatts and we use them for hours, the joule, as a unit of energy, is too small. For example, an electric fire of power 3 kW, used for 2 hours would use 3000 × 2 × 60 × 60 = 21 600 000 joules, i.e. 21.6 million joules of energy! Instead, in many parts of the world, electrical energy is purchased in kilowatt-hours (kW h).

One kilowatt-hour is the energy expended when work is done at the rate of 1 kilowatt for a time of 1 hour.

1 kW h = 1.0 kW × 1 hour

= 1000 W × 3600 s

1 kW h = 3.6 × 10^6 W s (1 W s is equivalent to 1 J)

The kilowatt-hour is sometimes referred to as the Unit of energy. Electricity meters in the home (Figure 6.14) are often shown as measuring Units, where 1 Unit = 1 kW h.

Example

Calculate the cost of using an electric fire, rated at 2.5 kW for a time of 6.0 hours, if 1 kW h of energy costs 7.0 cents.

energy used = 2.5 × 6.0 = 15 kW h

cost = 15 × 7.0

= **105 cents**

Now it's your turn

14 A television set is rated at 280 W. Calculate the cost of watching a three-hour film if 1 kilowatt-hour of electrical energy costs 8 cents.

15 An electric kettle is rated at 2.4 kW. Electrical energy costs 8 cents per kW h. The kettle takes 1.0 minute to boil sufficient water for two mugs of coffee. Calculate the cost of making this amount of coffee on three separate occasions.

16 Electrical energy generating companies sometimes measure their output in gigawatt-years. Calculate the number of kilowatt-hours in 6.0 gigawatt-years.

www.iranino.co

IRANINOVIN

Summary

- When a force moves its point of application in the direction of the force, work is done.
- *Work done* = $Fx\cos\theta$, where θ is the angle between the direction of the force F and the displacement x.
- When a gas expands at constant pressure:
  work done = pressure × change in volume, or $W = p\Delta V$
- Energy is needed to do work; energy is the ability to do work.
- Potential energy is the energy stored in a body due to its position or shape; examples are elastic potential energy and gravitational potential energy.
- When an object of mass m moves vertically through a distance Δh, then the change in gravitational potential energy is given by: $\Delta E_p = mg\Delta h$ where g is the acceleration of free fall.
- Kinetic energy is the energy stored in a body due to its motion.
- For an object of mass m moving with speed v, the kinetic energy is given by: $E_k = \frac{1}{2}mv^2$
- Energy cannot be created or destroyed. It can only be converted from one form to another.
- Efficiency = useful energy output/total energy input
- Power is defined as the rate of doing work or work done per unit time.
- *Work done = power × time taken*
- The unit of power is the watt (W).
- 1 watt = 1 joule per second
- *Power = force × velocity*
- Electrical energy may be measured in kilowatt-hours (kW h).
- 1 kW h is the energy expended when work is done at the rate of 1000 watts for a time of 1 hour. 1kW h is equivalent to 3.6MJ.

Examination style questions

1 A force F moves its point of application by a distance x in a direction making an angle θ with the direction of the force, as shown in Fig. 6.15.

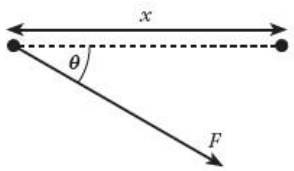


Fig. 6.15

The force does an amount W of work. Copy and complete the following table.

F/N	x/m	θ/°	W/J
15	6.0	0	
15	6.0	90	
15	6.0	30	
46		23	6.4
2.4×10^3	1.6×10^2		3.1×10^5
	2.8	13	7.1×10^3

2 An elastic band is stretched so that its length increases by 2.4 cm. The force required to stretch the band increases linearly from 6.3 N to 9.5 N. Calculate:

a the average force required to stretch the elastic band,
b the work done in stretching the band.

3 When water boils at an atmospheric pressure of 101 kPa, 1.00 cm^3 of liquid becomes 1560 cm^3 of steam. Calculate the work done against the atmosphere when a saucepan containing 550 cm^3 of water is allowed to boil dry.

4 Name each of the following types of energy:

a energy used in muscles,
b energy stored in the Sun,
c energy of water in a mountain lake,
d energy captured by a wind turbine,
e energy produced when a firework explodes,
f energy of a compressed gas.

5 A child of mass 35 kg moves down a sloping path on a skate board. The sloping path makes an angle of 4.5° with the horizontal. The constant speed of the child along the path is 6.5 m s^{-1}. Calculate:

a the vertical distance through which the child moves in 1.0 s,
b the rate at which potential energy is being lost (g = 10 m s^{-2}).

www.iranino.co

IRANINOVIN

6 A stone of mass 120 g is dropped down a well. The surface of the water in the well is 9.5 m below ground level. The acceleration of free fall of the stone is 9.8 m s^{-2}. Calculate, for the stone falling from ground level to the water surface:

a the loss of potential energy,

b its speed as it hits the water, assuming all the potential energy has been converted into kinetic energy.

7 An aircraft of mass 3.2 × 10^5 kg accelerates along a runway. Calculate the change in kinetic energy, in MJ, when the aircraft accelerates:

a from zero to 10 m s^{-1},

b from 30 m s^{-1} to 40 m s^{-1},

c from 60 m s^{-1} to 70 m s^{-1}.

8 In order to strengthen her legs, an athlete steps up on to a box and then down again 30 times per minute. The girl has mass 50 kg and the box is 35 cm high. The exercise lasts 4.0 minutes and as a result of the exercise, her leg muscles generate 120 kJ of heat energy. Calculate the efficiency of the leg muscles (g = 10 m s^{-2}).

9 By accident, the door of a refrigerator is left open. Use the law of conservation of energy to explain whether the temperature of the room will rise, stay constant or fall after the refrigerator has been working for a few hours.

10 The lights in a school laboratory have a total power of 600 W and are left on for 7.0 hours each day. In order to reduce fuel bills, it is decided to have the lights switched on only when there are people in the laboratory. This amounts to a total time of 4.5 hours per day. Assuming that the laboratory is used for 200 days each year, calculate the saving, if 1 kW h of energy costs 7.0 cents.

11 A car travels in a straight line at speed v along a horizontal road. The car moves against a resistive force F given by the equation

$$F = 400 + kv^2$$

where F is in newtons, v in m s^{-1} and k is a constant.

At speed v = 15 m s^{-1}, the resistive force F is 1100 N.

a Calculate, for this car:

i the power necessary to maintain the speed of 15 m s^{-1},

ii the total resistive force at a speed of 30 m s^{-1},

iii the power required to maintain the speed of 30 m s^{-1}.

b Determine the energy expended in travelling 1.2 km at a constant speed of:

i 15 m s^{-1},

ii 30 m s^{-1}.

c Using your answers to part **b**, suggest why, during a fuel shortage, the maximum permitted speed of cars may be reduced.

12 a Distinguish between gravitational potential energy and elastic potential energy. *[2]*

b A ball of mass 65 g is thrown vertically upwards from ground level with a speed of 16 m s^{-1}. Air resistance is negligible.

i Calculate, for the ball,

1 the initial kinetic energy, *[2]*

2 the maximum height reached. *[2]*

ii The ball takes time t to reach maximum height. For time $t/2$ after the ball has been thrown, calculate the ratio

$$\frac{\text{potential energy of ball}}{\text{kinetic energy of ball}}$$ *[3]*

iii State and explain the effect of air resistance on the time taken for the ball to reach maximum height. *[1]*

Cambridge International AS and A level Physics, 9702/23 Oct/Nov 2013 Q 4

13 a Explain what is meant by *work done*. *[1]*

b A car is travelling along a road that has a uniform downhill gradient, as shown in Fig. 6.16.

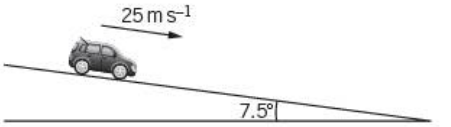


Fig. 6.16

The car has a total mass of 850 kg. The angle of the road to the horizontal is 7.5°. Calculate the component of the weight of the car down the slope. *[2]*

c The car in **b** is travelling at a constant speed of 25 m s^{-1}. The driver then applies the brakes to stop the car. The constant force resisting the motion of the car is 4600 N.

i Show that the deceleration of the car with the brakes applied is 4.1 m s^{-2}. *[2]*

ii Calculate the distance the car travels from when the brakes are applied until the car comes to rest. *[2]*

iii Calculate

1 the loss of kinetic energy of the car, *[2]*

2 the work done by the resisting force of 4600 N. *[1]*

iv The quantities in **iii** part **1** and in **iii** part **2** are not equal. Explain why these two quantities are not equal. *[1]*

Cambridge International AS and A level Physics, 9702/21 May/June 2011 Q 2

14 a Distinguish between gravitational potential energy and electric potential energy. *[2]*

b A body of mass m moves vertically through a distance h near the Earth's surface. Use the defining equation for work done to derive an expression for the gravitational potential energy change of the body. *[2]*

c Water flows down a stream from a reservoir and then causes a water wheel to rotate, as shown in Fig. 6.17.

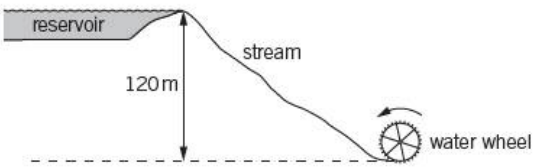


Fig. 6.17

As the water falls through a vertical height of 120 m, gravitational potential energy is converted to different forms of energy, including kinetic energy of the water.

www.iranino.co

IRANINOVIN

At the water wheel, the kinetic energy of the water is only 10% of its gravitational potential energy at the reservoir.

i Show that the speed of the water as it reaches the wheel is 15 m s^{-1}. [2]

ii The rotating water wheel is used to produce 110 kW of electrical power. Calculate the mass of water flowing per second through the wheel, assuming that the production of electric energy from the kinetic energy of the water is 25% efficient. [3]

Cambridge International AS and A level Physics, 9702/21 Oct/Nov 2011 Q 4

15 A ball is thrown vertically down towards the ground with an initial velocity of 4.23 m s^{-1}. The ball falls for a time of 1.51 s before hitting the ground. Air resistance is negligible.

a i Show that the downwards velocity of the ball when it hits the ground is 19.0 m s^{-1}. [2]

ii Calculate, to three significant figures, the distance the ball falls to the ground. [2]

b The ball makes contact with the ground for 12.5 ms and rebounds with an upwards velocity of 18.6 m s^{-1}. The mass of the ball is 46.5 g.

i Calculate the average force acting on the ball on impact with the ground. [4]

ii Use conservation of energy to determine the maximum height the ball reaches after it hits the ground. [2]

c State and explain whether the collision the ball makes with the ground is elastic or inelastic. [1]

Cambridge International AS and A level Physics, 9702/21 May/June 2012 Q 2

www.iranino.co

IRANINOVIN

9 Deformation of solids

By the end of this topic, you will be able to:

9.1 (a) appreciate that deformation is caused by a force and that, in one dimension, the deformation can be tensile or compressive
(b) describe the behaviour of springs in terms of load, extension, elastic limit, Hooke's law and the spring constant (force per unit extension)
(c) define and use the terms stress, strain and the Young modulus
(d) describe an experiment to determine the Young modulus of a metal in the form of a wire

9.2 (a) distinguish between elastic and plastic deformation of a material
(b) understand that the area under the force–extension graph is equal to work done
(c) deduce the strain energy in a deformed material from the area under the force–extension graph

Starting points

- When forces are applied to a solid body, its shape or size may change.
- The change of shape or size is called **deformation**.
- The deformation is called a **tensile** deformation if an object is stretched or a **compressive** deformation if the object squashed.

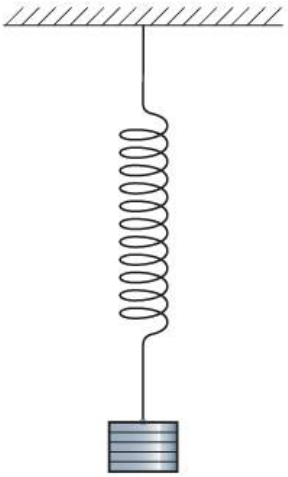
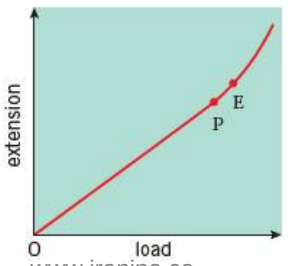


Figure 9.1 A loaded helical spring



www.iranino.co

Figure 9.2 Extension of a loaded spring

9.1 & 9.2 Force and deformation

Hooke's law

A helical spring, attached to a fixed point, hangs vertically and has weights attached to its lower end, as shown in Figure 9.1. As the magnitude of the weight is increased the spring becomes longer. The increase in length of the spring is called the **extension** of the spring and the weight attached to the spring is called the **load**.

If the load is increased greatly, the spring will change its shape permanently. However, for small loads, when the load is removed, the spring returns to its original length. The spring is said to have undergone an **elastic change**.

In an elastic change, a body returns to its original shape and size when the load on it is removed.

Figure 9.2 shows the variation with load of the extension of the spring. The section of the line from the origin to the point P is straight. In this region, the extension of the spring is proportional to the load. The point P is referred to as the **proportionality limit**. The point E is referred to as the **elastic limit** and is usually just beyond the point P. The spring is deformed permanently and the change is said to be **plastic** for points beyond E.

The elastic limit is the maximum force that can be applied to a wire/spring such that the wire/spring returns to its original length when the force is removed.

The fact that there is a straight line relationship between load and extension is expressed in Hooke's law. It should be appreciated that, although we have used a spring as an illustration, the law applies to any object, provided the proportionality limit has not been exceeded.

IRANINOVIN

Hooke's law states that, provided the proportionality limit is not exceeded, the extension of a body is proportional to the applied load.

The law can be expressed in the form of an equation

force $F \propto$ extension e

Removing the proportionality sign gives

$F = ke$

where k is a constant, known as the **spring constant** (or force constant).

The spring constant is the force per unit extension.

The unit of the constant is newton per metre (N m^{-1}).

Example

An elastic cord has an unextended length of 25 cm. When the cord is extended by applying a force at each end, the length of the cord becomes 40 cm for forces of 0.75 N. Calculate the force constant of the cord.

extension of cord = 15 cm

force constant = 0.75/0.15 (extension in metres)

= **5.0 N m^{-1}**

Now it's your turn

1 Explain what is meant by *extended elastically*.
2 Calculate the spring constant for a spring which extends by a distance of 3.5 cm when a load of 14 N is hung from its end.
3 A steel wire extends by 1.5 mm when it is under a tension of 45 N. Calculate:
(a) the force constant of the wire,
(b) the tension required to produce an extension of 1.8 mm, assuming that the proportionality limit is not exceeded.

Strain energy

When an object has its shape changed by forces acting on it, the object is said to be **strained**. Work has to be done by the forces to cause this strain. Provided that the elastic limit is not exceeded, the object can do work as it returns to its original shape when the forces are removed. Energy is stored in the body as potential energy when it is strained. This particular form of potential energy is called elastic potential energy or strain potential energy, or simply **strain energy**.

Elastic potential energy (strain energy) is energy stored in a body due to change of shape.

Consider the spring shown in Figure 9.1. To produce an extension e, the force applied at the lower end of the spring increases linearly with extension from zero to a value F. The average force is $\frac{1}{2}F$ and the work done W by the force is therefore

W = *average force* × *extension* (see Topic 6, page 80)

$= \frac{1}{2}Fe$

However, the force constant k is given by the equation

$F = ke$

Therefore, substituting for F,

strain energy $W = \frac{1}{2}ke^2$

The energy is in joules if k is in newtons per metre and e is in metres.

www.iranino.co

IRANINOVIN

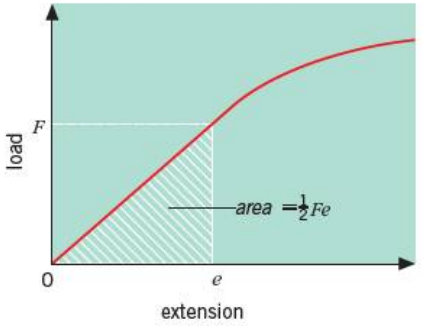


Figure 9.3 Strain energy is given by the area under the graph.

A graph of load (y-axis) against extension (x-axis) enables strain energy to be found even when the graph is not linear (see Figure 9.3). We have shown that strain energy is given by

$$\text{strain energy} = \tfrac{1}{2}Fe$$

The expression $\frac{1}{2}Fe$ represents the area between the straight line on Figure 9.3 and the x-axis. This means that strain energy is represented by the area under the line on a graph of load (y-axis) plotted against extension (x-axis).

Example

A spring has a spring constant 65 N m^{-1} and is extended elastically by 1.2 cm. Calculate the strain energy stored in the spring.

$$\text{strain energy } W = \tfrac{1}{2}ke^2$$
$$= \tfrac{1}{2} \times 65 \times (1.2 \times 10^{-2})^2$$
$$= \mathbf{4.7 \times 10^{-3}\,J}$$

Now it's your turn

4 A wire has a force constant of 5.5 × 10^4 N m^{-1}. It is extended elastically by 1.4 mm. Calculate the strain energy stored in the wire.

5 A rubber band has a force constant of 180 N m^{-1}. The work done in extending the band is 0.16 J. Calculate the extension of the band.

The Young modulus

The difficulty with using the force constant is that the constant is different for each specimen of a material having a different shape. It would be far more convenient if we had a constant for a particular material which would enable us to find extensions knowing the constant and the dimensions of the specimen. This is possible using the **Young modulus**.

We have already mentioned the term **strain**. When an object of original length L_o is extended by an amount e, the strain (ε) is defined as

$$\text{strain} = \frac{\text{extension}}{\text{original length}}$$

$$\varepsilon = e/L_o$$

Strain is the ratio of two lengths and does not have a unit.

The strain produced within an object is caused by a **stress**. In our case, we are dealing with changes in length and so the stress is referred to as a **tensile stress**. When a tensile force F acts normally to an area A, the stress (σ) is given by

$$\text{stress} = \frac{\text{force}}{\text{area normal to the force}}$$

$$\sigma = F/A$$

The unit of tensile stress is newton per square metre (N m^{-2}). This unit is also the unit of pressure and so an alternative unit for stress is the pascal (Pa).

In Figure 9.3, we plotted a graph of load against extension. Since load is related to stress and extension is related to strain, a graph of stress plotted against strain would have the same basic shape, as shown in Figure 9.4. Once again, there is a straight line region between the origin and P, the proportionality limit. In this region, changes of strain with stress are proportional.

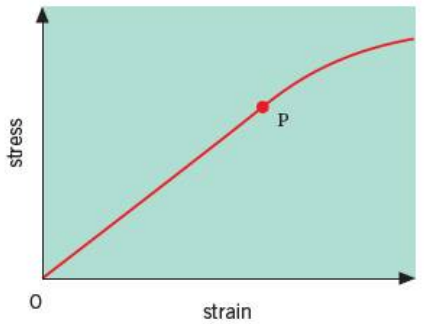


Figure 9.4 Stress–strain graph

www.iranino.co

IRANINOVIN

In the region where the changes are proportional, it can be seen that

$stress \propto strain$

or, removing the proportionality sign,

$stress = E \times strain$

The constant E is known as the Young modulus of the material.

$$\text{Young modulus } E = \frac{stress}{strain}$$

The unit of the Young modulus is the same as that for stress because strain is a ratio and has no unit.

This definition for the Young modulus can be used to derive the expression

$$E = (F/A) \times (L_o/e) = (FL_o)/(Ae)$$

This expression is used to determine the Young modulus of a metal.

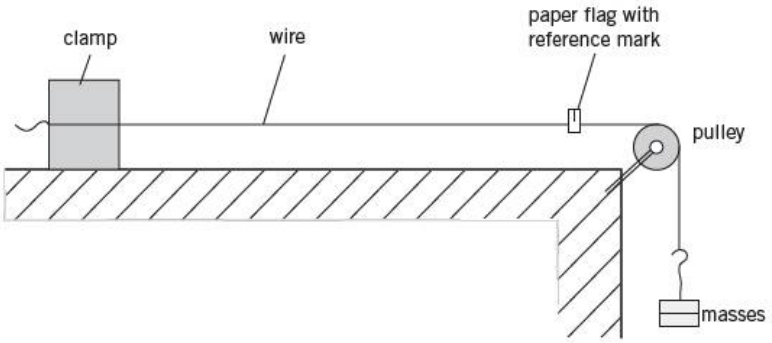


Figure 9.5 Simple experiment to measure the Young modulus of a wire

The Young modulus of a metal in the form of a wire may be measured by applying loads to a wire and measuring the extensions caused. The original length and the cross-sectional area must also be measured. A suitable laboratory arrangement is shown in Figure 9.5. A copper wire is often used. This is because, for wires of the same diameter under the same load, a copper wire will give larger, more measurable, extensions than a steel wire. (Why is this?) A paper flag with a reference mark on it is attached to the wire at a distance of approximately one metre from the clamped end. The original length L_o is measured from the clamped end to the reference mark, using a metre rule. The diameter d of the wire is measured using a micrometer screw gauge, and the cross-sectional area A calculated from $A = \frac{1}{4}\pi d^2$. Extensions e are measured as masses m are added to the mass-carrier. (Think of a suitable way of measuring these extensions.) The load F is calculated from $F = mg$. A graph of F (y-axis) against e (x-axis) has gradient EA/L_o, so the Young modulus E is equal to gradient $\times$ (L_o/A). This method is only applicable where Hooke's law is valid and the graph obtained is a straight line. Care should be taken not to exceed the limit of proportionality when extending the wire.

Some values of the Young modulus for different materials are shown in Table 9.1.

Table 9.1 Young modulus for different materials

material	Young modulus E/Pa
aluminium	7.0×10^{10}
copper	1.1×10^{11}
steel	2.1×10^{11}
glass	4.1×10^{10}
rubber	5×10^{8}

Example

A steel wire of diameter 1.0 mm and length 2.5 m is suspended from a fixed point and a mass of weight 45 N is suspended from its free end. The Young modulus of the material of the wire is 2.1×10^{11} Pa. Assuming that the proportionality limit of the wire is not exceeded, calculate:

(a) the applied stress,
(b) the strain,
(c) the extension of the wire.

www.iranino.co

IRANINOVIN

(a) area = $\pi \times (0.5 \times 10^{-3})^2$
$= 7.9 \times 10^{-7}\,m^2$
stress = force/area
$= 45/7.9 \times 10^{-7}$
$= \mathbf{5.7 \times 10^7\,Pa}$

(b) strain = stress/Young modulus
$= 5.7 \times 10^7/2.1 \times 10^{11}$
$= \mathbf{2.7 \times 10^{-4}}$

(c) extension = strain × length
$= 2.7 \times 10^{-4} \times 2.5$
$= 6.8 \times 10^{-4}\,m$
$= \mathbf{0.68\,mm}$

Now it's your turn

6 A copper wire of diameter 1.78 mm and length 1.40 m is suspended from a fixed point and a mass of weight 32.0 N is suspended from its free end. The Young modulus of the material of the wire is 1.10×10^{11} Pa. Assuming that the proportionality limit of the wire is not exceeded, calculate:
(a) the applied stress,
(b) the strain,
(c) the extension of the wire.

7 An elastic band of area of cross-section 2.0 mm^2 has an unextended length of 8.0 cm. When stretched by a force of 0.40 N, its length becomes 8.3 cm. Calculate the Young modulus of the elastic.

Summary

- Forces on an object can cause tensile deformation (stretching) or compressive deformation (squashing).
- An elastic change occurs when an object returns to its original shape and size when the load is removed from it.
- Hooke's law states extension is proportional to load provided the limit of proportionality is not exceeded.
- The spring constant (force constant) k is the ratio of force to extension.
- Elastic potential energy (strain energy) is energy stored in a body due to change of shape.
- *Strain energy* = $\frac{1}{2}ke^2$
- *Tensile strain = extension/original length*
- *Tensile stress = force/cross-sectional area*
- *Young modulus = stress/strain*

Examination style questions

1 A spring has an unextended length of 12.4 cm. When a load of 4.5 N is suspended from the spring, its length becomes 13.3 cm. Calculate:
a the spring constant of the spring,
b the length of the spring for a load of 3.5 N.

2 The elastic cord of a catapult has a force constant of 700 N m^{-1}. Calculate the energy stored in the elastic cord when it is extended by 15 cm.

3 Two wires each have length 1.8 m and diameter 1.2 mm. One wire has a Young modulus of 1.1×10^{11} Pa and the other 2.2×10^{11} Pa. One end of each wire is attached to the same fixed point and the other end of each wire is attached to the same load of 75 N so that each has the same extension. Assuming that the proportionality limit of the wires is not exceeded, calculate the extension of the wires.

4 Explain what is meant by *plastic deformation*.

5 a Explain what is meant by *strain energy (elastic potential energy)*. [2]
b A spring that obeys Hooke's law has a spring constant k. [3]
Show that the energy E stored in the spring when it has been extended elastically by an amount x is given by
$$E = \frac{1}{2}kx^2.$$

www.iranino.co

IRANINOVIN

c A light spring of unextended length 14.2 cm is suspended vertically from a fixed point, as illustrated in Fig. 9.6.

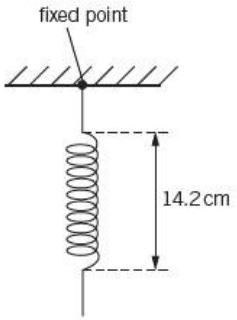


Fig. 9.6

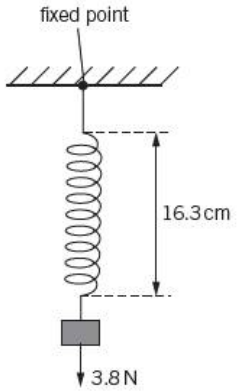


Fig. 9.7

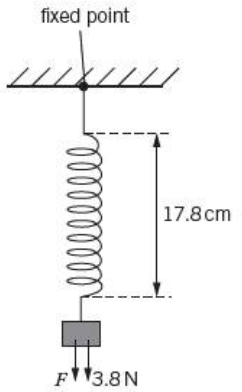


Fig. 9.8

A mass of weight 3.8 N is hung from the end of the spring, as shown in Fig. 9.7. The length of the spring is now 16.3 cm.

An additional force F then extends the spring so that its length becomes 17.8 cm, as shown in Fig. 9.8.

The spring obeys Hooke's law and the elastic limit of the spring is not exceeded.

i Show that the spring constant of the spring is 1.8 N cm^{-1}. *[1]*

ii For the extension of the spring from a length of 16.3 cm to a length of 17.8 cm,

1 calculate the change in the gravitational potential energy of the mass on the spring, *[2]*

2 show that the change in elastic potential energy of the spring is 0.077 J, *[1]*

3 determine the work done by the force F. *[1]*

Cambridge International AS and A level Physics, 9702/22 Oct/Nov 2009 Q 4

6 a Define, for a wire,

i *stress*, *[1]*

ii *strain*. *[1]*

b A wire of length 1.70 m hangs vertically from a fixed point, as shown in Fig. 9.9.

The wire has cross-sectional area 5.74 × 10^{-8} m^2 and is made of a material that has a Young modulus of 1.60 × 10^{11} Pa. A load of 25.0 N is hung from the wire.

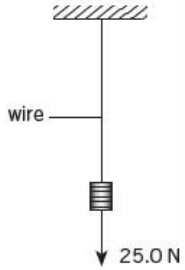


Fig. 9.9

i Calculate the extension of the wire. *[3]*

ii The same load is hung from a second wire of the same material. This wire is twice the length but the **same volume** as the first wire. State and explain how the extension of the second wire compares with that of the first wire. *[3]*

Cambridge International AS and A Level Physics, 9702/21 May/June 2011 Q 4

www.iranino.co

IRANINOVIN

7 **a** State Hooke's law. *[1]*

b The variation with extension x of the force F for a spring A is shown in Fig. 9.10.
The point L on the graph is the elastic limit of the spring.

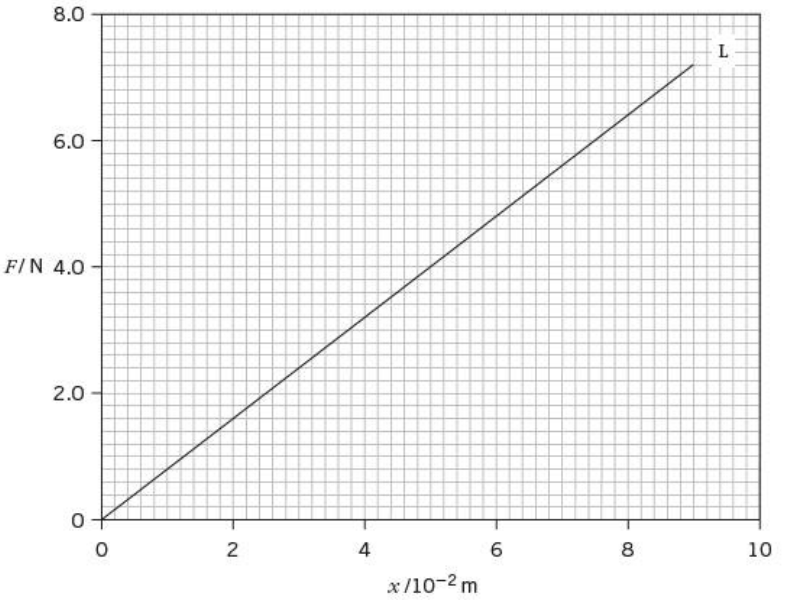


Fig. 9.10

i Describe the meaning of *elastic limit*. *[1]*

ii Calculate the spring constant k_A for spring A. *[1]*

iii Calculate the work done in extending the spring with a force of 6.4 N. *[2]*

c A second spring B of spring constant $2k_A$ is now joined to spring A, as shown in Fig. 9.11.

A force of 6.4 N extends the combination of springs.

For the combination of springs, calculate

i the total extension, *[1]*

ii the spring constant. *[1]*

Cambridge International AS and A Level Physics, 9702/21 Oct/Nov 2011 Q 6

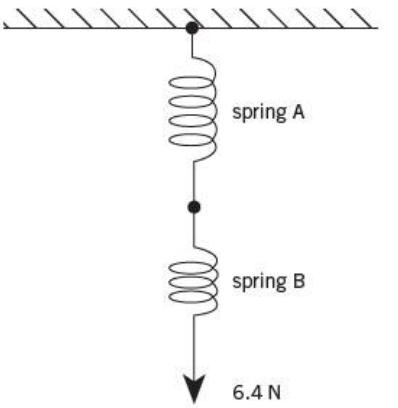


Fig. 9.11

www.iranino.co

IRANINOVIN

14 Waves

By the end of this topic, you will be able to:

14.1 (a) describe what is meant by wave motion as illustrated by vibration in ropes, springs and ripple tanks
(b) understand and use the terms displacement, amplitude, phase difference, period, frequency, wavelength and speed
(c) deduce, from the definitions of speed, frequency and wavelength, the equation $v = f\lambda$
(d) recall and use the equation $v = f\lambda$
(e) understand that energy is transferred due to a progressive wave
(f) recall and use the relationship intensity is proportional to (amplitude)^2

14.2 (a) compare transverse and longitudinal waves
(b) analyse and interpret graphical representations of transverse and longitudinal waves

14.3 (a) determine the frequency of sound using a calibrated cathode-ray oscilloscope (c.r.o)

14.4 (a) understand that when a source of waves moves relative to a stationary observer, there is a change in observed frequency
(b) use the expression $f_o = f_s v/(v \pm v_s)$ for the observed frequency when a source of sound waves moves relative to a stationary observer
(c) appreciate that the Doppler shift is observed with all waves, including sound and light

14.5 (a) state that all electromagnetic waves travel with the same speed in free space and recall the orders of magnitude of the wavelengths of the principal radiations from radio waves to γ-rays

Note: 14.3 (b) is covered in AS Level Topic 15.

Starting points

- Review basic properties of waves.
- Review knowledge of the electromagnetic spectrum.

14.1 Wave motion

This topic will introduce some general properties of waves. We will meet two broad classifications of waves, transverse and longitudinal, based on the direction in which the particles vibrate relative to the direction in which the wave transmits energy. We will define terms such as amplitude, wavelength and frequency for a wave, and derive the relationship between speed, frequency and wavelength. We will look at demonstrations of some properties of waves, such as reflection and refraction.

Wave motion is a means of moving energy from place to place. For example, electromagnetic waves from the Sun carry the energy that plants need to survive and grow. The energy carried by sound waves causes our ear drums to vibrate. The energy carried by seismic waves (earthquakes) can devastate vast areas, causing land to move and buildings to collapse. Waves which move energy from place to place without the transfer of matter are called **progressive waves**.

Vibrating objects act as sources of waves. For example, a vibrating tuning fork sets the air close to it into oscillation, and a sound wave spreads out from the fork. For a radio wave, the vibrating objects are electrons.

There are two main groups of waves. These are **transverse** waves and **longitudinal** waves.

A transverse wave is one in which the vibrations of the particles in the wave are at right angles to the direction in which the energy of the wave is travelling.

www.iranino.co

IRANINOVIN

Figure 14.1 shows a transverse wave moving along a rope. The particles of the rope vibrate up and down, whilst the energy travels at right angles to this, from A to B. There is no transfer of matter from A to B. Examples of transverse waves include light waves, surface water waves and secondary seismic waves (S-waves).

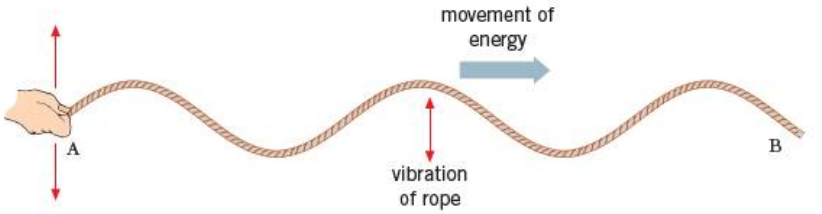


Figure 14.1 Transverse wave on a rope

A longitudinal wave is one in which the direction of the vibrations of the particles in the wave is along the direction in which the energy of the wave is travelling.

Figure 14.2 shows a longitudinal wave moving along a stretched spring (a 'slinky'). The coils of the spring vibrate along the length of the spring, whilst the energy travels along the same line, from A to B. Note that the spring itself does not move from A to B. Examples of longitudinal waves include sound waves and primary seismic waves (P-waves).

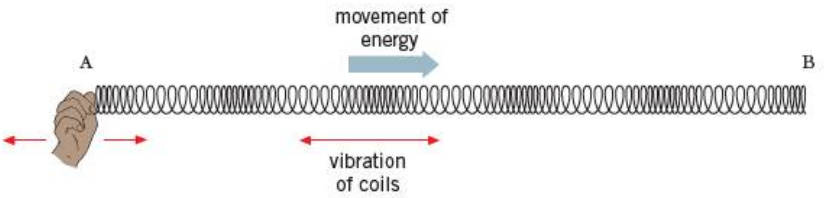


Figure 14.2 Longitudinal wave on a slinky spring

14.2 Graphical representation of waves

The **displacement** of a particle on a wave is its distance in a specified direction from its rest position.

Displacement is a vector quantity; it can be positive or negative. A transverse wave may be represented by plotting displacement y on the y-axis against distance x along the wave, in the direction of energy travel, on the x-axis. This is shown in Figure 14.3. It can be seen that the graph is a snapshot of what is actually observed to be a transverse wave.

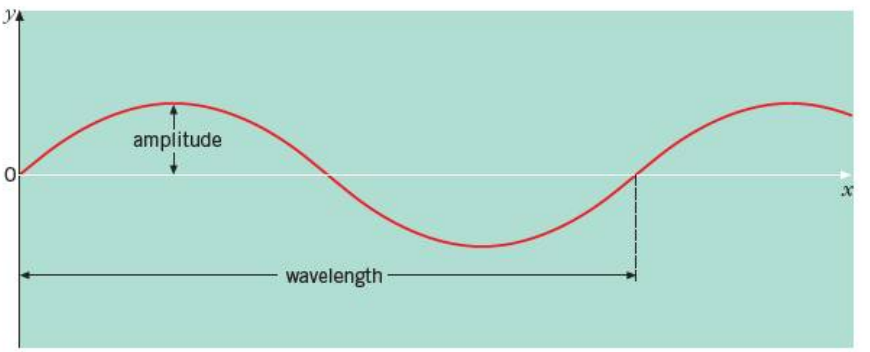


Figure 14.3 Displacement–distance graph for a transverse wave

For a longitudinal wave, the displacement of the particles is along the direction of energy travel. However, if these displacements are plotted on the y-axis of a graph of displacement against distance, the graph has exactly the same shape as for a transverse

www.iranino.co

IRANINOVIN

wave (see Figure 14.3). This is very useful, in that one graph can represent both types of wave. Using this graph, wave properties may be treated without reference to the type of wave.

The **amplitude** of the wave motion is defined as the maximum displacement of a particle in the wave.

Also, it can be seen that the wave repeats itself. That is, the wave can be constructed by repeating a section of the wave. The length of the smallest repetition unit is called the **wavelength**.

The wavelength is the shortest distance between two peaks or the shortest distance between two troughs. It is the shortest distance between points which are vibrating in phase (see below) with each other. It is the distance moved by the wavefront during one oscillation of the source of the waves.

The usual symbol for wavelength is λ, the Greek letter lambda.

Another way to represent both waves is to plot a graph of displacement y of a point on the wave against time t. This is shown in Figure 14.4. Again, the wave repeats itself after a certain interval of time. The time for one complete vibration is called the **period** T of the wave.

The period of the wave is the time for a particle in the wave to complete one vibration, or one cycle.

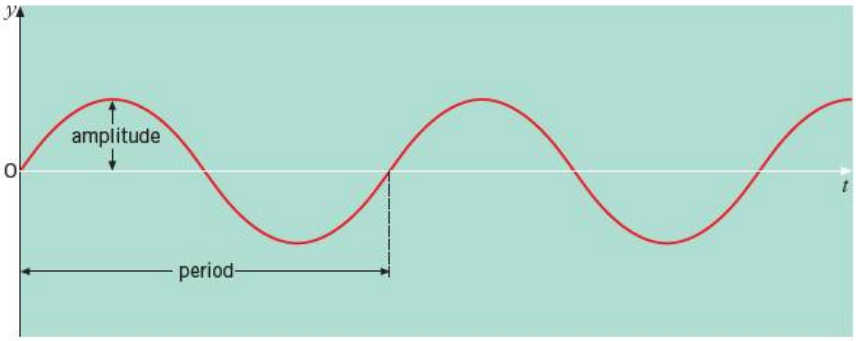


Figure 14.4 Displacement–time graph for a wave

The number of complete vibrations (cycles) per unit time is called the **frequency** f of the wave.

For waves on ropes and springs, displacement and amplitude are measured in mm, m or other units of length. Period is measured in seconds (s). Frequency has the unit per second (s^{-1}) or hertz (Hz).

A term used to describe the relative positions of the crests or troughs of two different waves of the same frequency is **phase**. When the crests and troughs of the two waves are aligned, the waves are said to be **in phase**. When crests and troughs are not aligned the waves are said to have a **phase difference**. When a crest and a trough of two waves are aligned the waves are said to be in **antiphase**. The quantitative measure of phase difference has the unit of angle (radians or degrees). Thus, when waves are out of phase with a crest aligned with a trough, one wave is half a cycle behind the other. Since one cycle is equivalent to 2π radians or 360°, the phase difference between waves that are exactly out of phase (in antiphase) is π radians or 180°.

Consider Figure 14.5, in which there are two waves of the same frequency, but with a phase difference between them. The period T corresponds to a phase angle of 2π rad or 360°. The two waves are out of step by a time t. Thus, phase difference is equal to $2\pi(t/T)$ rad = $360(t/T)°$. A similar argument may be used for waves of wavelength λ which are out of step by a distance x. In this case the phase difference is $2\pi(x/\lambda)$ rad = $360(x/\lambda)°$.

www.iranino.co

IRANINOVIN

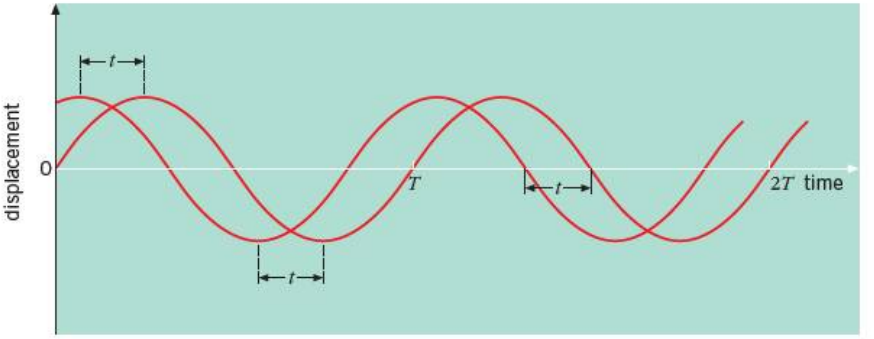


Figure 14.5 Phase difference

One of the characteristics of a progressive wave is that it carries energy. The amount of energy passing through unit area per unit time is called the **intensity** of the wave. Thus the intensity of a wave is the power per unit area. The intensity is proportional to the square of the amplitude of a wave. Thus, doubling the amplitude of a wave increases the intensity of the wave by a factor of four. The intensity also depends on the frequency: intensity is proportional to the square of the frequency.

For a wave of amplitude A and frequency f, the intensity I is proportional to A^2f^2.

If the waves from a point source spread out equally in all directions, we have what is called a **spherical wave**. As the wave travels further from the source, the energy it carries passes through an increasingly large area. Since the surface area of a sphere is $4\pi r^2$, the intensity is $W/4\pi r^2$, where W is the power of the source. The intensity of the wave thus decreases with increasing distance from the source. The intensity I is proportional to $1/r^2$, where r is the distance from the source.

This relationship assumes that there is no absorption of wave energy.

Wave equation

From the definition of wavelength λ, in one cycle of the source the wave energy moves a distance λ. The time taken for one cycle is the time period T. Referring to Topic 3, speed v is the distance moved per unit time. Therefore

$$v = \lambda/T$$

If f is the frequency of the wave, then $f = 1/T$. Therefore

$$v = f\lambda$$

or

speed = frequency × wavelength

This is an important relationship between the speed of a wave and its frequency and wavelength.

Examples

1 A tuning fork of frequency 170 Hz produces sound waves of wavelength 2.0 m. Calculate the speed of sound.

Using $v = f\lambda$, we have $v = 170 \times 2.0 =$ **340 m s^{-1}**.

2 The amplitude of a wave in a rope is 15 mm. If the amplitude were changed to 20 mm, keeping the frequency the same, by what factor would the power carried by the rope change?

Intensity is proportional to the square of the amplitude. Here the amplitude has been increased by a factor of 20/15, so the power carried by the wave increases by a factor of $(20/15)^2 =$ **1.8**.

www.iranino.co

IRANINOVIN

Now it's your turn

1 Water waves of wavelength 0.080 m have a frequency 5.0 Hz. Calculate the speed of these water waves.
2 The speed of sound is 340 m s^{-1}. Calculate the wavelength of the sound wave produced by a violin when a note of frequency 500 Hz is played.
3 A sound wave has twice the intensity of another sound wave of the same frequency. Calculate the ratio of the amplitudes of the waves.

Properties of wave motions

Although there are many different types of waves (light waves, sound waves, electrical waves, mechanical waves, etc.) there are some basic properties which they all have in common. All waves can be reflected and refracted. All waves can be diffracted, and can produce interference patterns (see Topic 15).

These properties may be demonstrated using a ripple tank similar to that shown in Figure 14.6. As the motor turns, the wooden bar vibrates, creating ripples on the surface of the water. The ripples are lit from above. This creates shadows on the viewing screen. The shadows show the shape and movement of the waves. Each shadow corresponds to a particular point on the wave, and is referred to as a **wavefront**.

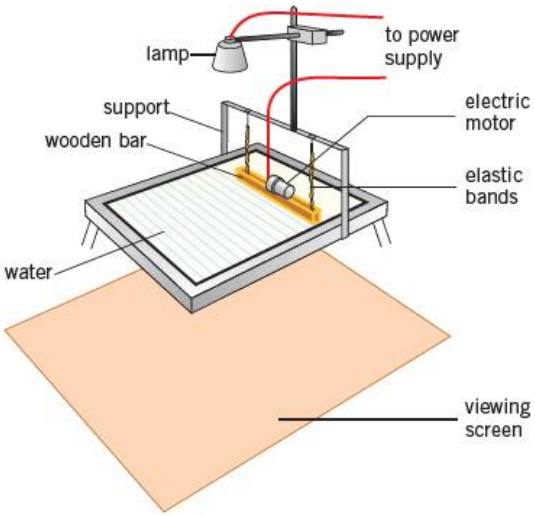


Figure 14.6 Ripple tank

Figure 14.7 illustrates the pattern of wavefronts produced by a low-frequency vibrator and one of higher frequency. Note that for the higher frequency the wavelength is less, since wave speed is constant and $v = f\lambda$.

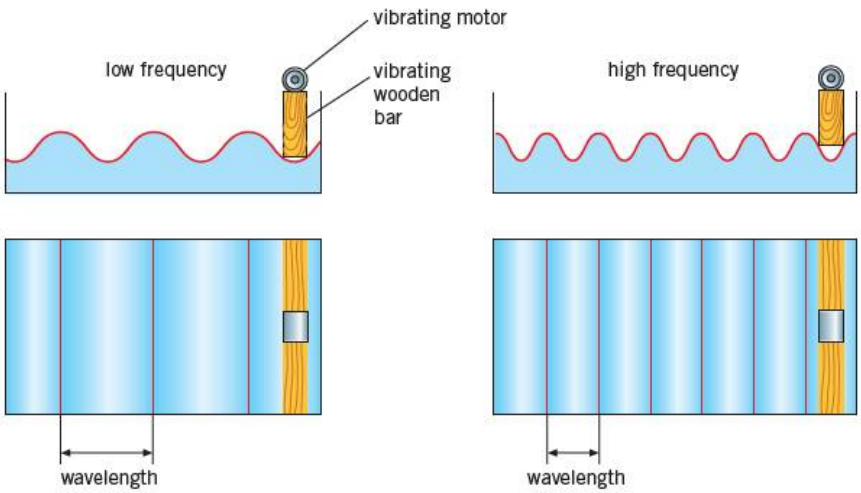


Figure 14.7 Ripple tank patterns for low- and high-frequency vibrations

www.iranino.com

IRANINOVIN

Circular waves may be produced by replacing the vibrating bar with a small dipper, or by allowing drops of water to fall into the ripple tank. A circular wave is illustrated in Figure 14.8. This pattern is characteristic of waves spreading out from a point source.

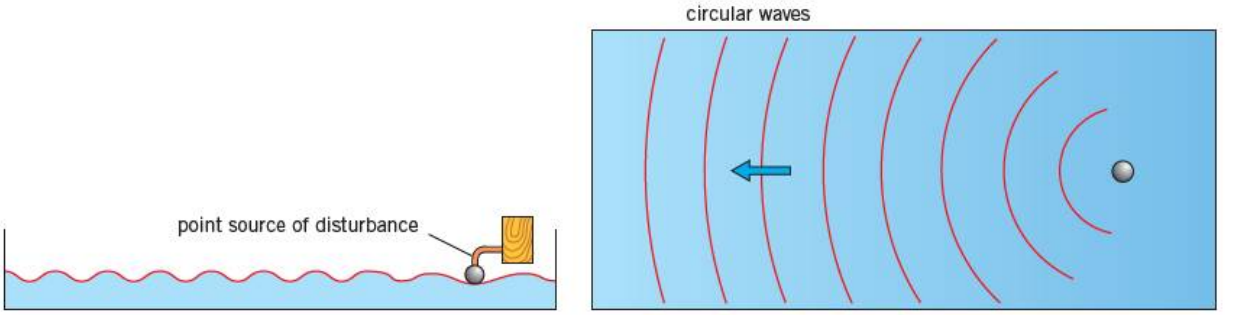


Figure 14.8 Ripple tank pattern for a point source

We shall now see how the ripple tank may be used to demonstrate the wave properties of reflection and refraction.

Reflection

As the waves strike a plane barrier placed in the water, they are reflected. The angle of reflection equals the angle of incidence, and there is no change in wavelength (see Figure 14.9a). If a curved barrier is used, the waves can be made to converge or diverge (Figure 14.9b).

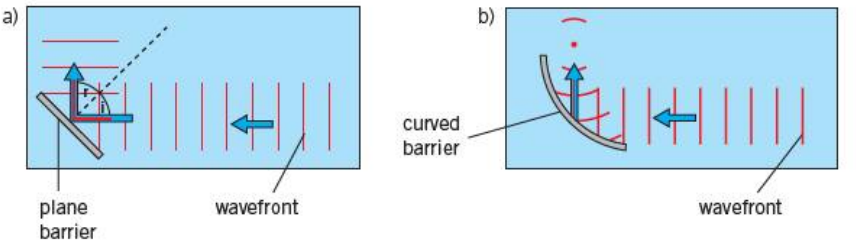


Figure 14.9 Ripple tank pattern showing reflection at a) a plane surface and b) a curved one

Refraction

If a glass block is submerged in the water, this produces a sudden change in the depth of the water. The speed of surface ripples on water depends on the depth of the water: the shallower the water, the slower the speed. Thus, the waves move more slowly as they pass over the glass block. The frequency of the waves remains constant, and so the wavelength decreases. If the waves are incident at an angle to the submerged block, they will change direction, as shown in Figure 14.10.

The change in direction of a wave due to a change in speed is called **refraction**.

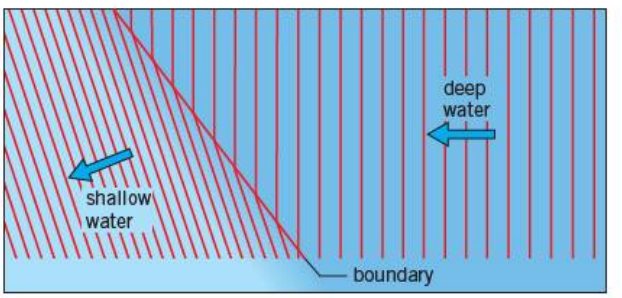


Figure 14.10 Ripple tank pattern showing refraction

www.iranino.co

As the waves re-enter the deeper water, their speed increases to the former value and they change direction once again.

14.3 The determination of the frequency of sound using a calibrated c.r.o.

A cathode *h* ray oscilloscope (c.r.o.) with a calibrated time-base may be used to determine the frequency of sound. A method of measuring the frequency of sound waves is illustrated in Figure 14.11. A signal generator and loudspeaker are used to produce a note of a single frequency. The microphone is connected to the Y-plates of the c.r.o.

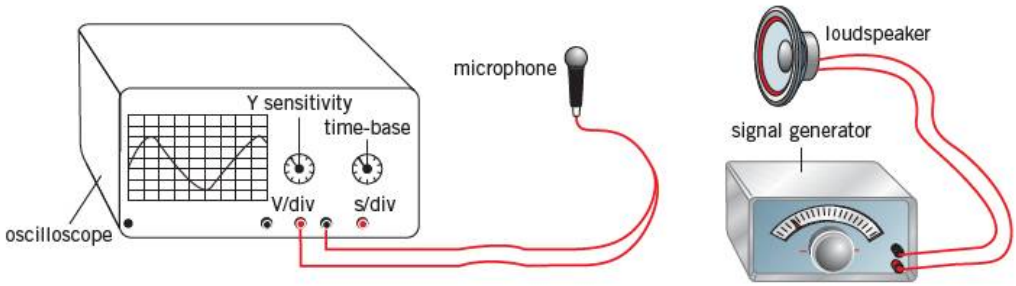


Figure 14.11 Measuring the frequency of sound using a c.r.o.

The microphone detects the sound and a trace on the c.r.o. can be obtained by adjusting the Y-plate sensitivity and the time-base settings. A typical trace is shown in Figure 14.12.

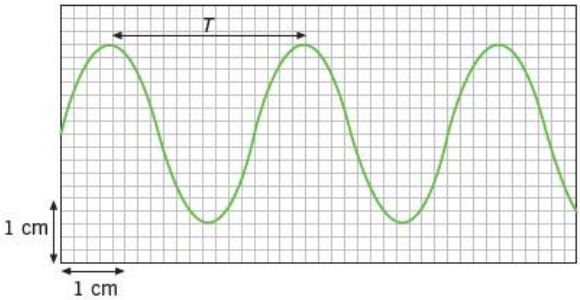


Figure 14.12 Measurement of the time period from the c.r.o. trace

The distance between peaks or troughs is measured using the scale on the c.r.o. display. The time-base setting is used to determine the time period T and frequency of the sound. The calculated value can be compared with that shown on the signal generator.

Example

The time-base setting for the c.r.o. used to obtain the trace in Figure 14.12 is 2.0 ms cm^{-1}. Determine for the sound

(a) the time period,
(b) the frequency.

(a) The distance for a time period is 3.0 cm.
Hence the period = 3.0 × 2.0 × 10^{-3} = **6.0 × 10^{-3} s**
(b) The frequency = 1/time period = 1/6.0 × 10^{-3} = **1 Hz**

Now it's your turn

4 The time-base on a c.r.o. is set on 0.50 ms cm^{-1}. The trace obtained for a sound wave shows three complete time periods in 7.2 cm. Calculate
(a) the time period,
(b) the frequency.

www.iranino.co

IRANINOVIN

14.4 Doppler effect

The whistle of a train or the siren of a police car appears to increase in frequency as it moves towards a stationary observer. The frequency change due to the relative motion between a source of sound or light and an observer is known as the **Doppler effect**.

When the observer and source of sound are both stationary, the number of waves per second reaching the observer will be the same frequency as the source (see Figure 14.13).

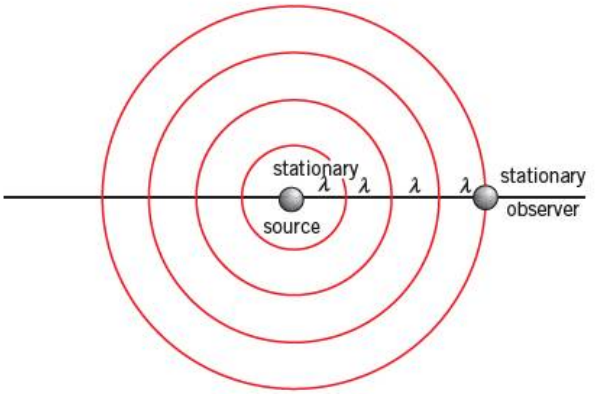


Figure 14.13 The source emits waves of wavelength λ. The observer is stationary and receives waves with the same wavelength λ.

When the source moves towards the observer the effect is to shorten the wavelength of the waves reaching the observer (see Figure 14.14).

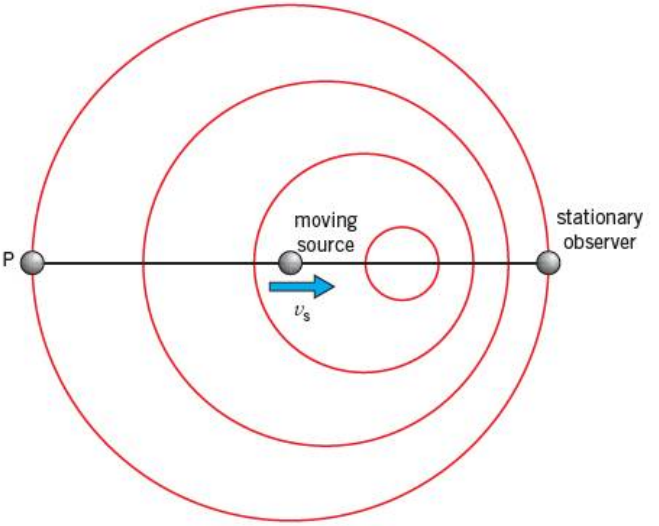


Figure 14.14 Source of sound moving towards a stationary observer

Let v be the speed of sound in air. A source of sound has a frequency f_s and wavelength λ. The source moves towards an observer at a speed v_s.

The period of oscillation of the source of sound is $T (= 1/f_s)$. In the time of one oscillation the source moves **towards** the observer a distance v_sT. Hence the wavelength is shortened by this distance. The wavelength of the sound received by the observer is $\lambda - v_sT$.

Hence the frequency observed $f_o = v/(\lambda - v_sT) = v/(v/f_s - v_s/f_s)$

$$f_o = f_s v/(v - v_s)$$

The source would move away from a stationary observer at position P on the left-hand side of Figure 14.14. The observed wavelengths would lengthen.

For a source of sound moving **away** from an observer the observed frequency can be shown to be

$f_o = f_s v/(v + v_s)$

www.iranino.co

IRANINOVIN

The frequency is *increased* when the source moves towards the observer and the frequency is *decreased* when the source moves away from the observer.

The above expressions apply only when the source of waves is sound. However, a change of frequency (Doppler shift) is observed with all waves, including light.

In astronomy, the wavelength tends to be measured rather than the frequency. If the measured wavelength of an observed spectral line (see Topic 25) is less than that measured for a stationary source, then the distance between the source (star) and detector is decreasing (**blue shift**). If the measured wavelength is greater than the value of a stationary source, then the source is moving away from the detector (**red shift**). The blue and red shifts are referred to in this way as red has the longest wavelength in the visible spectrum and blue the shortest.

Example

A police car travels towards a stationary observer at a speed of 15 m s^{-1}. The siren on the car emits a sound of frequency 250 Hz. Calculate the observed frequency. The speed of sound is 340 m s^{-1}.

Observed frequency $f_o = f_s v/(v - v_s) = 250 \times 340/(340 - 15)$

$= \mathbf{260\,Hz}$

Now it's your turn

5 The sound emitted from the siren of an ambulance has a frequency of 1500 Hz. The speed of sound is 340 m s^{-1}. Calculate the difference in frequency heard by a stationary observer as the ambulance travels towards and then away from the observer at a speed of 30 m s^{-1}.

14.5 The electromagnetic spectrum

Visible light is just a small region of the **electromagnetic spectrum**. All electromagnetic waves are transverse waves, consisting of electric and magnetic fields which oscillate at right angles to each other and to the direction in which the wave is travelling. This is illustrated in Figure 14.15. Electromagnetic (e.m.) waves show all the properties common to wave motions: they can be reflected and refracted. They can be diffracted and they obey the principle of superposition and produce interference patterns (see Topic 15). In a vacuum (free space) all electromagnetic waves travel at the same speed, 3.0 × 10^8 m s^{-1}.

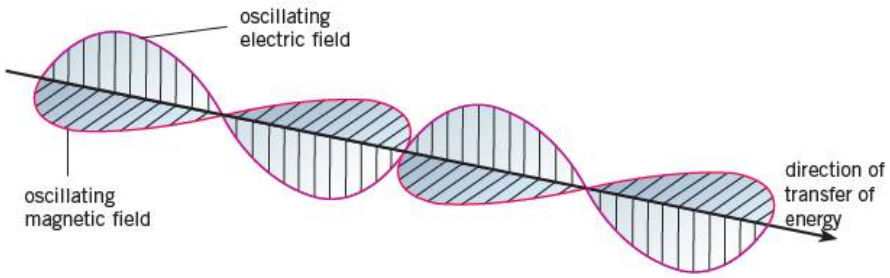


Figure 14.15 Oscillating electric and magnetic fields in an electromagnetic wave

The complete electromagnetic spectrum is divided into a series of regions based on the properties of electromagnetic waves in these regions, as illustrated in Figure 14.16. It should be noted that there is no clear boundary between regions.

www.iranino.co

IRANINOVIN

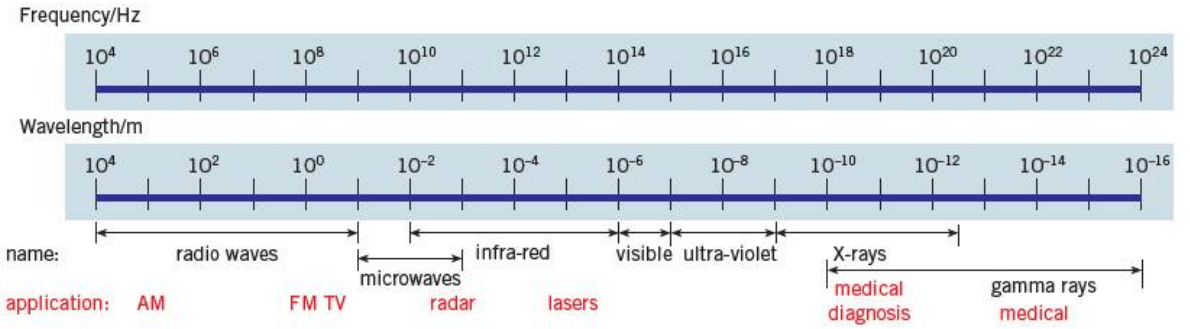


Figure 14.16 The electromagnetic spectrum

Table 14.1 Wavelengths of the principal radiations of the electromagnetic spectrum

radiation	approximate wavelength range/m
γ-rays	10^{-10}–10^{-16} and shorter
X-rays	10^{-9}–10^{-12} and shorter
ultra-violet	10^{-7}–10^{-9}
visible	4×10^{-7}–7×10^{-7}
infra-red	10^{-2}–10^{-6}
microwaves	10^{-3}–10^{-1}
radio waves	10^{-1}–10^{4} and longer

Examples

1 Calculate the frequency in MHz of a radio wave of wavelength 250 m.

The speed of all electromagnetic waves is 3.0×10^8 m s^{-1}.

For a wave $v = f\lambda$

$f = 3 \times 10^8/250 = 1.2 \times 10^6$ Hz = **1.2 MHz**.

2 Calculate the wavelength in nm of an X-ray wave of frequency 2.0×10^{18} Hz.

For a wave $v = f\lambda$

$\lambda = 3 \times 10^8/2 \times 10^{18} = 1.5 \times 10^{-10}$ = **0.15 nm**

Now it's your turn

6 The speed of light is 3.0×10^8 m s^{-1}. Calculate the frequency of red light of wavelength 6.5×10^{-7} m.

7 A beam of red light has an amplitude that is 2.5 times the amplitude of a second beam of the same colour. Calculate the ratio of the intensities of the waves.

8 Calculate the wavelength of microwaves of frequency 8.0 GHz.

Summary

- A progressive wave travels outwards from the source, carrying energy but without transferring matter.
- In a transverse wave, the oscillations of the particles are at right angles to the direction in which the wave carries energy.
- In a longitudinal wave, the oscillations of the particles are along the direction in which the wave carries energy.
- The intensity of a wave is the energy passing through unit area per unit time. Intensity is proportional to the square of the amplitude (and to the square of the frequency).
- The speed v, frequency f and wavelength λ of a wave are related by $v = f\lambda$.
- When a source of waves moves relative to a stationary observer there is a change in the observed frequency.
- The Doppler shift is observed with all waves, including sound and light.
- Properties of wave motion (reflection and refraction) can be observed in a ripple tank.
- All wavelengths of electromagnetic radiation have the same speed in free space $c = 3.0 \times 10^8$ m s^{-1}.

www.iranino.co

IRANINOVIN

Examination style questions

1 A certain sound wave in air has a speed 340 m s^{-1} and wavelength 1.7 m. For this wave, calculate:
   a the frequency,
   b the period.

2 The speed of electromagnetic waves in vacuum (or air) is 3.0×10^8 m s^{-1}.
   a The visible spectrum extends from a wavelength of 400 nm (blue light) to 700 nm (red light). Calculate the range of frequencies of visible light.
   b A typical frequency for v.h.f. television transmission is 250 MHz. Calculate the corresponding wavelength.

3 Two waves travel with the same speed and have the same amplitude, but the first has twice the wavelength of the second. Calculate the ratio of the intensities transmitted by the waves.

4 A student stands at a distance of 5.0 m from a point source of sound, which is radiating uniformly in all directions. The intensity of the sound wave at her ear is 6.3×10^{-6} W m^{-2}.
   a The receiving area of the student's ear canal is 1.5 cm^2. Calculate how much energy passes into her ear in 1 minute.
   b The student moves to a point 1.8 m from the source. Calculate the new intensity of the sound.

5 Assume that waves spread out uniformly in all directions from the epicentre of an earthquake. The intensity of a particular earthquake wave is measured as 5.0×10^6 W m^{-2} at a distance of 40 km from the epicentre. What is the intensity at a distance of only 2 km from the epicentre?

6 a i Define, for a wave,
   1 wavelength λ, [1]
   2 frequency f. [1]
   ii Use your definitions to deduce the relationship between λ, f and the speed v of the wave. [1]

b Plane waves on the surface of water are represented by Fig. 14.17 at one particular instant of time.

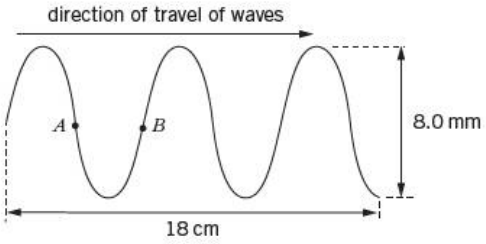


Fig. 14.17

The waves have frequency 2.5 Hz. Determine, for the waves,
   i the amplitude, [1]
   ii the speed, [2]
   iii the phase difference between points A and B. [1]

c The wave in **b** was produced in a ripple tank. Describe briefly, with the aid of a sketch diagram, how the wave may be observed. [2]

Cambridge International AS and A level Physics, 9702/23 Oct/Nov 2013 Q 5

7 A student is studying a water wave in which all the wavefronts are parallel to one another. The variation with time t of the displacement x of a particular particle in the wave is shown in Fig. 14.18.

The distance d of the oscillating particles from the source of the waves is measured. At a particular time, the variation of the displacement x with this distance d is shown in Fig. 14.19.

a Define, for a wave, what is meant by
   i *displacement*, [1]
   ii *wavelength*. [1]

b Use Figs. 14.18 and 14.19 to determine, for the water wave,
   i the period T of vibration, [1]
   ii the wavelength λ, [1]
   iii the speed v. [2]

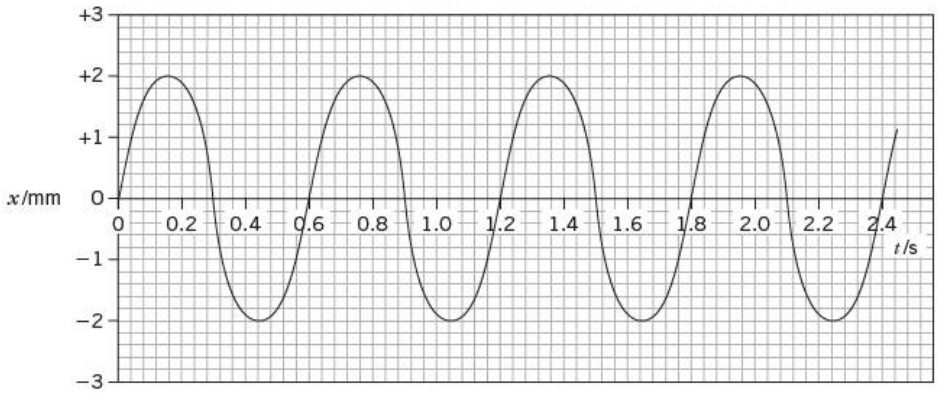


Fig. 14.18

www.iranino.co

IRANINOVIN

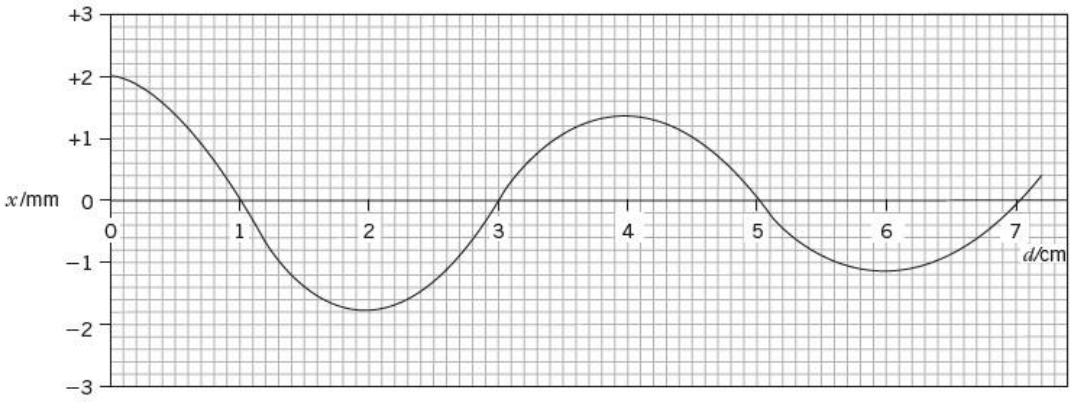


Fig. 14.19

c i Use Figs. 14.18 and 14.19 to state and explain whether the wave is losing power as it moves away from the source. *[2]*

ii Determine the ratio

$$\frac{\text{Intensity of wave at source}}{\text{Intensity of wave 6.0 cm from source}}$$ *[3]*

Cambridge International AS and A Level Physics, 9702/22 Oct/Nov 2010 Q 5

8 a State one property of electromagnetic waves that is **not** common to other transverse waves. *[1]*

b The seven regions of the electromagnetic spectrum are represented by blocks labelled A to G in Fig. 14.20. A typical wavelength for the visible region D is 500 nm.

i Name the principle radiations and give a typical wavelength for each of the regions B, E and F. *[3]*

ii Calculate the frequency corresponding to a wavelength of 500 nm. *[2]*

Cambridge International AS and A Level Physics, 9702/23 Oct/Nov 2012 Q 5

9 a By a reference to vibrations of the points on a wave and to its direction of energy transfer, distinguish between transverse waves and longitudinal waves. *[2]*

b The variation with distance x of the displacement y of a transverse wave is shown in Fig. 14.21.

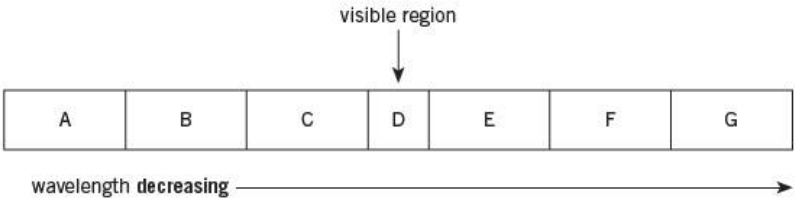


Fig. 14.20

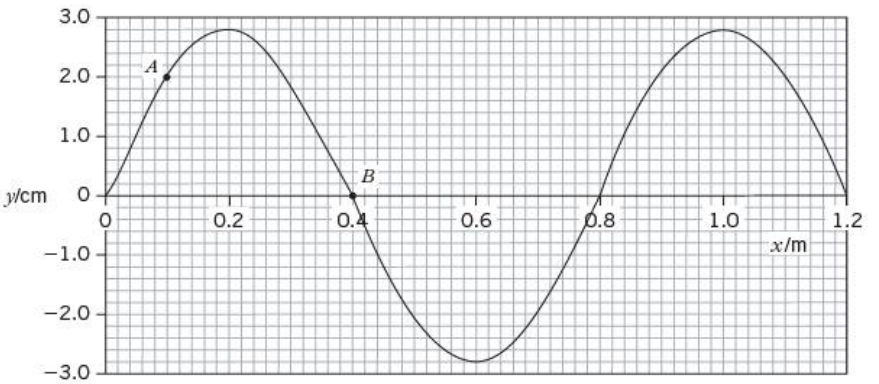


Fig. 14.21

www.iranino.co

IRANINOVIN

i Use Fig. 14.21 to determine
  1 the amplitude of the wave, [1]
  2 the phase difference between the points labelled A and B. [2]
ii Determine the amplitude of a wave with twice the intensity of that shown in Fig. 14.21. [1]

Cambridge International AS and A Level Physics, 9702/23 Oct/Nov 2011 Q 5

10 a A transverse progressive wave travels along a stretched string from left to right. The shape of part of the string at a particular instant is shown in Fig. 14.22.

The frequency of the wave is 15 Hz.

For this wave, use Fig. 14.22 to determine

i the amplitude, [1]
ii the phase difference between the points P and Q on the string, [1]
iii the speed of the wave. [2]

b The period of vibration of the wave is T.
The wave moves forward from the position shown in Fig. 14.22 for a time $0.25T$. Sketch the new position of the wave. [2]

Cambridge International AS and A Level Physics, 9702/23 May/June 2011 Q 6

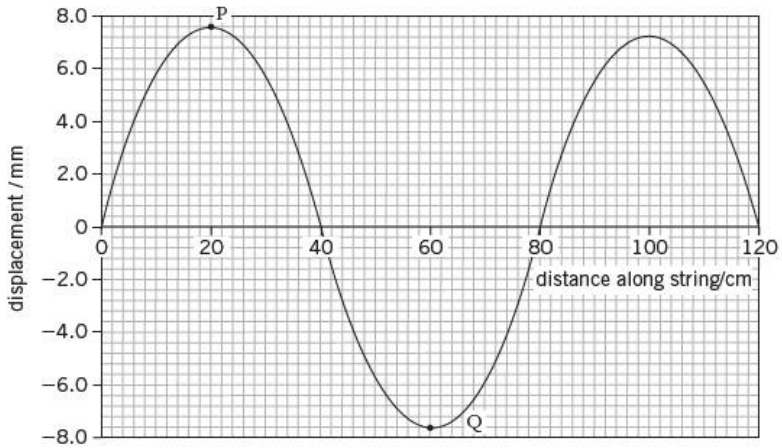


Fig. 14.22

www.iranino.co

IRANINOVIN

15 Superposition

By the end of this topic, you will be able to:

14.3 (b) determine the wavelength of sound using stationary waves

15.1 (a) explain and use the principle of superposition in simple applications
(b) show an understanding of experiments that demonstrate stationary waves using microwaves, stretched strings and air columns
(c) explain the formation of a stationary wave using a graphical method and identify nodes and antinodes

15.2 (a) explain the meaning of the term diffraction
(b) show an understanding of experiments which demonstrate diffraction including the diffraction of water waves in a ripple tank with both a wide gap and a narrow gap

15.3 (a) understand the terms interference and coherence
(b) show an understanding of experiments which demonstrate two-source interference using water ripples, light and microwaves
(c) understand the conditions required if two-source interference fringes are to be observed
(d) recall and solve problems using the equation $\lambda = ax/D$ for double-slit interference using light

15.4 (a) recall and solve problems using the formula $d \sin \theta = n\lambda$
(b) describe the use of a diffraction grating to determine the wavelength of light

Starting points

- Basic properties of waves such as reflection and refraction.
- Knowledge of the terms used to describe waves.
- Graphical representation of waves.
- The properties of waves, interference and diffraction are introduced in this topic.

15.3 Interference

Any moment now the unsuspecting fisherman in Figure 15.1 is going to experience the effects of interference. The amplitude of oscillation of his boat will be significantly affected by the two approaching waves and their interaction when they reach his position.



www.iranino.co

Figure 15.1

IRANINOVIN

If two or more waves overlap, the resultant displacement is the sum of the individual displacements. Remember that displacement is a vector quantity. The overlapping waves are said to **interfere**. This may lead to a resultant wave of either a larger or a smaller displacement than either of the two component waves.

Interference can be demonstrated in a ripple tank (see Topic 14) by using two point sources. Figure 15.2 shows such an interference pattern, and Figure 15.3 shows how it arises.

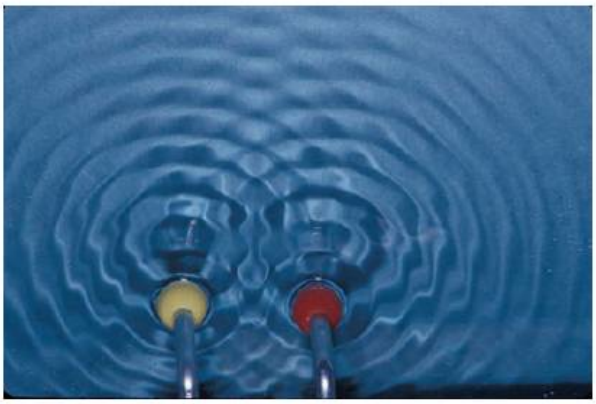


Figure 15.2 The interference pattern obtained using two point sources to produce circular water waves in a ripple tank

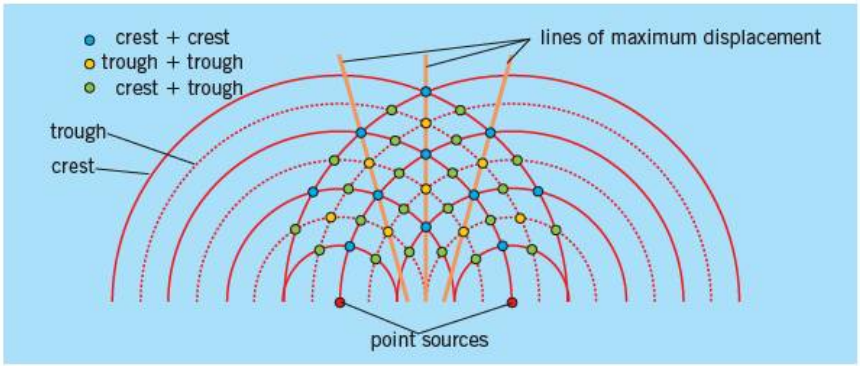


Figure 15.3 Two-source interference of circular waves

Figure 15.4 shows two waves arriving at a point at the same time. If they arrive **in phase** – that is, if their crests arrive at exactly the same time – they will interfere **constructively**. A resultant wave will be produced which has crests much higher than either of the two individual waves, and troughs which are much deeper. If the two incoming waves have the same frequency and equal amplitude A, the resultant wave produced by constructive interference has an amplitude of $2A$. The frequency of the resultant is the same as that of the incoming waves.

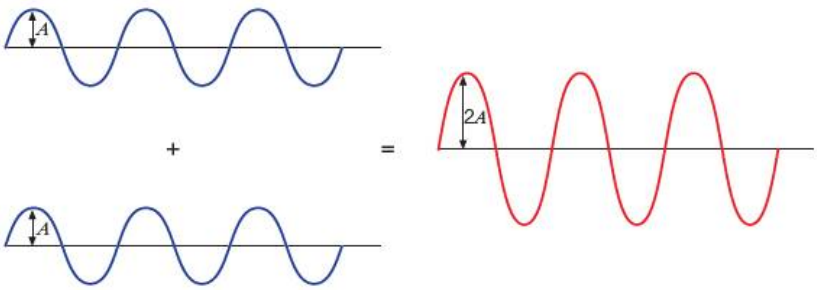


Figure 15.4 Constructive interference

If the two waves arrive **in antiphase** (with a phase difference of π radians or 180°) the peaks of one wave arrive at the same time as the troughs from the other and they will interfere **destructively**. The resultant wave will have a smaller amplitude. In the case shown in Figure 15.5, where the incoming waves have equal amplitude, the resultant wave has zero amplitude.

www.iranino.co

IRANINOVIN

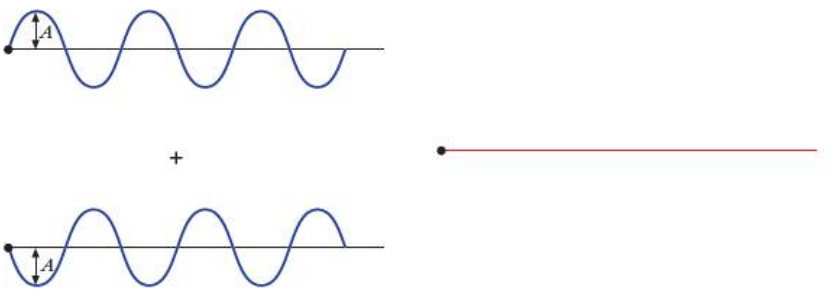


Figure 15.5 Destructive interference

This situation is an example of the **principle of superposition of waves**. The principle describes how waves, which meet at the same point in space, interact.

> The principle of superposition states that, when two or more waves meet at a point, the resultant displacement at that point is equal to the sum of the displacements of the individual waves at that point.

Because displacement is a vector, we must remember to add the individual displacements taking account of their directions. The principle applies to all types of wave.

If we consider the effect of superposition at a number of points in space, we can build up a pattern showing some areas where there is constructive interference, and hence a large wave disturbance, and other areas where the interference is destructive, and there is little or no wave disturbance.

Figure 15.6 illustrates the interference of waves from two point sources A and B. The point C is equidistant from A and B: a wave travelling to C from A moves through the same distance as a wave travelling to C from B (path difference is zero). If the waves started in phase at A and B, they will arrive in phase at C (phase difference is zero). They combine constructively, producing a maximum disturbance at C.

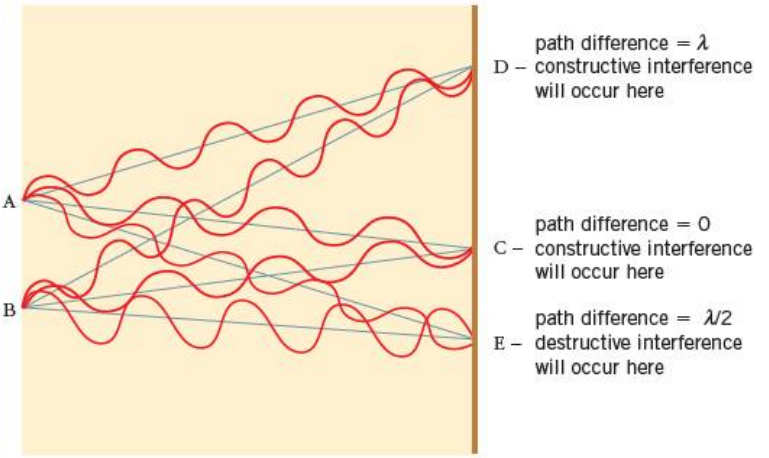


Figure 15.6 Producing an interference pattern

At other places, such as D, the waves will have travelled different distances from the two sources. There is a **path difference** between the waves arriving at D. If this path difference is a whole number of wavelengths (1λ, 2λ, 3λ, etc.) the waves arrive in phase and interfere constructively, producing maximum disturbance again. The equivalent **phase differences** between the waves are 2π, 4π, 6π, etc. or 360°, 720°, 1080°, etc. However, at places such as E, the path difference is an odd number of half-wavelengths ($\lambda/2$, $3\lambda/2$, $5\lambda/2$, etc.). The waves arrive at E in antiphase, and interference is destructive, producing a minimum resultant disturbance. The equivalent **phase differences** between the waves are π, 3π , 5π etc. or 180°, 540°, 900°, etc. The maxima and minima disturbances are called fringes. The collection of fringes produced by the superposition of overlapping waves is called an **interference pattern**. One is shown in Figure 15.2 for water waves.

www.iranino.co

IRANINOVIN

Producing an interference pattern with sound waves

Figure 15.7 shows an experimental arrangement to demonstrate interference with sound waves from two loudspeakers connected to the same signal generator and amplifier, so that each speaker produces a note of the same frequency. The demonstration is best carried out in the open air (on playing-fields, for example) to avoid reflections from walls, but it should be a windless day. Moving about in the space around the speakers, you pass through places where the waves interfere constructively and you can hear a loud sound. In places where the waves interfere destructively, the note is much quieter than elsewhere in the pattern.

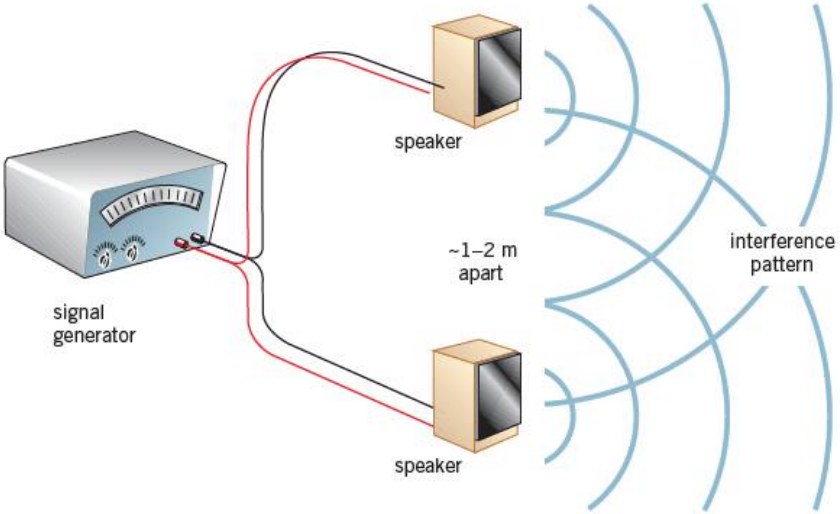


Figure 15.7 Demonstration of interference with sound waves

Producing an interference pattern with light waves

If you try to set up a demonstration with two separate light sources, such as car headlights, you will find that it is not possible to produce an observable interference pattern (Figure 15.8). A similar demonstration works with sound waves from two loudspeakers, each connected to separate signal generators. What has gone wrong?

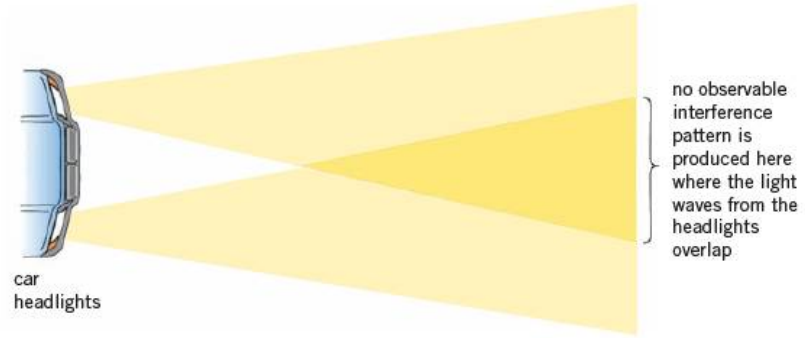


Figure 15.8 Failure of an interference demonstration with light

To produce an observable interference pattern, the two wave sources must have the same **single frequency**, not a mixture of frequencies as is the case for light from car headlights. They must also have a **constant phase relationship**. In the sound experiment, the waves from the two loudspeakers have the same frequency and a constant phase relationship because the loudspeakers are connected to the same oscillator and amplifier. If the waves emitted from the speakers are in phase when the experiment begins, they stay in phase for the whole experiment.

Wave sources which maintain a constant phase relationship are described as **coherent** sources.

www.iranino.co

IRANINOVIN

Coherent sources are illustrated in Figure 15.9.

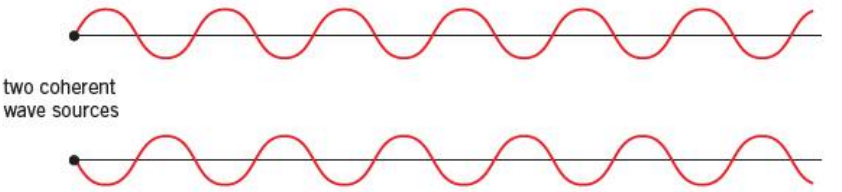


Figure 15.9 Coherent wave trains

Light is emitted from sources as a series of pulses or packets of energy. These pulses last for a very short time, about a nanosecond (10^{-9} s). Between each pulse there is an abrupt change in the phase of the waves. Waves from two separate sources may be in phase at one instant, but out of phase in the next nanosecond. The human eye cannot cope with such rapid changes, so the pattern is not observable. Separate light sources, even of the same frequency, produce incoherent waves (Figure 15.10).

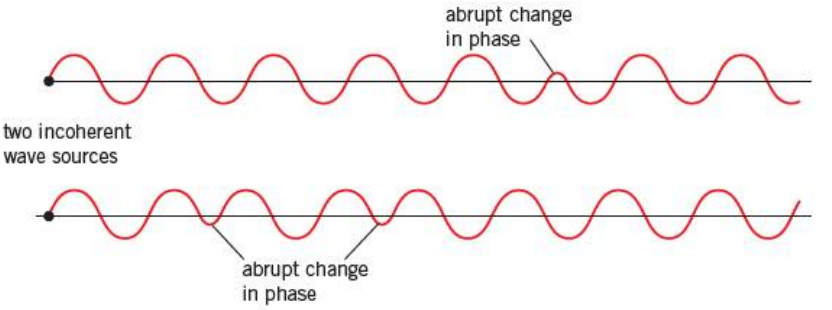


Figure 15.10 Incoherent light trains

To obtain observable interference patterns, it is not essential for the amplitudes of the waves from the two sources to be the same. However, if the amplitudes are not equal, a completely dark fringe will never be obtained, and the contrast of the pattern is reduced.

Young's double-slit experiment

In 1801 Thomas Young demonstrated how light waves could produce an interference pattern. The experimental arrangement is shown in Figure 15.11 (not to scale).

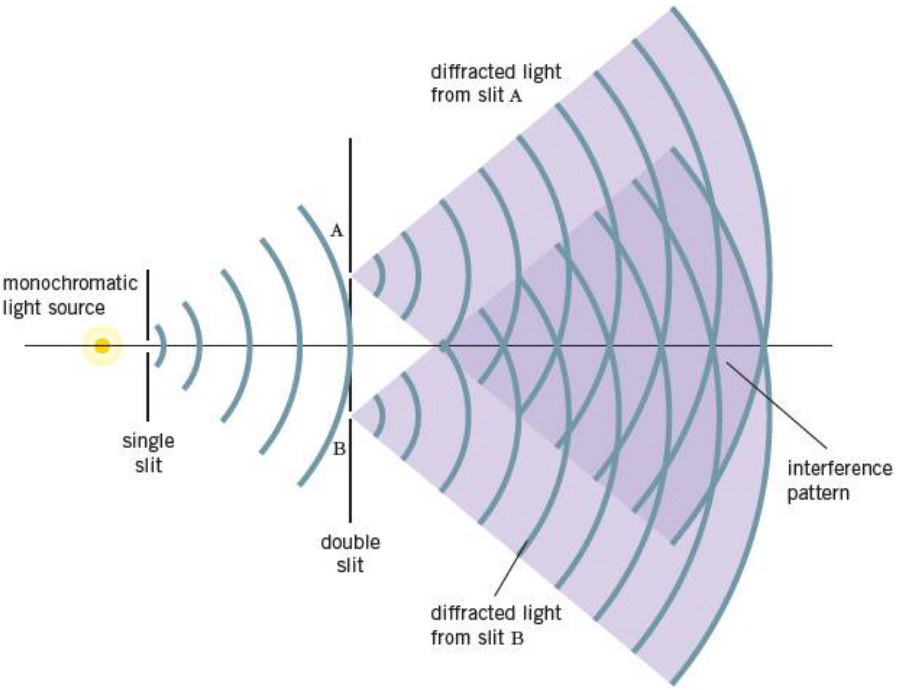


Figure 15.11 Young's double-slit experiment

www.iranino.co

IRANINOVIN

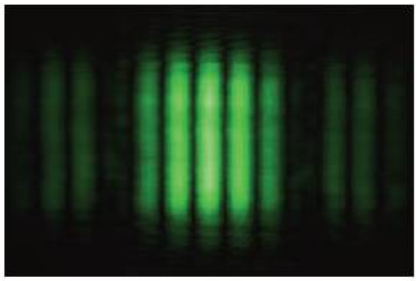


Figure 15.12 Fringe pattern in Young's experiment

A monochromatic light source (a source of one colour, and hence one wavelength λ) is placed behind a single slit to create a small, well defined source of light. Light from this source is diffracted at the slit (see Topic 15, page 126), producing two light sources at the double slit A and B. Because these two light sources originate from the same primary source, they are coherent and create a sustained and observable interference pattern, as seen in the photograph of the dark and bright interference fringes in Figure 15.12. Bright fringes are seen where constructive interference occurs – that is, where the path difference between the two diffracted waves from the sources A and B is $n\lambda$, where n is a whole number. Dark fringes are seen where destructive interference occurs. The condition for a dark fringe is that the path difference should be $(n + \frac{1}{2})\lambda$.

The distance x on the screen between successive bright fringes is called the fringe width. The fringe width is related to the wavelength λ of the light source by the equation

$$x = \lambda D/a \text{ or } \lambda = ax/D$$

where D is the distance from the double slit to the screen and a is the distance between the centres of the slits. Note that, because the wavelength of light is so small (of the order of 10^{-7} m), to produce observable fringes D needs to be large and a as small as possible. (This is another reason why you could never see an interference pattern from two sources such as car headlamps.) See the Example below.

Although Young's original double-slit experiment was carried out with light, the conditions for constructive and destructive interference apply for any two-source situation. The same formula applies for all types of wave, including sound waves, water waves and microwaves, provided that the fringes are detected at a distance of many wavelengths from the two sources.

Example

Calculate the observed fringe width for a Young's double-slit experiment using light of wavelength 600 nm and slits 0.50 mm apart. The distance from the slits to the screen is 0.80 m.

Using $x = \lambda D/a$, $x = 600 \times 10^{-9} \times 0.80/0.50 \times 10^{-3} = 9.6 \times 10^{-4}$ m

$= \mathbf{0.96\ mm}$

Now it's your turn

1 Calculate the wavelength of light which produces fringes of width 0.50 mm on a screen 60 cm from two slits 0.75 mm apart.

2 Radar waves of wavelength 50 mm are emitted from two aerials and create a fringe pattern 1.0 km from the aerials. Calculate the distance between the aerials if the fringe spacing is 80 cm.



Figure 15.13 Cello being bowed

www.iranino.co

White-light fringes

If the two slits in Young's experiment are illuminated with white light, each of the different wavelengths making up the white light produces its own fringe pattern. At the centre of the pattern, where the path difference for all waves is zero, there will be a white maximum with a black fringe on each side. Thereafter, the maxima and minima of the different colours overlap in such a way as to produce a pattern of coloured fringes. Only a few will be visible; a short distance from the centre so many wavelengths overlap that they combine to produce what is effectively white light again.

15.1 Stationary waves

The notes we hear from a cello are created by the vibrations of its strings (Figure 15.13). The wave patterns on the vibrating strings are called **stationary waves** (or **standing waves**). The waves in the air which carry the sound to our ears transfer energy and are therefore called **progressive waves**.

IRANINOVIN

Figure 15.14 shows a single transverse pulse travelling along a 'slinky' spring. The pulse is reflected when it reaches the fixed end. If a second pulse is sent along the slinky (Figure 15.15), the reflected pulse will pass through the outward-going pulse, creating a new pulse shape. Interference will take place between the outward and reflected pulses.

If the interval between outward pulses is reduced, a progressive wave is generated. When the wave reaches the fixed end, it is reflected. We now have two progressive waves of equal frequency and amplitude travelling in opposite directions on the same spring. The waves interfere, producing a wave pattern (Figure 15.16) in which the crests and troughs do not move. This pattern is called a **stationary** or **standing wave**, because it does not move.

A stationary wave is the result of interference between two waves of equal frequency and amplitude, travelling along the same line with the same speed but in opposite directions.

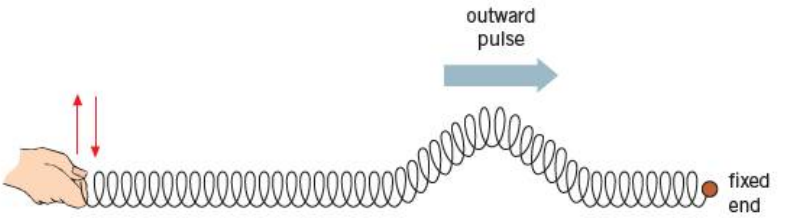


Figure 15.14 Single transverse pulse travelling along a slinky

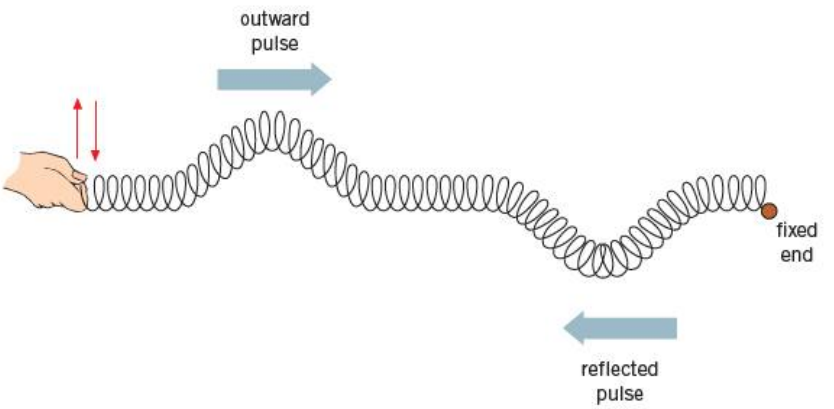


Figure 15.15 Reflected pulse about to meet an outward-going pulse

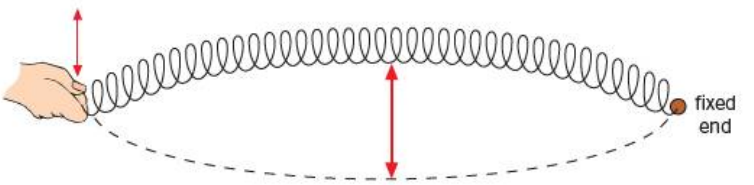


Figure 15.16 A stationary wave is created when two waves travelling in opposite directions interfere.

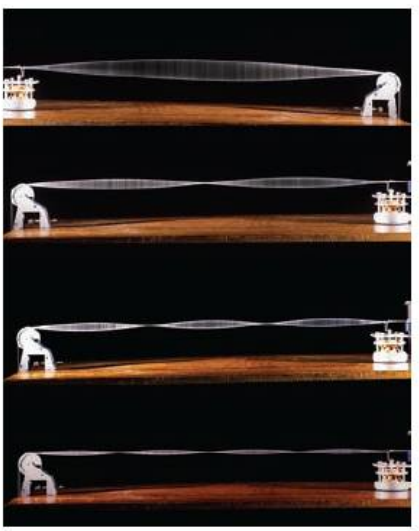


Figure 15.17 First four modes of vibration of a string

www.iranino.co

Stationary waves on strings

If a string is plucked and allowed to vibrate freely, there are certain frequencies at which it will vibrate. The amplitude of vibration at these frequencies is large. This is known as a **resonance** effect.

It is possible to investigate stationary waves in a more controlled manner using a length of string under tension and a vibrator driven by a signal generator. As the frequency of the vibrator is changed, different standing wave patterns are formed. Some of these are shown in Figure 15.17.

IRANINOVIN

Figure 15.18 shows the simplest way in which a stretched string can vibrate. The wave pattern has a single loop. This is called the **fundamental mode** of vibration, or the **first harmonic**. At the ends of the string there is no vibration. These points are called **nodes**. At the centre of the string, the amplitude is a maximum. A point of maximum amplitude is called an **antinode**. Nodes and antinodes do not move along the string.

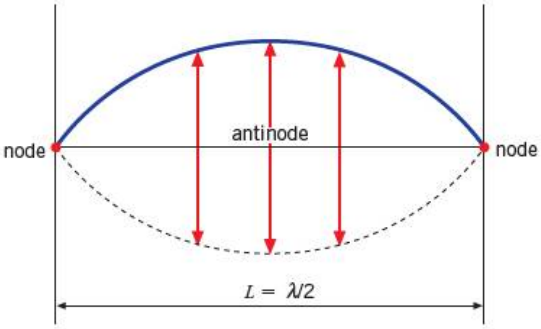


Figure 15.18 Fundamental mode of vibration of a stretched string

The wavelength λ of the standing wave in the fundamental mode is $2L$. From the wave equation $c = f\lambda$, the frequency f_1 of the note produced by the string vibrating in its fundamental mode is given by $f_1 = c/2L$, where c is the speed of the progressive waves which have interfered to produce the stationary wave.

Figure 15.19 shows the second mode of vibration of the string. The stationary wave pattern has two loops. This mode is sometimes called the **first overtone**, or the **second harmonic** (don't be confused!). The wavelength of this second mode is L. Applying the wave equation, the frequency f_2 is found to be c/L.

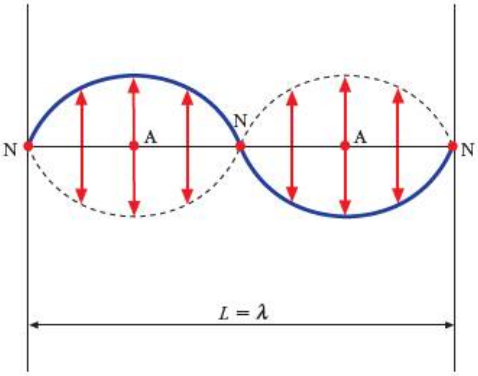


Figure 15.19 Second mode of vibration of a stretched string

Figure 15.20 shows the third mode (the second overtone, or third harmonic). This is a pattern with three loops. The wavelength is $2L/3$, and the frequency f_3 is $3c/2L$.

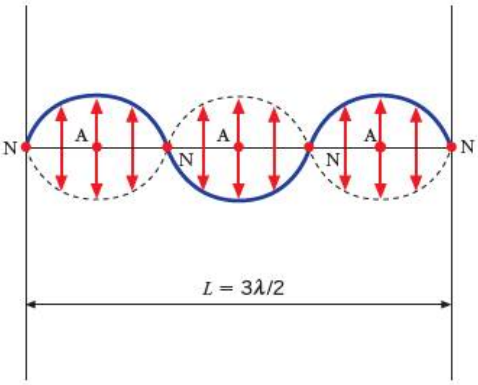


Figure 15.20 Third mode of vibration of a stretched string

The general expression for the frequency f_n of the nth mode (or the nth harmonic, or $(n - 1)$th overtone) is

$$f_n = nc/2L \qquad n = 1, 2, 3, \ldots$$

www.iranino.co

IRANINOVIN

The key features of a stationary wave pattern on a string, which distinguish it from a progressive wave, are as follows.

- The amplitude of vibration varies with position along the string: it is zero at a node, and maximum at an antinode. In a progressive wave, all points have the same amplitude.
- The nodes and antinodes do not move along the string, whereas in a progressive wave, the crests and troughs do move along it.
- Between adjacent nodes, all points of the stationary wave vibrate in phase. That is, all particles of the string are at their maximum displacement at the same instant. In a progressive wave, phase varies continuously along the wave. All neighbouring points along the wave are out of phase with each other, they reach their maximum displacement at different times.
- There are only certain frequencies of stationary waves possible on the string. The allowed frequencies depend on the length L of the string and c, the speed of the progressive waves that form the stationary wave.

Stationary waves explained by interference

Let us explain the formation of a stationary wave using the principle of superposition. The set of graphs in Figure 15.21 represents two progressive waves of equal amplitude and frequency travelling in opposite directions. The red-dashed line is travelling from left to right, and the blue-dashed line is going from right to left. When the string is clamped (at the right-hand end), the effect of the clamp is to change the phase of the reflected wave by π rad (180°).

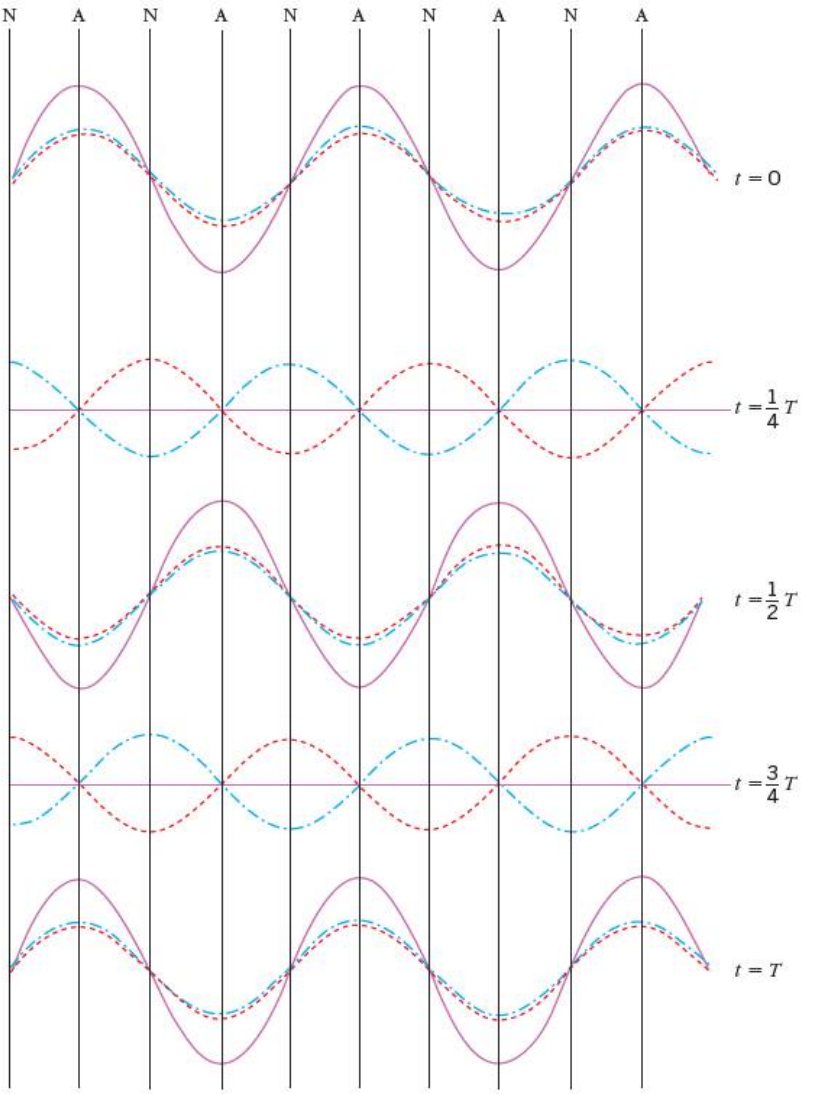


Figure 15.21 Formation of a stationary wave by superposition of two progressive waves travelling in opposite directions

www.iranino.co

IRANINOVIN

The top graph catches the waves at an instant at which they are in phase. Superposition gives the purple curve, which has twice the amplitude of either of the progressive graphs. The second graph is the situation a quarter of a period (cycle) later, when the two progressive waves have each moved a quarter of a wavelength in opposite directions. This has brought them to a situation where the movement of one relative to the other is half a wavelength, so that the waves are exactly out of phase. The resultant, obtained by superposition, is zero everywhere. In the third graph, half a period from the start, the waves are again in phase, with maximum displacement for the resultant. The process continues through the fourth graph, showing the next out-of-phase situation, with zero displacement of the resultant everywhere. Finally, the fifth graph, one period on from the first, brings the waves into phase again.

We can see how there are some positions, the nodes N, where the displacement of the resultant is zero *throughout* the cycle. The displacement of the resultant at the antinodes A fluctuates from a maximum value when the two progressive waves are in phase to zero when they are out of phase.

Stationary waves in air

Figure 15.22 shows an experiment to demonstrate the formation of stationary waves in air. A fine, dry powder (such as cork dust or lycopodium powder) is sprinkled evenly along the transparent tube. A loudspeaker powered by a signal generator is placed at the open end. The frequency of the sound from the loudspeaker is gradually increased. At certain frequencies, the powder forms itself into evenly spaced heaps along the tube. A stationary wave has been set up in the air, caused by the interference of the sound wave from the loudspeaker and the wave reflected from the closed end of the tube. At nodes (positions of zero amplitude) there is no disturbance, and the powder can settle into a heap. At antinodes the disturbance is at a maximum, and the powder is dispersed.

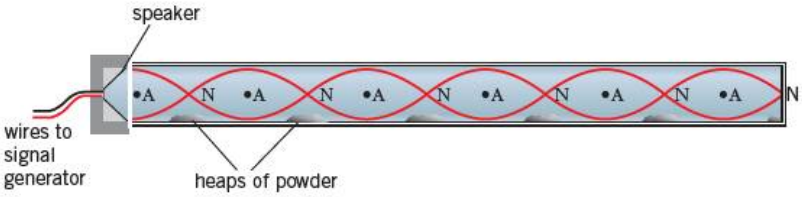


Figure 15.22 Demonstration of stationary waves in air

For stationary waves in a closed pipe, the air cannot move at the closed end, and so this must always be a node N. However, the open end is a position of maximum disturbance, and this is an antinode A. (In fact, the antinode is slightly outside the open end. The distance of the antinode from the end of the tube is called the end-correction. The value of the end-correction depends on the diameter of the tube.)

Figure 15.23 shows the simplest way in which the air in a pipe, closed at one end, can vibrate. Figure 15.23a illustrates the motion of some of the air particles in the tube. Their amplitude of vibration is zero at the closed end, and increases with distance up the tube to a maximum at the open end. This representation is tedious to draw, and Figure 15.23b is the conventional way of showing the amplitude of vibration: the amplitudes along the axis of the tube are plotted as a continuous curve. One danger of using diagrams like Figure 15.23b is that they give the impression that the sound wave is transverse rather than longitudinal. So be warned! The mode illustrated in Figure 15.23 is the fundamental mode (the first harmonic). The wavelength of this stationary wave (ignoring the end-correction) is $4L$, where L is the length of the pipe. Using the wave equation, the frequency f_1 of the fundamental mode is given by $f_1 = c/4L$, where c is the speed of the sound in air.

www.iranino.co

IRANINOVIN

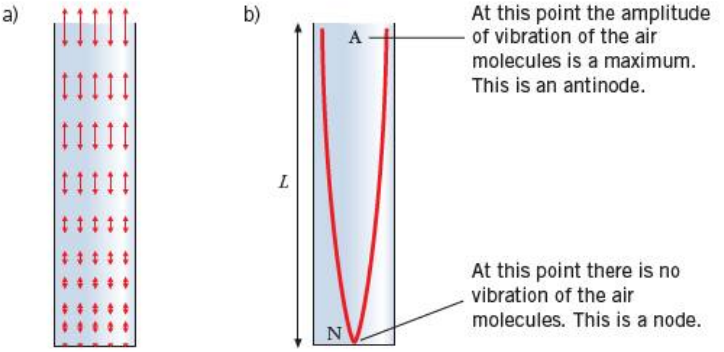


Figure 15.23 Fundamental mode of vibration of air in a closed pipe

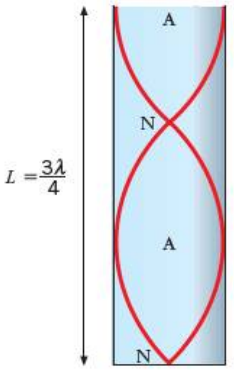


Figure 15.24 Second mode of vibration of air in a closed pipe

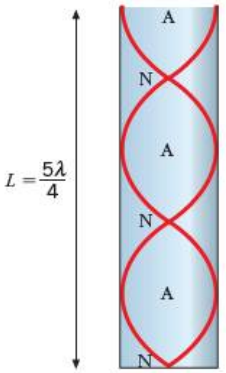


Figure 15.25 Third mode of vibration of air in a closed pipe

Other modes of vibration are possible. Figures 15.24 and 15.25 show the second mode (the first overtone, or second harmonic) and the third mode (the second overtone, or third harmonic). The corresponding wavelengths are $4L/3$ and $4L/5$, and the frequencies are $f_2 = 3c/4L$ and $f_3 = 5c/4L$.

The general expression for the frequency f_n of the nth mode of vibration of the air in the closed tube (the nth harmonic, or the $(n - 1)$th overtone) is

$$f_n = \frac{(2n - 1)c}{4L}$$

This is another example of resonance. The particular frequencies at which stationary waves are obtained in the pipe are the resonant frequencies of the pipe. The particular frequencies of stationary waves possible in the pipe depend on the length L of the pipe and c, the speed of sound in air.

Stationary waves using microwaves

Stationary waves can be demonstrated using microwaves. A source of microwaves faces a metal reflecting plate, as shown in Figure 15.26. A small detector is placed between source and reflector. The reflector is moved towards or away from the source until the signal picked up by the detector fluctuates regularly as it is moved slowly back and forth. The minima are nodes of the stationary wave pattern, and the maxima are antinodes. The distance moved by the detector between successive nodes is half the wavelength of the microwaves.

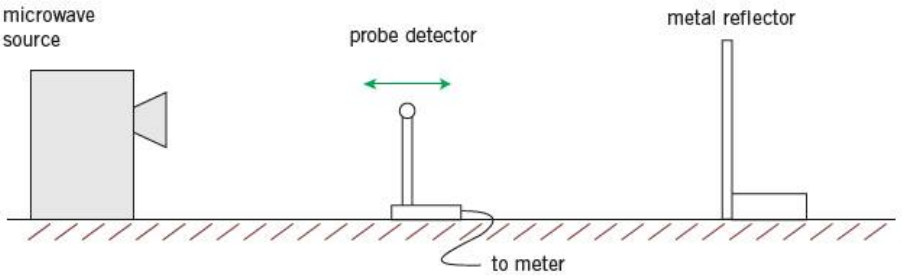


Figure 15.26 Using microwaves to demonstrate stationary waves

14.3 Measuring the speed of sound using stationary waves

The principle of resonance in a tube closed at one end can be used to measure the speed of sound in air. A glass tube is placed in a cylinder of water. By raising the tube, the length of the column of air can be increased. A vibrating tuning fork of known frequency

www.iranino.co

IRANINOVIN

f is held above the open end of the glass tube, causing the air in it to vibrate. The tube is gradually raised, increasing the length of the air column. At a certain position the note becomes much louder. This is known as the first position of resonance, and occurs when a stationary wave corresponding to the fundamental mode is established inside the tube. The length L_1 of the air column is noted. The tube is raised further until a second resonance position is found. This corresponds to the second mode of vibration. The length L_2 at this position is also noted. The two resonance positions are illustrated in Figure 15.27.

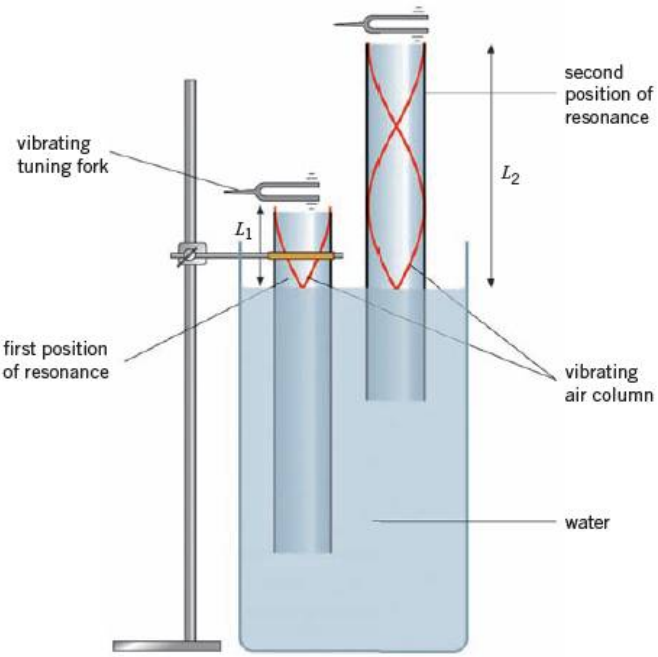


Figure 15.27 Speed of sound by the resonance tube method

At the first position of resonance, $\lambda/4 = L_1 + e$, where e is the end-correction of the tube (to allow for the fact that the antinode is slightly above the open end of the tube). At the second position of resonance, $3\lambda/4 = L_2 + e$. By subtracting these equations, we can eliminate e to give

$$\lambda/2 = L_2 - L_1$$

From the wave equation, the speed of sound c is given by $c = f\lambda$. Thus

$$c = 2f(L_2 - L_1)$$

Figure 15.28 illustrates a method of measuring the wavelength and the speed of sound using stationary waves in free air, rather than in a resonance tube. The signal generator and loudspeaker produce a note of known frequency f.

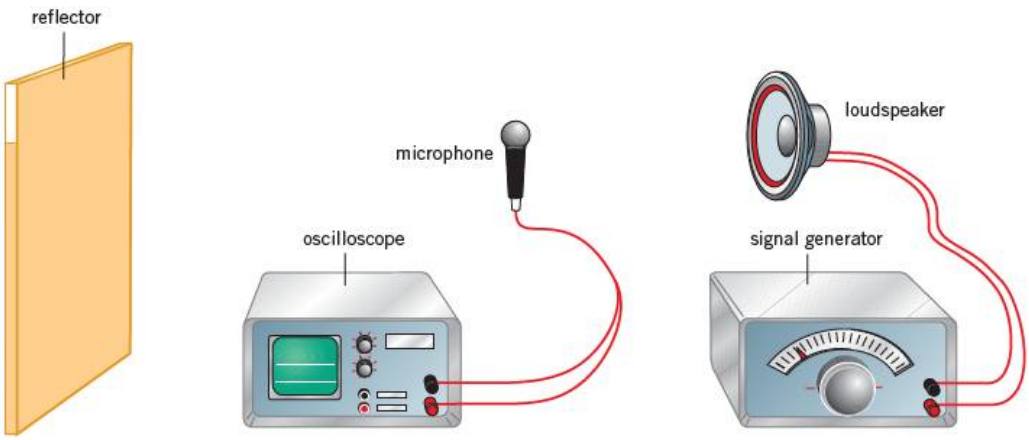


Figure 15.28 Speed of sound using stationary waves in free air

www.iraninovin.com

IRANINOVIN

The reflector is moved slowly back and forth until the trace on the oscilloscope has a minimum amplitude. When this happens, a stationary wave has been set up with one of its nodes in the same position as the microphone. The microphone is now moved along the line between the loudspeaker and the reflector. The amplitude of the trace on the oscilloscope will increase to a maximum, and then decrease to a minimum. The microphone has been moved from one node, through an antinode, to the next node. The distance d between these positions is measured. We know that the distance between nodes is $\lambda/2$. The speed of sound can then be calculated using $c = f\lambda$, giving $c = 2fd$.

Examples

1 A string 75 cm long is fixed at one end. The other end is moved up and down with a frequency of 15 Hz. This frequency gives a stationary wave pattern with three complete loops on the string. Calculate the speed of the progressive waves which have interfered to produce the stationary wave.

The three-loop pattern corresponds to the situation where the length L of the string is $3\lambda/2$ (see Figure 15.20). The wavelength λ is thus $2 \times 0.75/3 = 0.50$ m. The frequency of the wave is 15 Hz, so by the wave equation $c = f\lambda = 15 \times 0.50 =$ **7.5 m s^{-1}**.

2 Find the fundamental frequency and first two overtones for an organ pipe which is 0.17 m long and closed at one end. The speed of sound in air is 340 m s^{-1}.

The frequencies of the fundamental and first two overtones of a tube of length L, closed at one end, are $c/4L$, $3c/4L$ and $5c/4L$ (see Figures 15.23–15.25). The frequencies are thus $340/4 \times 0.17 =$ **500 Hz**, $3 \times 340/4 \times 0.17 =$ **1500 Hz** and $5 \times 340/4 \times 0.17 =$ **2500 Hz**.

Now it's your turn

3 A violin string vibrates with a fundamental frequency of 440 Hz. What are the frequencies of its first two overtones?

4 The speed of waves on a certain stretched string is 48 m s^{-1}. When the string is vibrated at frequency of 64 Hz, stationary waves are set up. Find the separation of successive nodes in the stationary wave pattern.

5 You can make an empty lemonade bottle resonate by blowing across the top. What fundamental frequency of vibration would you expect for a bottle 25 cm deep? The speed of sound in air is 340 m s^{-1}.

6 A certain organ pipe, closed at one end, can resonate at consecutive frequencies of 640 Hz, 896 Hz and 1152 Hz. Deduce its fundamental frequency.

15.2 & 15.4 Diffraction

When waves pass through a narrow gap, they spread out. This spreading out is called **diffraction**. The extent of diffraction depends on the width of the gap compared with the wavelength. It is most noticeable if the width of the gap is approximately equal to the wavelength. Diffraction can be demonstrated in a ripple tank (see Topic 14) by using the apparatus shown in Figures 14.6 and 14.7.

Diffraction is illustrated in Figure 15.29. Note that diffraction may also occur at an edge.

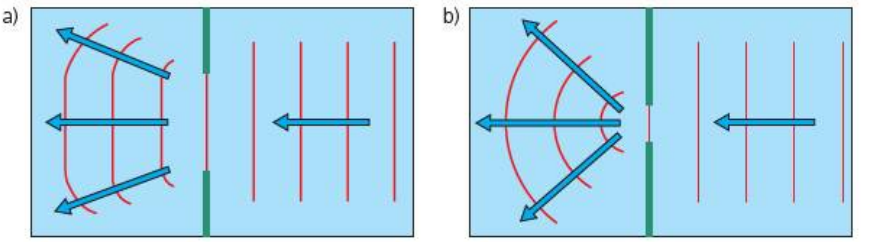


Figure 15.29 Ripple tank pattern showing diffraction at a) a wide gap, b) a narrow gap

Diffraction is defined as the spreading of a wave into regions where it would not be seen if it moved only in straight lines after passing through a narrow slit or past an edge.

www.iranino.co

IRANINOVIN

Although we often hear the statement 'light travels in straight lines', there are occasions when this appears not to be the case. Newton tried to explain the fact that when light travels through an aperture, or passes the edge of an obstacle, it deviates from the straight-on direction and appears to spread out. We have seen from the ripple tank demonstration (Figure 15.29) that water waves spread out when they pass through an aperture. This shows that water waves can be diffracted. The fact that light undergoes diffraction is powerful evidence that light has wave properties. Newton's attempt to explain diffraction was not, in fact, based on a wave theory of light. The Dutch scientist Christian Huygens, a contemporary of Newton, favoured the wave theory, and used it to account for reflection, refraction and diffraction. (It was not until 1815 that the French scientist Augustin Fresnel developed the wave theory of light to explain diffraction in detail.)

The experiment illustrated in Figure 15.29 shows that the degree to which waves are diffracted depends upon the size of the obstacle or aperture and the wavelength of the wave. The greatest effects occur when the wavelength is about the same size as the aperture. The wavelength of light is very small (green light has wavelength 5×10^{-7} m), and therefore diffraction effects can be difficult to detect.

Huygens' explanation of diffraction

If we let a single drop of water fall into a ripple tank, it will create a circular wavefront which will spread outwards from the disturbance (Figure 14.8). Huygens put forward a wave theory of light which was based on the way in which circular wavefronts advance. He suggested that, at any instant, all points on a wavefront could be regarded as secondary disturbances, giving rise to their own outward-spreading circular wavelets. The envelope, or tangent curve, of the wavefronts produced by the secondary sources gives the new position of the original wavefront. This construction is illustrated in Figure 15.30 for a circular wavefront.

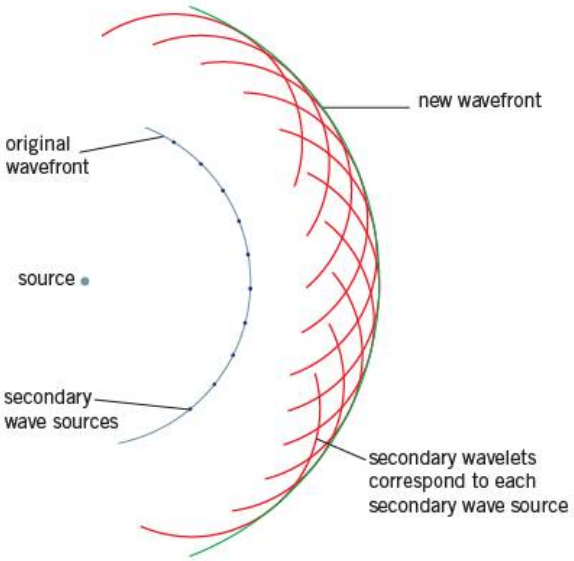


Figure 15.30 Huygens' construction for a circular wavefront

Now think about a plane (straight) wavefront. If the wavefront is restricted in any way, for example by passing through an aperture in the form of a slit, some of the wavelets making up the wavefront are removed, causing the edges of the wavefront to be curved. If the wavelength is small compared with the size of the aperture, the wavefronts which pass through the aperture show curvature only at their ends, and the diffraction effect is relatively small. If the aperture is comparable with the wavelength, the diffracted wavefronts become circular, centred on the slit. Note that there is no change of wavelength with diffraction. This effect is illustrated in Figure 15.29.

Figure 15.31 shows the diffraction pattern created by a single slit illuminated by monochromatic light. The central region of the pattern is a broad, bright area with

www.iranino.co

IRANINOVIN

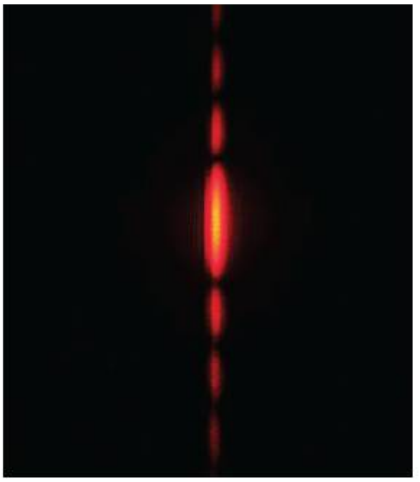


Figure 15.31 Diffraction of light at a single slit

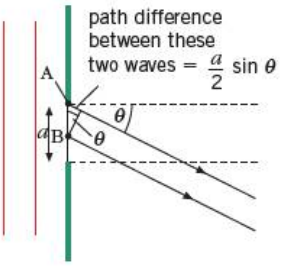


Figure 15.32 Light leaving a single slit

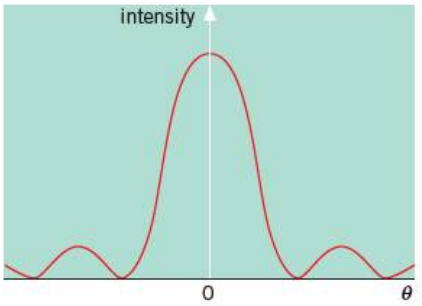


Figure 15.33 Light intensity graph for single slit diffraction

narrow, dark fringes on either side. Beyond these is a further succession of bright and dark areas. The bright areas become less and less intense as we move away from the centre.

This single-slit diffraction pattern has many features that we associate with an interference pattern. But how can a single slit produce an interference-type pattern? The explanation that follows is based on Huygens' wavelet idea.

Figure 15.32 shows plane wavefronts arriving at a single slit of width a. Each point on the wavefront passing through the slit can be considered to be a source of secondary wavelets.

One such source is at A, at the top edge of the slit, and a second is at B, at the centre of the slit, a distance $a/2$ along the wavefront from A. These two sources behave like the sources in a two-source interference experiment. The wavelets spreading out from these points overlap and create an interference pattern. In the straight-on direction, there is no path difference between the waves from A and B. Constructive interference occurs in this direction, giving a bright fringe in the centre of the pattern. To either side of the central fringe there are directions where the path difference between the waves from A and B is an odd number of half-wavelengths. This is the condition for destructive interference, resulting in dark fringes. The condition for constructive interference is that the path difference should be a whole number of wavelengths. Thus, the dark fringes alternate with bright fringes.

This argument can be applied to the whole of the slit. Every wavelet spreading out from a point in the top half of the slit can be paired with one coming from a point $a/2$ below it in the lower half of the slit. When wavelets from points right across the aperture are added up, we find that there are certain directions in which constructive interference occurs, and other directions in which the interference is destructive. Figure 15.33 is a graph of the intensity of the diffraction pattern as a function of the angle θ at which the light is viewed. It shows that most of the intensity is in the central area, and that this is flanked by dark and bright fringes.

We can use Figure 15.32 to derive an expression for the angle at which the first dark fringe is obtained. Remember that the condition for destructive interference is that the path difference between the two rays should be half a wavelength. The path difference between the two rays shown in Figure 15.32 is $\frac{1}{2}a\sin\theta$. If this is to be $\frac{1}{2}\lambda$, we have

$$\sin\theta = \lambda/a$$

This is the condition to observe the first dark fringe at angle θ. More generally,

$$\sin\theta = \frac{n\lambda}{a}$$

where n is a whole number called the *order* of the dark fringe being considered, counting outwards from the centre.

Although we have been concentrating on a diffraction pattern obtained with light, the derivation above applies to any type of wave passing through a rectangular aperture.

The wavelength of light is generally small compared with the width of slits or other apertures, so the diffraction angle θ is also small. Provided that θ is only a few degrees (less than about 5°), the approximation $\sin\theta = \theta$ may be used (remember that θ must be in radians!). Very often, the single-slit diffraction equation for light is expressed in the form

$$\theta = \frac{n\lambda}{a}$$

making use of the $\sin\theta$ approximation. But take care! This approximate form may not apply for the diffraction of other types of wave, such as sound or water waves, where the wavelength may be closer in magnitude to the width of the aperture, and diffraction angles are larger.

www.iranino.co

IRANINOVIN

Example

Calculate the angle between the centre of the diffraction pattern and the first minimum when light of wavelength 600 nm passes through a slit 0.10 mm wide.

Using $a \sin\theta = n\lambda$, we have $\sin\theta = n\lambda/a$. Substituting, $\sin\theta = 1 \times 6.0 \times 10^{-7}/1.0 \times 10^{-4}$ (don't forget to convert the nm and mm to m) = 0.0060, and $\theta = \mathbf{0.34°}$

Using the $\sin\theta = \theta$ approximation, we would have obtained $\theta = n\lambda/a = 1 \times 6.0 \times 10^{-7}/1.0 \times 10^{-4} = \mathbf{0.0060\,rad}$ (which is equal to 0.34°).

Now it's your turn

7 Calculate the angle between the centre of the diffraction pattern and the first minimum when a sound wave of wavelength 1.0 m passes through a door 1.2 m wide.

8 Calculate the wavelength of water waves which, on passing through a gap 50 cm wide, create a diffraction pattern such that the angle between the centre of the pattern and the second-order minimum is 60°.

The diffraction grating

A **diffraction grating** is a plate on which there is a very large number of parallel, identical, very closely spaced slits. If monochromatic light is incident on this plate, a pattern of narrow bright fringes is produced (Figure 15.34).

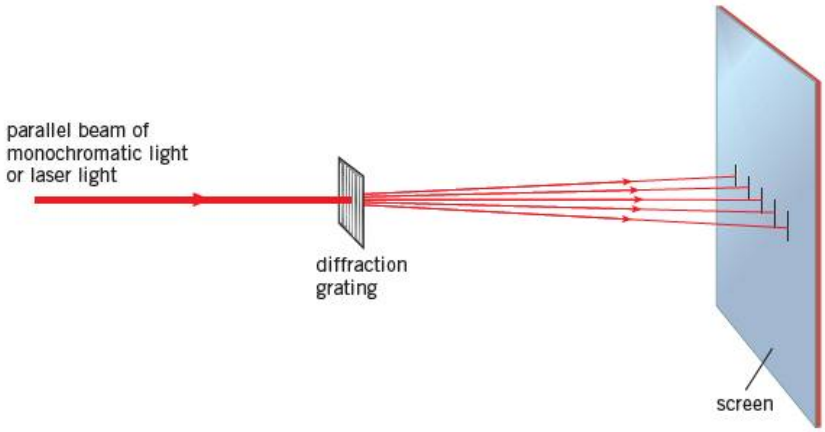


Figure 15.34 Arrangement for obtaining a fringe pattern with a diffraction grating

Although the device is called a *diffraction* grating, we shall use straightforward superposition and interference ideas in obtaining an expression for the angles at which the maxima of intensity are obtained.

Figure 15.35 shows a parallel beam of light incident normally on a diffraction grating in which the spacing between adjacent slits is d. Consider first rays 1 and 2 which are incident on adjacent slits. The path difference between these rays when they emerge at an angle θ is $d \sin\theta$. To obtain constructive interference in this direction from these two rays, the condition is that the path difference should be an integral number of wavelengths. The path difference between rays 2 and 3, 3 and 4, and so on, will also be $d \sin\theta$. The condition for constructive interference is the same. Thus, the condition for a maximum of intensity at angle θ is

$$d \sin\theta = n\lambda$$

where λ is the wavelength of the monochromatic light used, and n is an integer.

www.iranino.co

IRANINOVIN

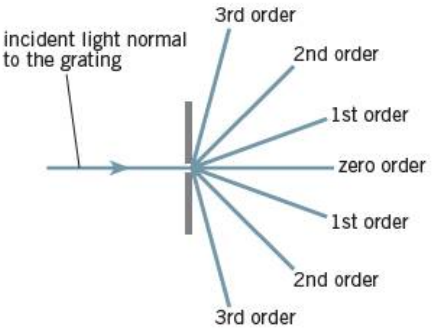


Figure 15.36 Maxima in the diffraction pattern of a diffraction grating

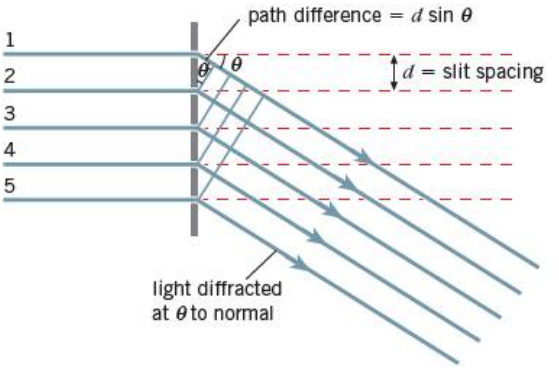


Figure 15.35

When $n = 0$, $\sin\theta = 0$ and θ is also zero; this gives the straight-on direction, or what is called the zero-order maximum. When $n = 1$, we have the first-order diffraction maximum, and so on (Figure 15.36).

Example

Monochromatic light is incident normally on a grating with 7.00×10^5 lines per metre. A second-order maximum is observed at an angle of diffraction of 40.0°. Calculate the wavelength of the incident light.

The slits on a diffraction grating are created by drawing parallel lines on the surface of the plate. The relationship between the slit spacing d and the number N of lines per metre is $d = 1/N$. For this grating, $d = 1/7.00 \times 10^5 = 1.43 \times 10^{-6}$ m. Using $n\lambda = d \sin\theta$, $\lambda = (d/n) \sin\theta = (1.43 \times 10^{-6}/2) \sin 40.0° =$ **460 nm**.

Now it's your turn

9 Monochromatic light is incident normally on a grating with 5.00×10^5 lines per metre. A third-order maximum is observed at an angle of diffraction of 78.0°. Calculate the wavelength of the incident light.

10 Light of wavelength 5.90×10^{-7} m is incident normally on a diffraction grating with 8.00×10^5 lines per metre. Calculate the diffraction angles of the first- and second-order diffraction images.

11 Light of wavelength 590 nm is incident normally on a grating with spacing 1.67×10^{-6} m. How many orders of diffraction maxima can be obtained?

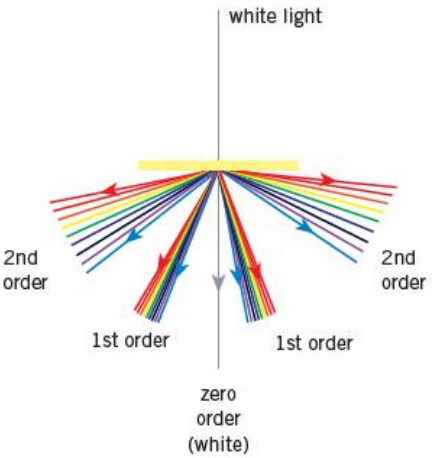


Figure 15.37 Production of the spectrum of white light with a diffraction grating

The diffraction grating with white light

If white light is incident on a diffraction grating, each wavelength λ making up the white light is diffracted by a different amount, as described by the equation $n\lambda = d \sin\theta$. Red light, because it has the longest wavelength in the visible spectrum, is diffracted through the largest angle. Blue light has the shortest wavelength, and is diffracted the least. Thus, the white light is split into its component colours, producing a continuous spectrum (Figure 15.37). The spectrum is repeated in the different orders of the diffraction pattern. Depending on the grating spacing, there may be some overlapping of different orders. For example, the red component of the first-order image may overlap with the blue end of the second-order spectrum.

An important use of the diffraction grating is in a **spectrometer**, a piece of apparatus used to investigate spectra. By measuring the angle at which a particular diffracted image appears, the wavelength of the light producing that image may be determined.

www.iranino.co

IRANINOVIN

Summary

- The principle of superposition of waves states that, when waves meet at the same point in space, the resultant displacement is given by the sum of the displacements of the individual waves.
- Constructive interference is obtained when the waves that meet are completely in phase, so that the resultant wave is of greater amplitude than any of its constituents.
- Destructive interference is obtained when the waves that meet are completely out of phase (in antiphase).
- To produce a sustained and observable interference pattern the sources must be monochromatic and coherent (have a constant phase relationship).
- Coherent sources have a constant phase difference between the vibrations of the sources.
- Young's double-slit experiment:
  - condition for constructive interference: path difference = $n\lambda$
  - condition for destructive interference: path difference $(n + \frac{1}{2})\lambda$
  - fringe width $x = \lambda D/a$, where a is the separation of the source slits and D is the distance of the screen from the slits.
- A stationary wave is the result of interference between two progressive waves of equal frequency and similar amplitude travelling along the same line with the same speed, but in opposite directions.
- Points of zero amplitude on a stationary wave are called nodes; points of maximum amplitude are called antinodes.
- For stationary waves on a stretched string, frequency f_n of the nth mode is given by $f_n = nc/2L$, where c is the speed of progressive waves on the string and L is the length of the string.
- For stationary waves in a gas in a tube closed at one end, frequency f_n of the nth mode is given by $f_n = (2n - 1)c/4L$, where c is the speed of sound in air and L is the length of the tube.
- Diffraction is the spreading out of waves after passing through an aperture or meeting the edge of an obstacle. It is most obvious when the size of the aperture and the wavelength of the wave are approximately the same.
- Properties of wave motion (diffraction and interference) can be observed in a ripple tank.
- Interference and diffraction of light is evidence that light has wave properties.
- Single slit diffraction: the condition for minima is $a \sin \theta = n\lambda$, where a is the width of the slit.
- The condition for a diffraction maximum in a diffraction grating pattern is $d \sin \theta = n\lambda$, where d is the grating spacing, θ is the angle at which the diffraction maximum is observed, n is an integer (the order of the image), and λ is the wavelength of the light.

Examination style questions

1 Compare a two-source experiment to demonstrate the interference of sound waves with a Young's double-slit experiment using light. What are the similarities and differences between the two experiments?

2 **a** Explain the term coherence as applied to waves from two sources.
b Describe how you would produce two coherent sources of light.
c A double-slit interference pattern is produced using slits separated by 0.45 mm, illuminated with light of wavelength 633 nm from a laser. The pattern is projected on to a wall 2.50 m from the slits. Calculate the fringe separation.

3 Fig. 15.38 shows the arrangement for obtaining interference fringes in a Young's double-slit experiment. Describe and explain what will be seen on the screen if the arrangement is altered in each of the following ways:
a the slit separation a is halved,
b the distance D from slits to screen is doubled,
c the monochromatic light source is replaced with a white-light source.

www.iranino.co

IRANINOVIN

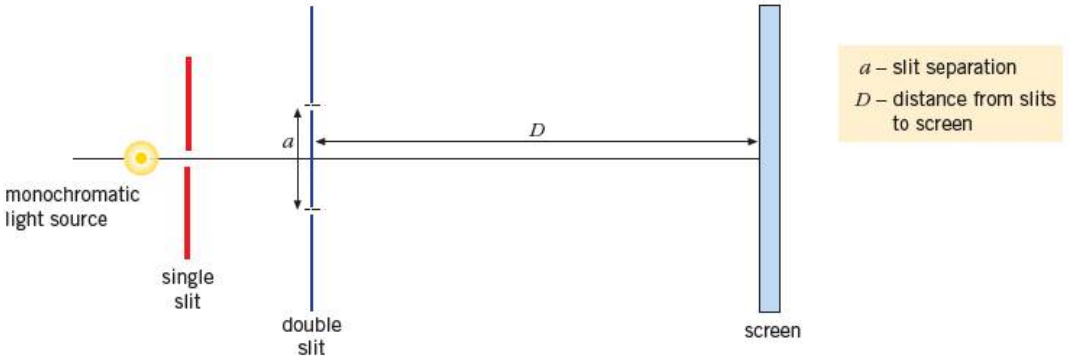


Fig. 15.38

4 State and explain four ways in which stationary waves differ from progressive waves.

5 A source of sound of frequency 2000 Hz is placed in front of a flat wall. When a microphone is moved away from the source towards the wall, a series of maxima and minima are detected.

- a Explain what has happened to create these maxima and minima.
- b The speed of sound in air is 340 m s^{-1}. Calculate the distance between successive minima.

6 A string is stretched between two fixed supports separated by 1.20 m. Stationary waves are generated on the string. It is observed that two stationary wave frequencies are 180 Hz and 135 Hz; there is no resonant frequency between these two. Calculate:

- a the speed of progressive waves on the stretched string,
- b the lowest resonant frequency of the string.

7 Blue and red light, with wavelengths 450 nm and 650 nm respectively, is incident normally on a diffraction grating which has 4.0×10^5 lines per metre.

- a Calculate the grating spacing.
- b Calculate the angle between the second-order maxima for these wavelengths.
- c For each wavelength, find the maximum order that can be observed.

8 Discuss any difference between the interference patterns formed by

- a two parallel slits 1 μm apart,
- b a diffraction grating with grating spacing 1 μm, when illuminated with monochromatic light.

9 Light of wavelength 633 nm passes through a slit 50 μm wide. Calculate the angular separation between the central maximum and the first minimum of the diffraction pattern.

10 A string is stretched between two fixed supports 3.5 m apart. Stationary waves are generated by disturbing the string. One possible mode of vibration of the stationary waves is shown in Fig. 15.39. The nodes and antinodes are labelled N and A respectively.

- a Distinguish between a node and an antinode in a stationary wave.

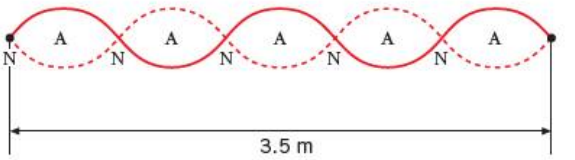


Fig. 15.39

- b State the phase difference between the vibrations of particles of the string at any two neighbouring antinodes.
- c Calculate the ratio of the frequency of the mode of vibration shown in Fig. 15.39 to the frequency of the fundamental mode of vibration of the string.
- d The frequency of the mode of vibration shown in Fig. 15.39 is 160 Hz. Calculate the speed of the progressive waves which produced this stationary wave.

11 A vibrating tuning fork of frequency 320 Hz is held over the open end of a resonance tube. The other end of the tube is immersed in water. The length of the air column is gradually increased until resonance first occurs. Taking the speed of sound in air as 340 m s^{-1}, calculate the length of the air column. (Neglect any end-correction.)

12 We can hear sounds round corners. We cannot see round corners. Both sound and light are waves. Explain why sound and light seem to behave differently.

13 Fig. 15.40 shows a narrow beam of monochromatic laser light incident normally on a diffraction grating. The central bright spot is formed at O.

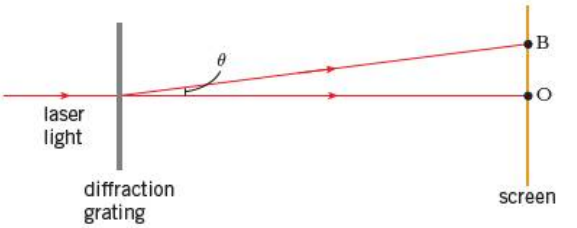


Fig. 15.40

www.iranino.co

IRANINOVIN

a Write down the relationship between the wavelength λ of the light and the angle θ for the first diffraction image formed at B. Identify any other symbol used.

b The screen is 1.1 m from the diffraction grating and the grating has 300 lines per mm. The laser light has wavelength 6.3×10^{-7} m. Find the distance OB from the central spot to the first bright image at B.

c The diffraction grating is now replaced by one which has 600 lines per mm. For this second grating, calculate the distance from the central spot to the first bright image.

14 a State two features of a stationary wave that distinguish it from a progressive wave. *[2]*

b A long tube is open at one end. It is closed at the other end by means of a piston that can be moved along the tube, as shown in Fig. 15.41.

A loudspeaker producing sound of frequency 550 Hz is held near the open end of the tube.

The piston is moved along the tube and a loud sound is heard when the distance L between the piston and the open end of the tube is 45 cm.

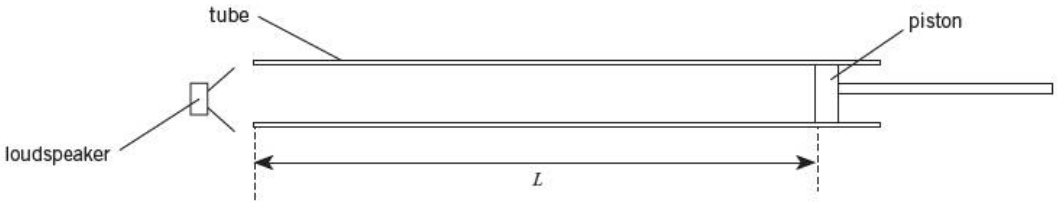


Fig. 15.41

The speed of sound in the tube is 330 m s^{-1}.

i Show that the wavelength of the sound in the tube is 60 cm. *[1]*

ii On a copy of Fig. 15.41, mark all the positions along the tube of

1 the displacement nodes (label these with the letter N), *[1]*

2 the displacement antinodes (label these with the letter A). *[2]*

c The frequency of the sound produced by the loudspeaker in **b** is gradually reduced.
Determine the lowest frequency at which a loud sound will be produced in the tube of length L = 45 cm. *[3]*

Cambridge International AS and A Level Physics, 9702/22 May/June 2010 Q 4

15 a State what is meant by the *diffraction* of a wave. *[2]*

b Plane wavefronts are incident on a slit, as shown in Fig. 15.42.

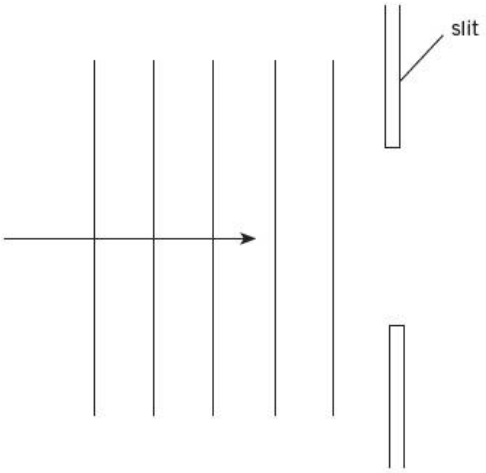


Fig. 15.42

Complete a copy of Fig. 15.42 to show four wavefronts that have emerged from the slit. *[2]*

c Monochromatic light is incident normally on a diffraction grating having 650 lines per millimetre, as shown in Fig. 15.43.
An image (the zero order) is observed for light that has an angle of diffraction equal to zero.

For incident light of wavelength 590 nm, determine the number of orders of diffracted light that can be observed on each side of the zero order. *[3]*

d The images in Fig. 15.43 are viewed, starting with the zero order and then with increasing order number.
State how the appearance of the images changes as the order number increases. *[1]*

Cambridge International AS and A Level Physics, 9702/21 Oct/Nov 2010 Q 5

www.iranino.co

IRANINOVIN

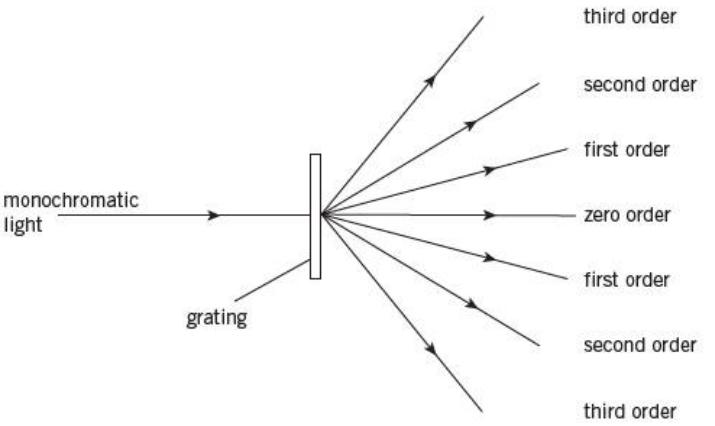


Fig. 15.43

16 a Fig. 15.44 shows the variation with time t of the displacement y of a wave W as it passes a point P. The wave has intensity I.

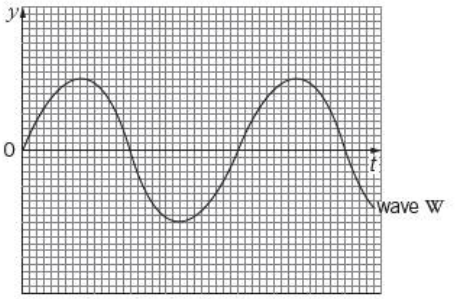


Fig. 15.44

A second wave X of the same frequency as wave W also passes point P. This wave has intensity $\frac{1}{2}I$. The phase difference between the two waves is 60°. On a copy of Fig. 15.44, sketch the variation with time t of the displacement y of wave X. *[3]*

b In a double-slit interference experiment using light of wavelength 540 nm, the separation of the slits is 0.700 mm. The fringes are viewed on a screen at a distance of 2.75 m from the double slit, as illustrated in Fig. 15.45 (not to scale).

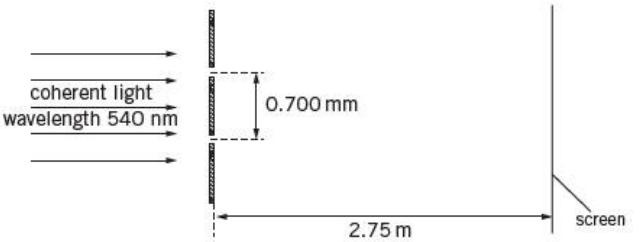


Fig. 15.45

Calculate the separation of the fringes observed on the screen. *[3]*

c State the effect, if any, on the appearance of the fringes observed on the screen when the following changes are made, separately, to the double-slit arrangement in **b**.

i The width of each slit is increased but the separation remains constant. *[3]*

ii The separation of the slits is increased. *[2]*

Cambridge International AS and A level Physics, 9702/02 Oct/Nov 2007 Q 5

17 a Explain the term *interference*. *[1]*

b A ripple tank is used to demonstrate interference between water waves.
Describe

i the apparatus used to produce two sources of coherent waves that have circular wavefronts, *[2]*

ii how the pattern of interfering waves may be observed. *[2]*

www.iranino.co

IRANINOVIN

c A wave pattern produced in **b** is shown in Fig. 15.46. Solid lines on Fig. 15.46 represent crests.

On a copy of Fig. 15.46,

i draw two lines to show where maxima would be seen (label each of these lines with the letter X), *[1]*

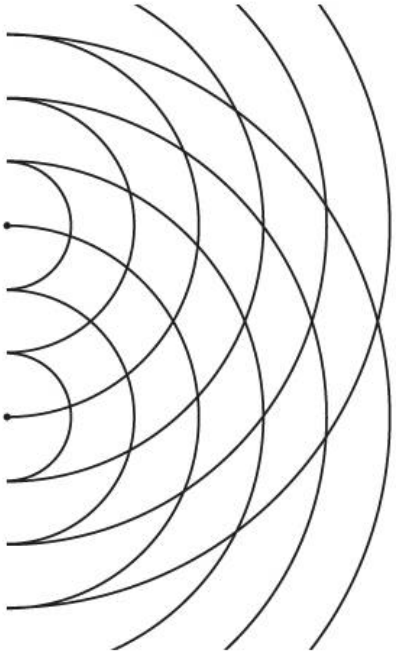


Fig. 15.46

ii draw one line to show where a minimum would be seen (label this line with the letter N). *[1]*

Cambridge International AS and A Level Physics, 9702/21 May/June 2011 Q 7

18 Fig.15.47 shows a string stretched between two fixed points P and Q.

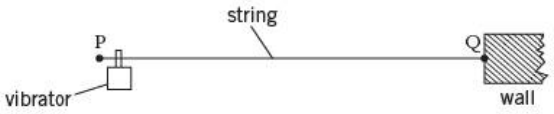


Fig. 15.47

A vibrator is attached near end P of the string. End Q is fixed to a wall. The vibrator has a frequency of 50 Hz and causes a transverse wave to travel along the string at a speed of 40 m s^{-1}.

a i Calculate the wavelength of the transverse wave on the string. *[2]*

ii Explain how this arrangement may produce a stationary wave on the string. *[2]*

b The stationary wave produced on PQ at one instant of time t is shown in Fig. 15.48. Each point on the string is at its maximum displacement.

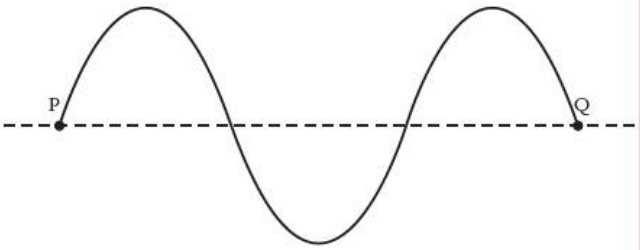


Fig. 15.48

i On a copy of Fig. 15.48, label all the nodes with the letter N and all the antinodes with the letter A. *[2]*

ii Use your answer in **a i** to calculate the length of string PQ. *[1]*

iii On a copy of Fig. 15.48, draw the stationary wave at time (t + 5.0 ms). Explain your answer. *[3]*

Cambridge International AS and A level Physics, 9702/22 May/June 2013 Q 5

www.iranino.co

IRANINOVIN

17 Electric fields

By the end of this topic, you will be able to:

17.1 **(a)** understand the concept of an electric field as an example of a field of force and define electric field strength as force per unit positive charge acting on a stationary point charge
(b) represent an electric field by means of field lines

17.2 **(a)** recall and use $E = \Delta V/\Delta d$ to calculate the field strength of the uniform field between charged parallel plates in terms of potential difference and separation
(b) calculate the forces on charges in uniform electric fields
(c) describe the effect of a uniform electric field on the motion of charged particles

Starting points

- There are two types of charge, positive and negative.
- Objects can be charged by friction or induction.
- Electric forces hold electrons in atoms, and bind atoms together in molecules and in solids.

17.1 Electric forces and fields

Some effects of static electricity are familiar in everyday life. For example, a balloon rubbed on a woollen jumper will stick to a wall, a TV screen that has been polished attracts dust, dry hair crackles (and may actually spark!) when brushed, and you may feel a shock when you touch the metal door-handle of a car on getting out after a journey in dry weather. All these are examples of insulated objects that have gained an electric charge by friction – that is, by being rubbed against other objects.

Insulators that are charged by friction will attract other objects. Some of the effects have been known for centuries. Greek scientists experimented with amber that was charged by rubbing it with fur. Our words *electron* and *electricity* come from the Greek word *ηλεκτρον*, which means amber. Today, electrostatics experiments are often carried out with plastic materials which are moisture-repellent and stay charged for longer.

Charging by friction can be hazardous. For example, a lorry which carries a bulk powder must be earthed before emptying its load. Otherwise, electric charge can build up on the tanker. This could lead to a spark from the tanker to earth, causing an explosion. Similarly, the pipes used for movement of highly flammable liquids (for example, petrol), are metal-clad. An aircraft moving through the air will also become charged. To prevent the first person touching the aircraft on landing from becoming seriously injured, the tyres are made to conduct, so that the aircraft loses its charge on touchdown.

There are two kinds of electric charge. Polythene becomes negatively charged when rubbed with wool, and cellulose acetate becomes positively charged, also when rubbed with wool. To understand why this happens, we need to go back to the model of the atom. An atom consists of a positively charged nucleus with negatively charged electrons orbiting it. When the polythene is rubbed with wool, friction causes some

www.iranino.co

IRANINOVIN

electrons to be transferred from the wool to the polythene. So the polythene has a negative charge, and the wool is left with a positive charge. Cellulose acetate becomes positive because it loses some electrons to the wool when it is rubbed. Polythene and cellulose acetate are poorly conducting plastics, and the charges remain static on the surface.

Putting two charged polythene rods close to each other, or two charged acetate rods close to each other, shows that similar charges repel each other (Figure 17.1). Conversely, unlike charges attract. A charged polythene rod attracts a charged acetate rod.

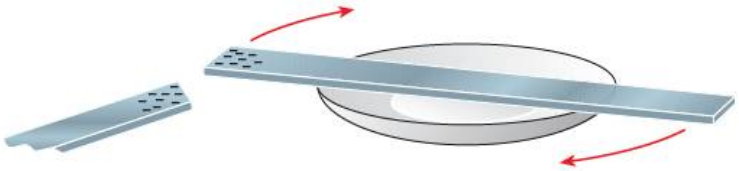


Figure 17.1 Like charges repel.

This is the basic law of force between electric charges:

Like charges repel, unlike charges attract.

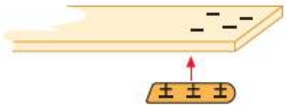


Figure 17.2 A charged rod can induce charges in an uncharged object.

Charged rods will also attract uncharged objects. A charged polythene rod will pick up small pieces of paper, for example. The presence of the negative charge on the rod causes a redistribution of the charges in the paper. Free electrons are repelled to the far side of the paper, so that the side nearest the rod is positive and is, therefore, attracted to the rod (Figure 17.2). The paper is said to be **charged by induction**. When the rod is removed, the free electrons will move back and cancel the positive charge.

Electric fields

Electric charges exert forces on each other, even when they are separated by a distance. The idea of an **electric field** is used to explain this action at a distance. An electric field is a region of space where a stationary charge experiences a force.

Electric fields are invisible, but they can be represented by electric lines of force, just as magnetic fields can be represented by magnetic lines of force (see Topic 22). The direction of the electric field is defined as the direction in which a positive charge would move if it were free to do so. So the lines of force are drawn with arrows that go from positive charge to negative charge.

For any electric field:
- the lines of force start on a positive charge, and end on a negative charge
- the lines of force are smooth curves which never touch or cross
- the strength of the electric field is indicated by the closeness of the lines: the closer they are, the stronger the field.

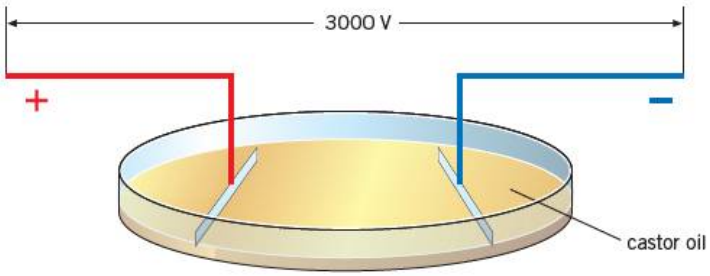


Figure 17.3 Apparatus for investigating electric field patterns

The apparatus in Figure 17.3 can be used to show electric field patterns. Semolina is sprinkled on a non-conducting oil and a high voltage supply is connected to the conducting plates. The semolina becomes charged by induction and lines up along

www.iranino.co

IRANINOVIN

the lines of force. Some electric field patterns are illustrated in Figure 17.4. The pattern for the charged conducting sphere (Figure 17.4c) is of particular importance. Although there are no electric field lines inside the conductor, the field lines appear to come from the centre. Thus, the charge on the surface of a conducting sphere appears as if it is all concentrated at the centre. This means that small conducting spheres may be used as an approximation to point charges.

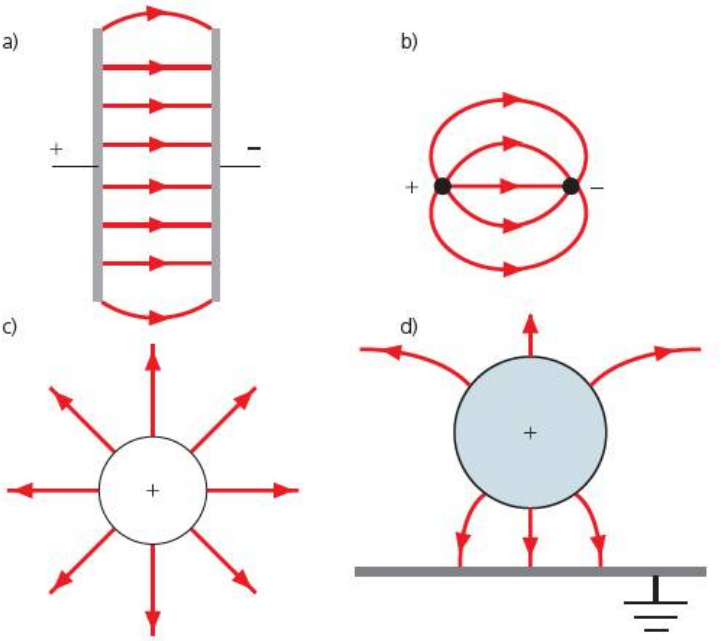


Figure 17.4 Some electric field patterns

Paint spraying makes use of some of the principles of electrostatics. It can be wasteful if the paint is not sprayed where it is needed. But if the spray is given a charge, and the object to be painted is given an opposite charge, the charged droplets follow the lines of force and end up on the surface. This is illustrated in Figure 17.5.

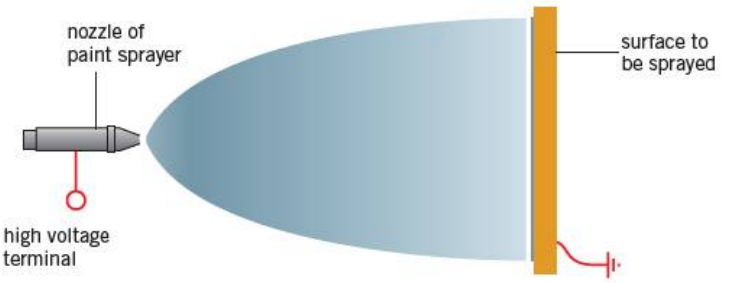


Figure 17.5 Paint spraying

Induced fields and charges

There is no resultant electric field in a conductor (unless a source of e.m.f. is connected across it). The reason for this is that electrons are free to move in the conductor. As soon as a charged body is placed near the conductor, the electric field of the charged body causes electrons in the conductor to move in the opposite direction to the field (because electrons have a negative charge). This is illustrated in Figure 17.6. These charges create a field in the opposite direction to the field due to the charged body. The induced charges will stop moving when the two fields are equal and opposite. Hence there is no resultant field in a conductor.

www.iranino.co

IRANINOVIN

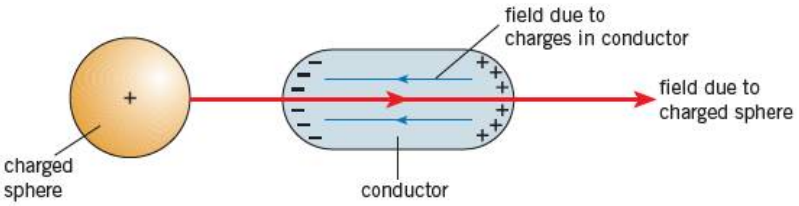


Figure 17.6 Induced charges

This effect may be used to charge a body by induction. The process is illustrated in Figure 17.7. If a positively charged rod is placed near an insulated conductor, induced charges appear on the conductor, as shown in Figure 17.7a. If the conductor is now earthed, as in Figure 17.7b, electrons move from earth to neutralise the induced positive charge. The electrons are held in position by the positively charged rod. When the earth connection is removed, the negative charge is still held in position by the positively charged rod, as in Figure 17.7c. Removal of the charged rod means that the electrons will distribute themselves over the surface of the sphere, as in Figure 17.7d. Note that if a positively charged rod is used, the final charge on the sphere is negative; if a negatively charged rod is used, the sphere becomes positively charged.

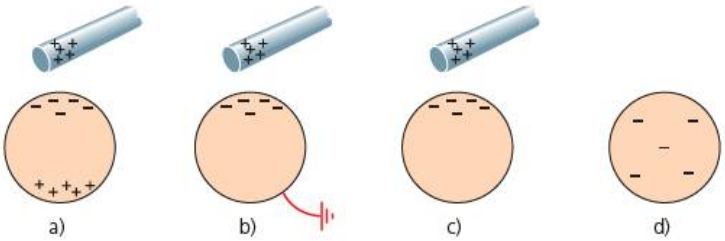


Figure 17.7 Charging by induction

17.2 Electric field strength

The **electric field strength** at a point is defined as the force per unit charge acting on a small stationary positive charge placed at that point.

If the force experienced by a positive test charge $+Q$ placed in the field is F, the field strength E is given by

$$E = \frac{F}{Q}$$

(Be careful not to be confused by the symbol E for field strength. E may also be used for energy.) A unit of field strength can be deduced from this equation: force is measured in newtons and charge in coulombs, so the SI unit of field strength is N C^{-1}. We shall see later that there is another common SI unit for electric field strength, V m^{-1}. These two units are equivalent.

A uniform electric field

A uniform field is where the electric field strength is the same at all points in the field.

In Figure 17.4a, the electric field pattern between the parallel plates consists of parallel, equally spaced lines, except near the edges of the plates. This shows that the field between charged parallel plates (as, for example, in a capacitor, see Topic 18) is uniform. This means that the force experienced by a charge is the same no matter where the charge is placed in the uniform field.

Figure 17.8 illustrates parallel plates a distance d apart with a potential difference V between them. A charge $+Q$ in the uniform field between the plates has a force F

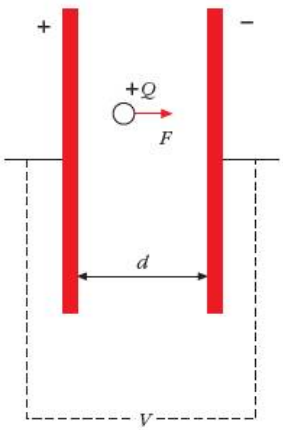


Figure 17.8

www.iranino.co

IRANINOVIN

acting on it. To move the charge towards the positive plate would require work to be done on the charge. Work is given by the product of force and distance. To move the charge from one plate to the other requires work W given by

$$W = Fd$$

From the definition of potential difference (page 148),

$$W = VQ$$

Thus $W = Fd = VQ$, or, rearranging,

$$F/Q = V/d$$

But F/Q is the force per unit charge, and this is the definition of electric field strength. Thus, for a uniform field, the field strength E is given by

$$E = \frac{V}{d}$$

This equation gives an alternative unit for electric field strength (remember that we have already derived N C^{-1} from the idea of force per unit charge). If we use the expression for a uniform field, the unit of potential difference is the volt and the unit of distance is the metre, so another unit for electric field strength is volts per metre (V m^{-1}). The two units N C^{-1} and V m^{-1} are equivalent.

Example

Two metal plates 5.0 cm apart have a potential difference of 1000 V between them. Calculate:

(a) the strength of the electric field between the plates,
(b) the force on a charge of 5.0 nC between the plates.

(a) From $E = V/d$, $E = 1000/5.0 \times 10^{-2} = \mathbf{2.0 \times 10^4\ V\ m^{-1}}$.
(b) From $F = EQ$, $F = 2.0 \times 10^4 \times 5.0 \times 10^{-9} = \mathbf{1.0 \times 10^{-4}\ N}$.

Now it's your turn

1 A glass rod rubbed with silk becomes positively charged. Explain what has happened to the glass. Explain also why the charged rod is able to attract small pieces of paper.

2 (a) Explain what is meant by an electric field.
(b) Sketch the electric field patterns
(i) between two negatively charged particles,
(ii) between a point positive charge and a negatively charged flat metal plate.

3 A positive and a negative charge of the same magnitude are on the same straight line as shown in Figure 17.9. State the direction of the electric field strength
(a) at point A,
(b) at point B,
(c) at point C,
(d) at point D.

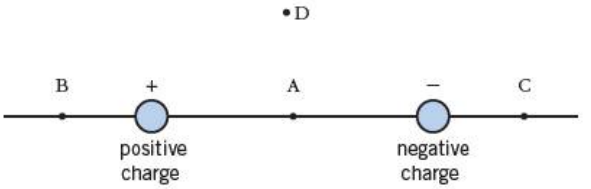


Figure 17.9

4 Two metal plates 15 mm apart have a potential difference of 750 V between them. The force on a small charged sphere placed between the plates is 1.2×10^{-7} N. Calculate:
(a) the strength of the electric field between the plates,
(b) the charge on the sphere.

www.iranino.co

IRANINOVIN

The electron charge

A uniform electric field was used by Millikan in 1912 to measure the electron charge. It involves the balancing of the gravitational force on a tiny, charged sphere by an electric force in the opposite direction, so that the sphere remains stationary. In Millikan's original experiment, the sphere was an oil droplet, which became charged by friction as it passed out of the jet of a spray. If you do this experiment in a school laboratory now, you will probably use a very small polystyrene sphere.

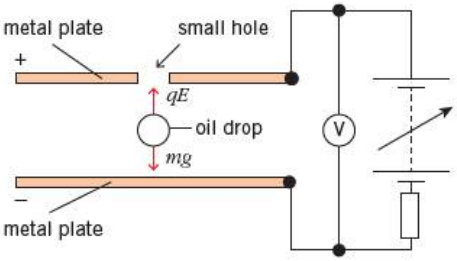


Figure 17.10 Principle of Millikan's experiment for e

Millikan's arrangement is sketched in Figure 17.10. Two horizontal and parallel metal plates are connected to a source of potential difference. This arrangement produces a uniform electric field between the plates. The electric field can be altered by varying the potential difference across the plates. The charged droplets are sprayed through a small hole in the top plate. The tendency is for a droplet to fall under its own weight, mg. The observer concentrates on one droplet and adjusts the electric field until the weight is exactly balanced by the electric force qE in the opposite direction. This gives the equilibrium condition $qE = qV/d = mg$. Millikan found that the charge q was always an integral multiple of a fundamental quantity of charge e equal to 1.6×10^{-19}C. This fundamental charge was ascribed to the charge of the electron. The experimental result that the charges on the droplets seems to be only in integral multiples of e means that charge is quantised, or exists only in discrete amounts.

Example

A negatively charged polystyrene sphere of mass 3.3×10^{-15}kg is held at rest between two parallel plates separated by 5.0mm when the potential difference between them is 170V. How many excess electrons are on the sphere?

Apply the equation for the equilibrium condition, $qV/d = mg$. This gives $q = 3.3 \times 10^{-15} \times 9.8 \times 5.0 \times 10^{-3}/170 = 9.5 \times 10^{-19}$C. Taking the electron charge e as 1.6×10^{-19}C, q corresponds to 5.9 excess electrons. Because you can't have 0.9 of an electron, the answer must be **6 excess electrons**. The discrepancy must be due to experimental inaccuracy or to rounding errors; we have been working to only two significant figures.

Now it's your turn

5 The charged sphere in the example above suddenly loses one of its excess electrons. The potential difference between the plates remains the same.
   (a) Describe the initial motion of the sphere.
   (b) Calculate the initial acceleration.

Motion of a charged particle in an electric field

Motion of charged particle moving parallel to the electric field

A charged particle at rest or moving along the direction of an electric field experiences an electric force. The direction of the force depends on the sign of the charge. A particle of charge $+q$ is accelerated in the direction of the electric field by the force. For a uniform field the force and, hence, the acceleration are constant. Hence, the equations of uniform acceleration given in Topic 3 can be used to determine the motion of the charged particle.

www.iranino.co

IRANINOVIN

A uniform electric field E is produced by a potential difference V across two horizontal and parallel metal plates separated by a distance d as shown in Figure 17.11. A charge of $+q$ is introduced at the top plate and moves to the bottom plate. The field does work on the charge and the charge gains kinetic energy.

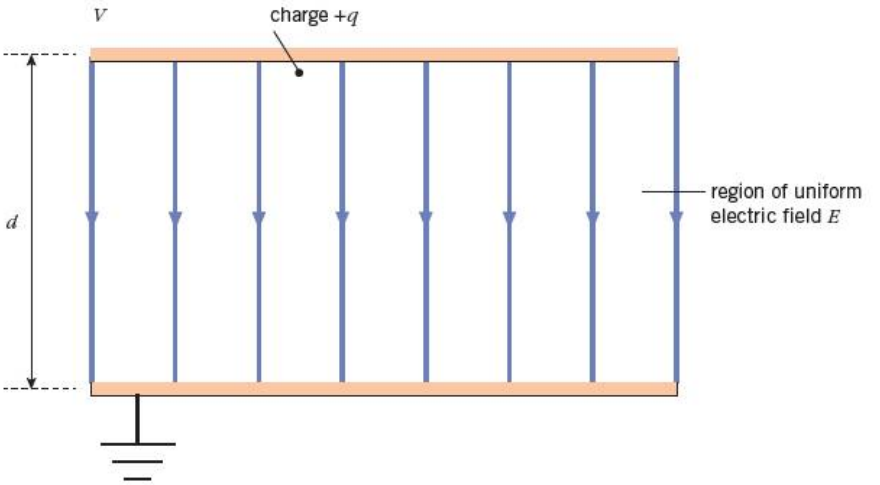


Figure 17.11

The work done = $Fd = Eqd = Vq$

Hence the gain in kinetic energy = Vq

The motion of the charged particle can also be determined using the link between work done and change in energy.

A negatively charged particle is accelerated in the opposite direction to the electric field.

Motion of charged particle with initial velocity perpendicular to the electric field

A charged particle $+q$ enters an electric field with a velocity v at right angles to the direction of the field (see Figure 17.12). The particle experiences a force at right angles to its initial direction. If the field direction is vertical and the initial direction is horizontal, the particle will follow a parabolic path as it passes through the field.

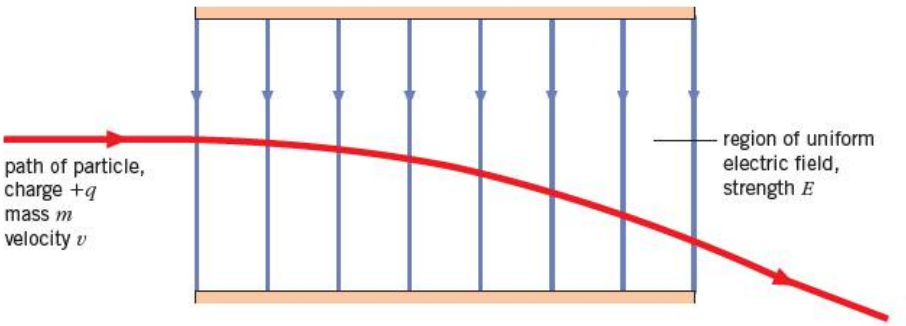


Figure 17.12

The analysis of the motion is similar to that of a particle of mass m moving in a gravitational field with a uniform velocity in one direction and a uniform acceleration in a perpendicular direction (as described for projectile motion in Topic 3).

www.iranino.co

IRANINOVIN

Example

A particle of mass 6.7×10^{-27} kg and charge $+3.2 \times 10^{-19}$ C is placed near the top plate in an electric field similar to that shown in Figure 17.11. The potential difference across the plates is 600 V and the separation of the plates is 12 mm. Determine

(a) the force on the particle,
(b) the acceleration of the particle,
(c) the speed of the particle when it reaches the bottom plate.

(a) force = $Eq = Vq/d = (600 \times 3.2 \times 10^{-19})/12 \times 10^{-3} =$ **1.6×10^{-14} N**
(b) acceleration = force/mass = $2.3 \times 10^{-12}/6.7 \times 10^{-27} =$ **2.4×10^{12} m s^{-2}**
(c) $v^2 = 0 + 2as = 2 \times 2.4 \times 10^{12} \times 12 \times 10^{-3}$ so speed $v =$ **2.4×10^{5} m s^{-1}**

Now it's your turn

6 An electron enters an electric field similar to that shown in Figure 17.12 with a horizontal velocity of 7.0×10^{7} m s^{-1}. The horizontal length of the plates is 2.0 cm. The electric field strength is 3.2×10^{5} V m^{-1}. Calculate the vertical displacement of the electron for the time it is between the plates.

Summary

- Insulators may be charged by friction.
- Like charges repel; unlike charges attract each other.
- When charged objects are placed near conductors, they cause a redistribution of charge in the conductor, thereby inducing charges.
- An electric field is a region of space where a stationary charge experiences a force.
- Electric field strength is the force per unit positive charge: $E = F/Q$
- The electric field between parallel plates is uniform. The field strength is given by: $E = V/d$
- The motion of a charged particle moving in the direction of the electric field can be analysed using the equations of constant acceleration.
- The motion of a charged particle moving with a uniform velocity in one direction and a uniform acceleration in a perpendicular direction is parabolic.

Examination style questions

1 Calculate the speed of an electron accelerated from rest through a distance of 40 mm by a uniform electric field of 3.0×10^{3} N C^{-1}.

2 An electron starts from rest from the bottom plate in Fig. 17.11. The potential difference across the plates is 1600 V and the separation of the plates is 15 mm. Calculate the time taken for the electron to reach the top plate.

3 For the electron in question **2** calculate:
- a the work done by the field on the electron,
- b the gain in kinetic energy,
- c the speed of the electron.

4 a Two horizontal metal plates are connected to a power supply, as shown in Fig. 17.13.

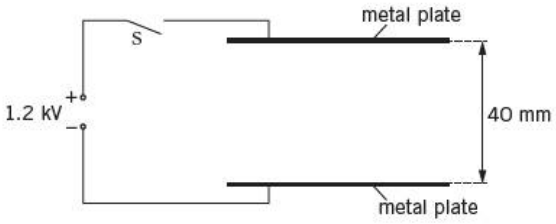


Fig. 17.13

The separation of the plates is 40 mm. The switch S is then closed so that a potential difference of 1.2 kV is applied across the plates.

i On a copy of Fig. 17.13, draw six field lines to represent the electric field between the metal plates. *[2]*

ii Calculate the electric field strength E between the plates. *[2]*

b The switch S is opened and the plates lose their charge. Two very small metal spheres A and B joined by an insulating rod are placed between the metal plates as shown in Fig. 17.14.

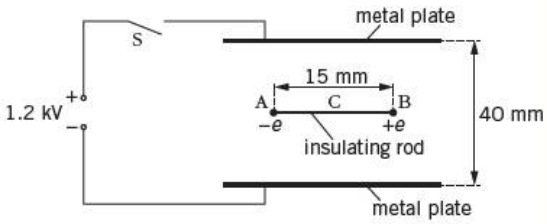


Fig. 17.14

www.iranino.co

IRANINOVIN

Sphere A has charge $-e$ and sphere B has charge $+e$, where e is the charge of a proton. The length AB is 15 mm. The rod is supported at its centre C so that the rod is horizontal and in equilibrium.
The switch S is then closed so that the potential difference of 1.2 kV is applied across the plates.

i There is a force acting on A due to the electric field between the plates. Show that this force is 4.8×10^{-15} N. [2]

ii The insulating rod joining A and B is fixed in the position shown in Fig. 17.14. Calculate the torque of the couple acting on the rod. [3]

iii The insulating rod is now released so that it is free to rotate about C. State and explain the position of the rod when it comes to rest. [2]

Cambridge International AS and A level Physics, 9702/23 Oct/Nov 2013 Q 7

5 Two vertical parallel metal plates are situated 2.50 cm apart in a vacuum. The potential difference between the plates is 350 V, as shown in Fig. 17.15.

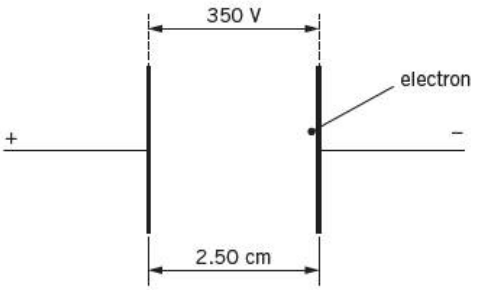


Fig. 17.15

An electron is initially at rest close to the negative plate and in the uniform electric field between the plates.

a i Calculate the magnitude of the electric field between the plates. [2]

ii Show that the force on the electron due to the electric field is 2.24×10^{-15} N. [2]

b The electron accelerates horizontally across the space between the plates. Determine

i the horizontal acceleration of the electron, [2]

ii the time to travel the horizontal distance of 2.50 cm between the plates. [2]

c Explain why gravitational effects on the electron need not be taken into consideration in your calculation in **b**. [2]

Cambridge International AS and A level Physics, 9702/21 May/June 2009 Q 6

6 **a** Define *electric field strength*. [1]

b Two flat parallel metal plates, each of length 12.0 cm, are separated by a distance of 1.5 cm, as shown in Fig. 17.16.

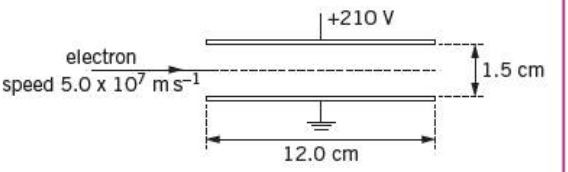


Fig. 17.16

The space between the plates is a vacuum. The potential difference between the plates is 210 V. The electric field may be assumed to be uniform in the region between the plates and zero outside this region. Calculate the magnitude of the electric field strength between the plates. [1]

c An electron initially travels parallel to the plates along a line mid-way between the plates, as shown in Fig. 17.16. The speed of the electron is 5.0×10^7 m s^{-1}. For the electron between the plates,

i determine the magnitude and direction of its acceleration, [4]

ii calculate the time for the electron to travel a horizontal distance equal to the length of the plates. [1]

d Use your answers in **c** to determine whether the electron will hit one of the plates or emerge from between the plates. [3]

Cambridge International AS and A level Physics, 9702/02 May/June 2007 Q 2

www.iranino.co

IRANINOVIN

19 Current of electricity

By the end of this topic, you will be able to:

19.1 (a) understand that electric current is a flow of charged carriers
(b) understand that the charge on charge carriers is quantised
(c) define the coulomb
(d) recall and use $Q = It$
(e) derive and use, for a current-carrying conductor, the expression $I = Anvq$, where n is the number density of charge carriers

19.2 (a) define potential difference and the volt
(b) recall and use $V = W/Q$
(c) recall and use $P = VI$, $P = I^2R$

19.3 (a) define resistance and the ohm
(b) recall and use $V = IR$
(c) sketch and discuss the I–V characteristics of a metallic conductor at constant temperature, a semiconductor diode and a filament lamp
(d) state Ohm's law
(e) recall and use $R = \rho l/A$

Starting points

- This topic considers fundamental ideas about electric charge and electric current.
- Examples of electric currents are in household wiring and electrical appliances.
- A potential difference is required for energy changes to occur in a circuit.
- Resistance controls the flow of charge in a circuit.

19.1 Charge and current

All matter is made up of tiny particles called atoms, each consisting of a positively charged nucleus with negatively charged electrons moving around it.

Charge is measured in units called **coulombs** (symbol C). The charge on an electron is -1.6×10^{-19} C. Normally atoms have equal numbers of positive and negative charges, so that their overall charge is zero. But for some atoms it is relatively easy to remove an electron, leaving an atom with an unbalanced number of positive charges. This is called a **positive ion**.

Robert Millikan in 1912 performed an experiment to determine the charge on an electron using charged oil droplets (see Topic 17). The experimental result showed that, no matter what the charge on the droplets, it seemed to occur only in integral multiples of e, the charge on an electron. The conclusion was that charge is not continuous but quantised, or exists only in discrete amounts. The photon is another example of a quantised physical quantity, introduced in Topic 25.

Atoms in metals have one or more electrons which are not held tightly to the nucleus. These **free** (or mobile) **electrons** wander at random throughout the metal. However, when a battery is connected across the ends of the metal, the free electrons drift towards the positive terminal of the battery, producing an **electric current**.

The size of the electric current is given by the rate of flow of charge. Electric current is an SI base quantity (see Topic 1). The SI base unit of current is the ampere (or amp for short), with symbol A. The SI units of all the other electrical quantities are derived from the SI base units.

www.iranino.co

IRANINOVIN

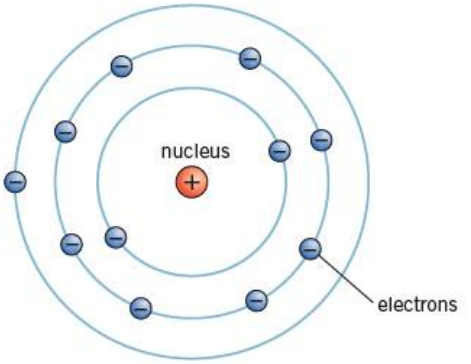


Figure 19.1 Atoms consist of a positively charged nucleus with negative electrons outside.

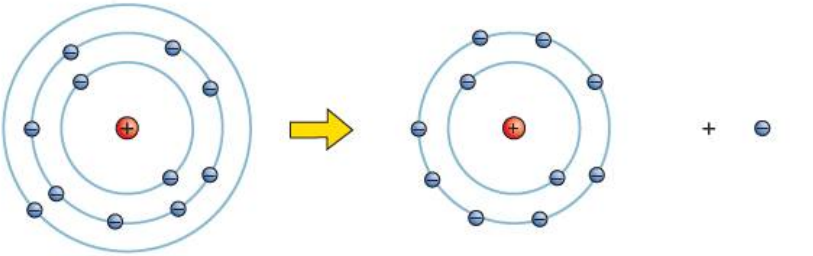


Figure 19.2 An atom with one or more electrons missing is a positive ion.

A current of 3 amperes means that 3 coulombs pass a point in the circuit every second. In 5 seconds, a total charge of 15 coulombs will have passed the point. So,

$$charge = current \times time$$

or

$$Q = It$$

where the charge Q is in coulombs when the current I is in amperes and the time t is in seconds. This gives a definition of the coulomb as follows.

The **coulomb** is that charge passing a point in a circuit when there is a current of one ampere for one second.

Example

The current in the filament of a torch bulb is 0.03 A. How much charge flows through the bulb in 1 minute?

Using $Q = It$, $Q = 0.03 \times 60$ (remember the time must be in seconds), so $Q =$ **1.8 C**.

Now it's your turn

1 Calculate the current when a charge of 240 C passes a point in a circuit in a time of 2 minutes.

2 In a silver-plating experiment, 9.65×10^4 C of charge is needed to deposit a certain mass of silver. Calculate the time taken to deposit this mass of silver when the current is 0.20 A.

3 The current in a wire is 200 mA. Calculate:
(a) the charge which passes a point in the wire in 5 minutes,
(b) the number of electrons needed to carry this charge (electron charge $e = -1.6 \times 10^{-19}$ C).

www.iranino.co

IRANINOVIN

Conventional current

Early studies of the effects of electricity led scientists to believe that an electric current is the flow of 'something'. In order to develop a further understanding of electricity, they needed to know the direction of flow. It was decided that this flow in the circuit should be from the positive terminal of the battery to the negative. This current is called the **conventional current**, and is in the direction of flow of positive charge. We now know, in a metal, that the electric current is the flow of electrons in the opposite direction, from the negative terminal to the positive terminal. However, laws of electricity had become so firmly fixed in the minds of people that the idea of conventional current has persisted. But be warned! Occasionally we need to take into account the fact that electron flow is in the *opposite* direction to conventional current.

Conduction in a current-carrying conductor

Figure 19.3 shows part of a conductor of cross-sectional area A through which there is a current I. In Figure 19.3a the charge carriers are positive and in Figure 19.3b the charge carriers are negative. The current in each of the conductors is from right to left but the charge carriers move in opposite directions shown by the drift speed v.

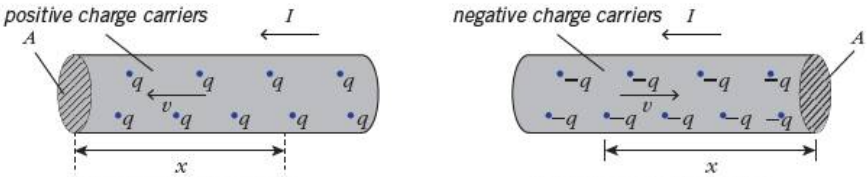


Figure 19.3 Conduction in a current-carrying conductor

The current is due to the movement of the charge carriers along the conductor. The flow of charge carriers is characterised by an average drift speed. The average drift speed of the charge carriers in the conductor is v and the number density (the number per unit volume) of the charge carriers is n. The charge on each of the charge carriers is q. (See Figure 19.3a)

The number of charge carriers in a length x of the conductor is nAx. The amount of charge that leaves this volume through the left-hand side of the conductor in a time t is $nAxq$.

Where the time interval t is x/v

The current I = charge/time = $\dfrac{nAxq}{x/v}$

$$I = nAvq$$

The same expression is obtained if the negative charge carriers are considered in the conductor shown in Figure 19.3b.

Example

A copper wire has 8.5×10^{28} charge carriers (free electrons) m^{-3}. The wire has a current of 2.0 A and a cross-sectional area of 1.2 mm^2. Calculate the average drift speed of the electrons.

$$v = I/nAe = 2/(8.5 \times 10^{28} \times 1.2 \times 10^{-6} \times 1.6 \times 10^{-19})$$
$$= \mathbf{1.2 \times 10^{-4}\ m\ s^{-1}}$$

Now it's your turn

4 The average drift speed in a metal wire is 6.5×10^{-4} m s^{-1} when the current is 0.80 A. The diameter of the wire is 0.50 mm. Calculate the number of 'free' electrons per unit volume in the wire.

www.iranino.co

IRANINOVIN

19.2 Potential difference

A cell makes one end of the circuit positive and the other negative. The cell is said to set up a **potential difference** across the circuit. Potential difference (p.d. for short) is measured in volts (symbol V), and is often called the voltage. You should never talk about the potential difference or voltage *through* a device, because it is in fact a difference *across* the ends of the device. The potential difference provides the energy to move charge through the device.

The potential difference between any two points in a circuit is a measure of the electrical energy transferred, or the work done, by each coulomb of charge as it moves from one point to the other. We already know that the unit of potential difference is the volt (V). Energy W is measured in joules, and charge Q in coulombs.

$$\text{potential difference} = \frac{\text{energy transferred [from electrical to other forms] (or work done)}}{\text{charge}}$$

or

$$V = \frac{W}{Q}$$

We can turn this relation round to get an expression for the electrical energy transferred or converted when a charge Q is moved through a potential difference V:

$$\text{energy transferred (work done)} = \text{potential difference} \times \text{charge}$$

$$W = VQ$$

The relation gives a definition of the volt as follows.

One **volt** is the potential difference between two points such that the energy transferred from electrical to other forms is one joule per coulomb of electric charge passing from one point to the other.

In Figure 19.4, one lamp is connected to a 240 V mains supply and the other to a 12 V car battery. Both lamps have the same current, yet the 240 V lamp glows more brightly. This is because the energy supplied to each coulomb of charge in the 240 V lamp is 20 times greater than for the 12 V lamp.

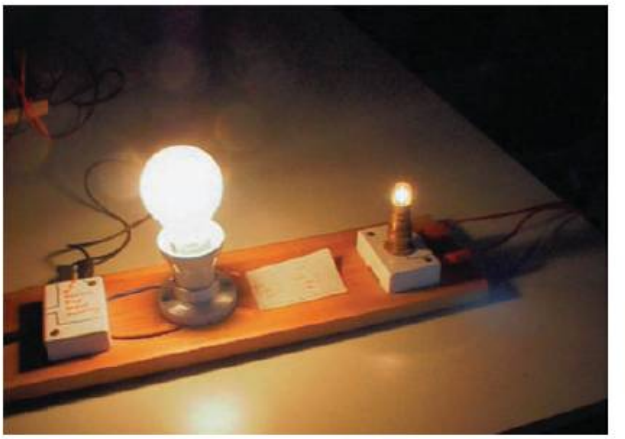


Figure 19.4 A 240 V, 100 W mains lamp is much brighter than a 12 V, 5 W car light, but both have the same current. (Do not try this experiment yourself as it involves a large voltage.)

Example

Electrons in a particular television tube are accelerated by a potential difference of 20 kV between the filament and the screen. The charge of the electron is -1.6×10^{-19} C. Calculate the gain in kinetic energy of each electron.

www.iranino.co

IRANINOVIN

Since $V = W/Q$, then $W = VQ$. The electrical energy transferred to the electron shows itself as the kinetic energy of the electron. Thus,

$$\text{kinetic energy} = VQ = 20 \times 10^3 \times 1.6 \times 10^{-19}$$
$$= \mathbf{3.2 \times 10^{-15}\,J}.$$

(Don't forget to turn the 20 kV into volts.)

Now it's your turn

5 An electron in a particle accelerator is said to have 1 MeV (1 mega-electronvolt) of energy when it has been accelerated through a potential difference of 1 million volts. Calculate the energy, in joules, gained by the electron.

6 A torch bulb is rated 2.2 V, 0.25 A. Calculate:
(a) the charge passing through the bulb in one second,
(b) the energy transferred by the passage of each coulomb of charge.

Electrical power

In Topic 6 we defined power P as the rate of doing work, or of transferring energy. The definition of potential difference V gives energy transferred per unit charge.

Therefore $P = W/t$ and $V = W/Q$

or $P = VQ/t$

Since $Q/t = I$

$$P = VI$$

power = potential difference × current

The power is measured in watts (W) when the potential difference is in volts (V) and the current is in amperes (A). A voltmeter can measure the p.d. across a device and an ammeter the current through it; the equation above can then be used to calculate the power in the device.

19.3 Resistance

Connecting wires in circuits are often made from copper, because copper offers little opposition to the movement of electrons. The copper wire is said to have a low electrical resistance. In other words, copper is a good conductor.

Some materials, such as plastics, are poor conductors. These materials are said to be insulators, because under normal circumstances they conduct little or no current.

The **resistance** R of a conductor is defined as the ratio of the potential difference V across the conductor to the current I in it.

or

$$R = \frac{V}{I}$$

where the resistance is in **ohms** when the potential difference is in volts and the current in amperes. The symbol for ohms is the Greek capital letter omega, Ω.

The relation gives a definition of the ohm as follows.

One **ohm** is one volt per ampere.

We have defined resistance for a conductor, but many devices have resistance. The general term for such a device is a **resistor**. (Note that the resistance of a resistor is measured in ohms, just as the volume of a tank is measured in m^3. We do not refer to the 'm^3' of a tank, or to the 'ohms' of a resistor.)

www.iranino.co

IRANINOVIN

The relation between resistance, potential difference and current means that, for a given potential difference, the resistance controls the size of the current in a circuit. A high resistance means a small current, while a low resistance means a large current.

Example

The current in an electric immersion heater in a school experiment is 6.3 A when the p.d. across it is 12 V. Calculate the resistance of the heater.

Since $R = V/I$, the resistance $R = 12/6.3 =$ **1.9 Ω**.

Now it's your turn

7 The current in a light-emitting diode is 20 mA when it has a potential difference of 2.0 V across it. Calculate its resistance.

Electrical heating

When an electric current passes through a resistor, it gets hot. This heating effect is sometimes called **Joule heating**. The electrical power P produced (dissipated) is given by $P = VI$. We can obtain alternative expressions for power in terms of the resistance R of the resistor. Since $R = V/I$, then

$$P = I^2R$$

and

$$P = V^2/R$$

For a given resistor, the power dissipated depends on the square of the current. Hence, if the current is doubled, the power will be four times as great. Similarly, a doubling of voltage will increase the power by a factor of four.

Example

An electric immersion heater used in a school experiment has a current of 6.3 A when the p.d. across it is 12 V. Calculate the power of the heater.

Since $P = VI$, power = 12 × 6.3 = **76 W**.

Now it's your turn

8 Show that a 100 W lamp connected to a mains supply of 240 V will have the same current as a 5 W car lamp connected to a 12 V battery. (See Figure 19.4 on page 148.)

9 An electric kettle has a power of 2.2 kW at 240 V. Calculate:
(a) the current in the kettle,
(b) the resistance of the kettle element.

Ohm's law

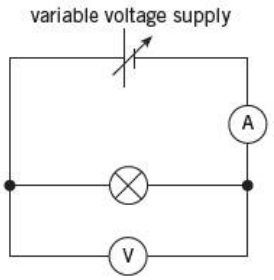


Figure 19.5 Circuit for plotting graphs of current against voltage

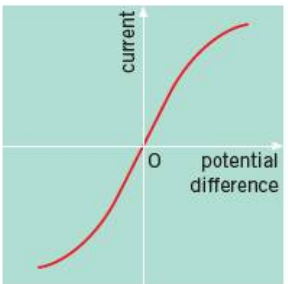


Figure 19.6 Current against voltage for a filament lamp

The relation between the potential difference across an electrical component and the current through it may be investigated using the circuit of Figure 19.5. For example, if a filament lamp is to be investigated, adjust the voltage across the lamp and measure the corresponding currents and voltages. The variation of current with potential difference is shown in Figure 19.6.

The resistance R of the lamp can be calculated from $R = V/I$. At first, the resistance is constant (where the graph is a straight line), but then the resistance increases with current (where the graph curves).

If the lamp is replaced by a length of constantan wire, the graph of the results is as shown in Figure 19.7. It is a straight line through the origin. This shows that, for constantan wire, the current is proportional to the voltage. The resistance of the wire is found to stay the same as the current increases. The difference between Figures 19.6

www.iranino.co

IRANINOVIN

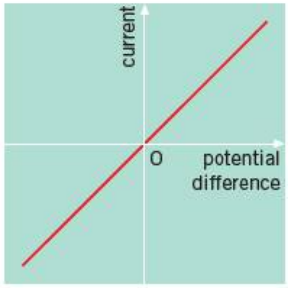


Figure 19.7 Current against voltage for a constantan wire



Figure 19.8 Georg Ohm

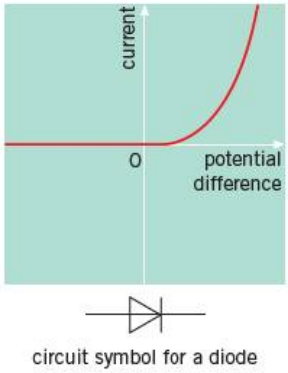


Figure 19.9 Current against voltage for a diode

and 19.7 is that the temperature of the constantan wire was constant for all currents used in the experiment, whereas the temperature of the filament of the lamp increased to about 1500°C.

Graphs like Figure 19.7 would be obtained for wires of any metal, provided that the temperature of the wires did not change during the experiment. The graph illustrates a law discovered by the German scientist Georg Simon Ohm (Figure 19.8). (Ohm's name is now used for the unit of resistance.)

Ohm's law states that, for a conductor at constant temperature, the current in the conductor is proportional to the potential difference across it.

Conductors where the current against voltage graph is a straight line through the origin, like that in Figure 19.7, are said to obey Ohm's law. It is found that Ohm's law applies to metal wires, provided that the current is not too large. What does 'too large' mean here? It means that the current must not be so great that there is a pronounced heating effect, causing an increase in temperature of the wire.

A lamp filament consists of a thin metal wire. Why does it not obey Ohm's law? (Figure 19.6 shows that the current against voltage graph is not a straight line.) This is because, as stated previously, the temperature of the filament does not remain constant. The increase in current causes the temperature to increase so much that the filament glows.

Current–voltage characteristics of a diode

When a diode is tested in the same way as the filament lamp (see the circuit in Figure 19.5), the current–voltage graph shown in Figure 19.9 is obtained.

Diodes are made from semiconducting material. The diode conducts when the current is in the direction of the arrowhead on the symbol. This condition is called **forward bias**. The potential on the left-hand side of the diode is more positive than the potential on the right-hand side. When the voltage is reversed, there is negative bias. This is called **reverse bias**. Figure 19.9 shows this important difference in the current–voltage graph when the voltage is reversed. The diode does not conduct when in the reverse bias condition. As a result, diodes are used to change alternating current into direct current in devices called **rectifiers** (see Topic 24).

Diodes do not obey Ohm's law. The resistance of the diode is very high for low voltages in the forward bias condition. The diode conducts with a forward bias voltage of about 0.5V. For voltages above this value, the resistance of the diode decreases. The straight line part of the graph in this region does not follow Ohm's law as the line does not go through the origin and, therefore, the resistance is not constant. The current is not proportional to the p.d..

Resistance and temperature

All solids are made up of atoms that constantly vibrate about their equilibrium positions. The higher the temperature, the greater the amplitude of vibration.

Electric current is the flow of free electrons through the material. As the electrons move, they collide with the vibrating atoms, so their movement is impeded. The more the atoms vibrate, the greater is the chance of collision. This means that the current is less and the resistance is greater.

A temperature rise can cause an increase in the number of free electrons. If there are more electrons free to move, this may outweigh the effect due to the vibrating atoms, and thus the flow of electrons, or the current, will increase. The resistance is therefore reduced. This is the case in semiconductors. Insulators, too, show a reduction in resistance with temperature rise.

For metals there is no increase in the number of free electrons. The increased amplitude of vibration of the atoms makes the resistance of metals increase with temperature.

www.iranino.co

IRANINOVIN

Resistivity

All materials have some resistance to a flow of charge. A potential difference across the material causes free charges inside to accelerate. As the charges move through the material, they collide with the atoms of the material which get in their way. They transfer some or all of their kinetic energy, and then accelerate again. It is this transfer of energy on collision that causes electrical heating.

As you may have thought, the longer a wire, the greater its resistance. This is because the charges have further to go through the material; there is more chance of collisions with the atoms of the material. In fact the resistance is proportional to the length of the wire, or $R \propto l$.

Also, the thicker a wire, the smaller its resistance. This is because there is a bigger area for the charges to travel through, with less chance of collision. In fact the resistance is inversely proportional to the cross-sectional area of the wire, or $R \propto 1/A$. These relations are illustrated in Figures 19.10 and 19.11.

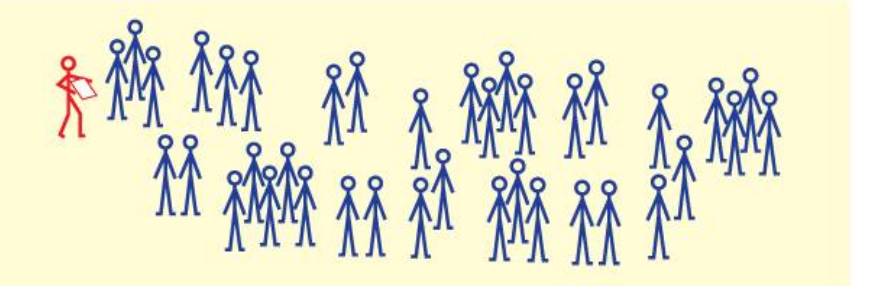


Figure 19.10 The longer the room, the greater the resistance the waiter meets.

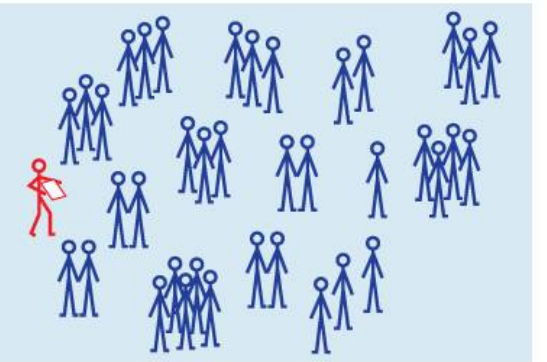


Figure 19.11 The wider the room, the easier it is for the waiter to pass through.

Finally, the resistance depends on the type of material. As previously stated, copper is a good conductor, whereas plastics are good insulators. Putting all of this together gives

$$R = \frac{\rho l}{A}$$

where ρ is a constant for a particular material at a particular temperature. ρ is called the **resistivity** of the material at that temperature.

The resistivity ρ of a material is numerically equal to the resistance between opposite faces of a cube of the material, of unit length and unit cross-sectional area.

So if $\rho = RA/l$ and with the resistance in ohms, the cross-sectional area in metres squared and the length in metres, then the resistivity is in ohm metres (Ω m).

www.iranino.co

IRANINOVIN

Table 19.1 Resistivity of some materials at room temperature

material	resistivity/Ω m
metals:	
copper	1.7×10^{-8}
gold	2.4×10^{-8}
aluminium	2.6×10^{-8}
semiconductors:	
germanium (pure)	0.6
silicon (pure)	2.3×10^{3}
insulators:	
glass	about 10^{12}
perspex	about 10^{13}
polyethylene	about 10^{14}
sulfur	about 10^{15}

Remember that A is the cross-sectional area through which the current is passing, not the surface area.

We have already seen that the resistance of a wire depends on temperature. Thus, resistivity also depends on temperature. The resistivity of a metal increases with increasing temperature, and the resistivity of a semiconductor decreases very rapidly with increasing temperature.

The values of the resistivity of some materials at room temperature are given in Table 19.1.

Note the enormous range of resistivity spanned by the materials in this list – a range of 23 orders of magnitude, from $10^{-8}\,\Omega$ m to $10^{15}\,\Omega$ m.

Note, too, that the resistivity is a property of a material, while the resistance is a property of a particular wire or device.

Example

Calculate the resistance per metre at room temperature of a constantan wire of diameter 1.25 mm. The resistivity of constantan at room temperature is $5.0 \times 10^{-7}\,\Omega$ m.

The cross-sectional area of the wire is calculated using πr^2.

$$\text{Area} = \pi \left(\frac{1.25 \times 10^{-3}}{2} \right)^2 = 1.23 \times 10^{-6}\,\text{m}^2$$

(don't forget to change the units from mm to m.)

The resistance per metre is given by R/l, and $R/l = \rho/A$. So

resistance per metre = $5.0 \times 10^{-7}/(1.23 \times 10^{-6})$ = **0.41 Ω m^{-1}**

Now it's your turn

10 Find the length of copper wire, of diameter 0.63 mm, which has a resistance of 1.00 Ω. The resistivity of copper at room temperature is $1.7 \times 10^{-8}\,\Omega$ m.

11 Find the diameter of a copper wire which has the same resistance as an aluminium wire of equal length and diameter 1.20 mm. The resistivities of copper and aluminium at room temperature are $1.7 \times 10^{-8}\,\Omega$ m and $2.6 \times 10^{-8}\,\Omega$ m respectively.

Summary

- Electric current is the rate of flow of charge: $I = Q/t$
- Conventional current is a flow of positive charge from positive to negative. In metals, current is carried by electrons, which travel from negative to positive.
- Charge on charge carriers is quantised.
- The coulomb is the unit of charge and is an ampere second.
- Charge can be calculated using $Q = It$
- For a current-carrying conductor $I = nAvq$
- Potential difference (or voltage) measures the electrical energy transferred by each coulomb of charge: $V = W/Q$
- Resistance R of a resistor is defined as: $R = V/I$
- Electrical power $P = VI = I^2R = V^2/R$
- Ohm's law: for a conductor at constant temperature, the current in the conductor is proportional to the potential difference across it.
- The resistance of a metallic conductor increases with increasing temperature; the resistance of a semiconductor decreases with increasing temperature.
- A diode has a low resistance when connected in forward bias, and a very high resistance in reverse bias.
- Resistivity ρ of a material is the resistance between opposite faces of a unit cube of the material: $R = \rho l/A$.

www.iranino.co

IRANINOVIN

Examination style questions

1 A 240 V heater takes a current of 4.2 A. Calculate:
   a the charge that passes through the heater in 3 minutes,
   b the rate at which heat energy is produced by the heater,
   c the resistance of the heater.

2 A small torch has a 3.0 V battery connected to a bulb of resistance 15 Ω.
   a Calculate:
      i the current in the bulb,
      ii the power delivered to the bulb.
   b The battery supplies a constant current to the bulb for 2.5 hours. Calculate the total energy delivered to the bulb.

3 The capacity of storage batteries is rated in ampere-hours (A h). An 80 A h battery can supply a current of 80 A for 1 hour, or 40 A for 2 hours, and so on. Calculate the total energy, in J, stored in a 12 V, 80 A h car battery.

4 An electric kettle is rated at 2.2 kW, 240 V. The supply voltage is reduced from 240 V to 230 V. Calculate the new power of the kettle.

5 The element of an electric kettle has resistance 26 Ω at room temperature. The element is made of nichrome wire of diameter 0.60 mm and resistivity 1.1×10^{-6} Ω m at room temperature. Calculate the length of the wire.

6 These are values of the current I through an electrical component for different potential differences V across it:

V/V	0	0.19	0.48	1.47	2.92	4.56	6.56	8.70
I/A	0	0.20	0.40	0.60	0.80	1.00	1.20	1.40

   a Draw a diagram of the circuit that could be used to obtain these values.
   b Calculate the resistance of the component at each value of current.
   c Plot a graph to show the variation with current of the resistance of the component.
   d Suggest what the component is likely to be, giving a reason for your answer.

7 The current in a 2.50 m length of wire of diameter 1.5 mm is 0.65 A when a potential difference of 0.40 V is applied between its ends. Calculate:
   a the resistance of the wire,
   b the resistivity of the material of the wire.

8 A household electric lamp is rated as 240 V, 60 W. The filament of the lamp is made from tungsten and is a wire of constant radius 6.0×10^{-6} m. The resistivity of tungsten at the normal operating temperature of the lamp is 7.9×10^{-7} Ω m.
   a For the lamp at its normal operating temperature,
      i calculate the current in the lamp,
      ii show that the resistance of the filament is 960 Ω. [3]
   b Calculate the length of the filament. [3]
   c Comment on your answer to **b**. [1]

Cambridge International AS and A level Physics, 9702/02 May/June 2004 Q 7

9 a Two resistors, each of resistance R, are connected first in series and then in parallel. Show that the ratio

$$\frac{\text{combined resistance of resistors connected in series}}{\text{combined resistance of resistors connected in parallel}}$$

is equal to 4. [1]

   b The variation with potential difference V of the current I in a lamp is shown in Fig. 19.12.

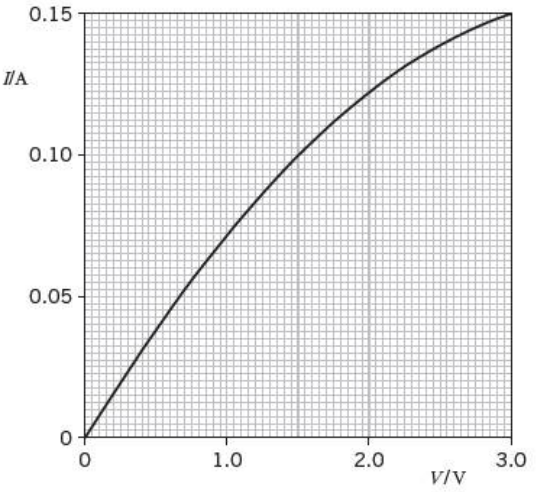


Fig. 19.12

Calculate the resistance of the lamp for a potential difference across the lamp of 1.5 V. [2]

   c Two lamps, each having the I–V characteristic shown in Fig. 19.12, are connected first in series and then in parallel with a battery of e.m.f. 3.0 V and negligible internal resistance.
Complete a copy of the table of Fig. 19.13 for the lamps connected to the battery. [4]

	p.d. across each lamp/V	resistance of each lamp/Ω	combined resistance of lamps/Ω
lamps connected in series	.....	.....	.....
lamps connected in parallel	.....	.....	.....

Fig. 19.13

   d i Use data from the completed Fig. 19.13 to calculate the ratio

$$\frac{\text{combined resistance of lamps connected in series}}{\text{combined resistance of lamps connected in parallel}}$$

[1]

      ii The ratios in **a** and **d i** are not equal.
By reference to Fig. 19.12, state and explain qualitatively the change in the resistance of a lamp as the potential difference is changed. [3]

Cambridge International AS and A level Physics, 9702/22 Oct/Nov 2009 Q 6

www.iranino.co

IRANINOVIN

10 a The output of a heater is 2.5 kW when connected to a 220 V supply.

i Calculate the resistance of the heater. *[2]*

ii The heater is made from a wire of cross-sectional area $2.0 \times 10^{-7}\,m^2$ and resistivity $1.1 \times 10^{-6}\,\Omega\,m$. Use your answer in **i** to calculate the length of the wire. *[3]*

b The supply voltage is changed to 110 V.

i Calculate the power output of the heater at this voltage, assuming there is no change in the resistance of the wire. *[1]*

ii State and explain quantitatively **one** way that the wire of the heater could be changed to give the same power as in **a**. *[2]*

Cambridge International AS and A Level Physics, 9702/21 May/June 2012 Q 4

11 a Define *charge*. *[1]*

b A heater is made from a wire of resistance 18.0 Ω and is connected to a power supply of 240 V. The heater is switched on for 2.60 Ms.
Calculate

i the power transformed in the heater, *[2]*

ii the current in the heater, *[1]*

iii the charge passing through the heater in this time, *[2]*

iv the number of electrons per second passing a given point in the heater. *[2]*

Cambridge International AS and A Level Physics, 9702/22 May/June 2013 Q 6

www.iranino.co

IRANINOVIN

20 D.C. circuits

By the end of this topic, you will be able to:

20.1 **(a)** recall and use appropriate circuit symbols
(b) draw and interpret circuit diagrams containing sources, switches, resistors, ammeters and voltmeters and/or any other type of component referred to in the syllabus
(c) define electromotive force (e.m.f.) in terms of the energy transferred by a source in driving unit charge round a complete circuit
(d) distinguish between e.m.f. and potential difference (p.d.) in terms of energy considerations
(e) understand the effects of the internal resistance of a source of e.m.f. on the terminal potential difference

20.2 **(a)** recall Kirchhoff's first law and appreciate the link to conservation of charge
(b) recall Kirchhoff's second law and appreciate the link to conservation of energy
(c) derive, using Kirchhoff's laws, a formula for the combined resistance of two or more resistors in series
(d) solve problems using the formula for the combined resistance of two or more resistors in series
(e) derive, using Kirchhoff's laws, a formula for the combined resistance of two or more resistors in parallel
(f) solve problems using the formula for the combined resistance of two or more resistors in parallel
(g) apply Kirchhoff's laws to solve simple circuit problems

20.3 **(a)** understand the principle of a potential divider circuit as a source of variable p.d.
(b) recall and solve problems using the principle of the potentiometer as a means of comparing potential differences

Starting points

- Basic knowledge of appropriate circuit symbols.
- Methods for drawing and interpreting circuit diagrams.
- Electric current is the rate of flow of charge.
- A potential difference is required to provide energy to move charge through a device.

20.1 Electrical circuits

When reporting an electrical experiment, or describing a circuit, it is essential to know exactly how the components are connected. This could be done by taking a photograph, but this technique has disadvantages; the photograph in Figure 19.4, for example, is not clear and does not show all the components. You could sketch a block diagram, in which the components are indicated as rectangular boxes labelled 'cell', 'ammeter', 'resistor', etc. The blocks would then be connected with lines to indicate the wiring. This is also unsatisfactory; it takes a lot of time to label all the boxes. It is much better to draw the circuit diagram using a set of symbols that is recognised by everyone and which do not need to be labelled.

Figure 20.1 shows the symbols that you are likely to need in school and college work, and which you will meet in examination questions. (You will have met many of them already.) It is important that you learn these so that you can recognise them straight away. The only labels you are likely to see on them will be the values of the components, for example 1.5V for a cell or 22Ω for a resistor.

www.iranino.co

IRANINOVIN

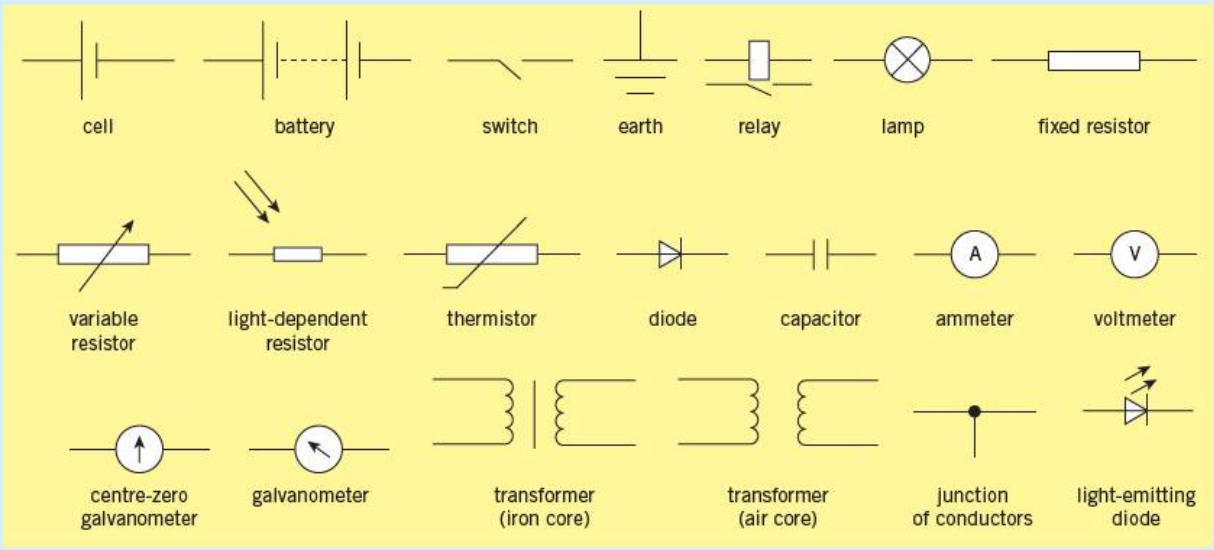


Figure 20.1 Circuit symbols

Electromotive force and potential difference

When charge passes through a power supply such as a battery, it **gains** electrical energy. The power supply is said to have an **electromotive force**, or **e.m.f.** for short. The electromotive force measures, in volts, the electrical energy gained by each coulomb of charge that passes through the power supply. Note that, in spite of its name, the e.m.f. is not a force. The energy gained by the charge comes from the chemical energy of the battery.

$$\text{e.m.f.} = \frac{\text{energy converted from other forms to electrical}}{\text{charge}}$$

When charge passes through a resistor, its electrical energy is converted to heat energy in the resistor. The resistor has a **potential difference (p.d.)** across it. The potential difference measures, in volts, the electrical energy converted for each coulomb of charge that passes through it.

$$\text{p.d.} = \frac{\text{energy converted from electrical to other forms}}{\text{charge}}$$

Example

Two lamps are connected in series to a battery. State the energy transformation that occurs in
(a) the battery,
(b) the lamps.

(a) chemical to electrical
(b) electrical to thermal (heat) and light.

Now it's your turn

1 Each lamp in the example above has a resistance R and the e.m.f. of the battery is E. The current in the circuit is I. State the rate of energy transformation in
(a) the battery,
(b) a lamp.

www.iranino.co

IRANINOVIN

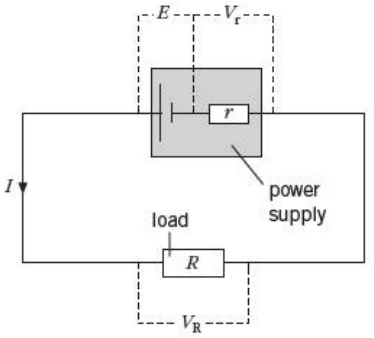


Figure 20.2

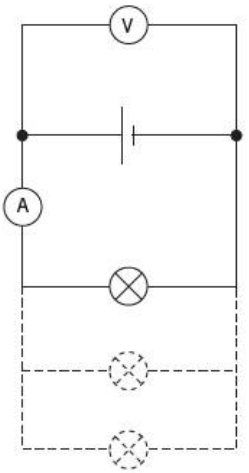


Figure 20.3 Effect of circuit current on terminal potential difference

Internal resistance

When a car engine is started with the headlights switched on, the headlights sometimes dim. This is because the car battery has resistance.

All power supplies have some resistance between their terminals, called **internal resistance**. This causes the charge circulating in the circuit to dissipate some electrical energy in the power supply itself. The power supply becomes warm when it delivers a current.

Figure 20.2 shows a power supply which has e.m.f. E and internal resistance r. It delivers a current I when connected to an external resistor of resistance R (called the load). V_R is the potential difference across the load, and V_r is the potential difference across the internal resistance. Using conservation of energy,

$$E = V_R + V_r$$

The potential difference V_R across the load is thus given by

$$V_R = E - V_r$$

V_R is called the **terminal potential difference**.

The terminal potential difference is the p.d. between the terminals of a cell when a current is being delivered.

The terminal potential difference is always less than the electromotive force when the power supply delivers a current. This is because of the potential difference across the internal resistance. The potential difference across the internal resistance is sometimes called the **lost volts**.

lost volts = e.m.f. – terminal p.d.

In contrast, the electromotive force is the terminal potential difference when the cell is on open circuit (when no current is delivered). This e.m.f. may be measured by connecting a very high resistance voltmeter across the terminals of the cell.

You can use the circuit in Figure 20.3 to show that the greater the current delivered by the power supply, the lower its terminal potential difference. As more lamps are connected in parallel to the power supply, the current increases and the lost volts, given by

lost volts = current × internal resistance

will increase. Thus the terminal potential difference decreases.

To return to the example of starting a car with its headlights switched on, a large current (perhaps 100A) is supplied to the starter motor by the battery. There will then be a large potential difference across the internal resistance; that is, the lost voltage will be large. The terminal potential difference will drop and the lights will dim.

In the terminology of Figure 20.2, $V_R = IR$ and $V_r = Ir$, so $E = V_R + V_r$ becomes

$E = IR + Ir$, or $E = I(R + r)$

The maximum current that a power supply can deliver will be when its terminals are short-circuited by a wire of negligible resistance, so that $R = 0$. In this case, the potential difference across the internal resistance will equal the e.m.f. of the cell. The terminal p.d is then zero. **Warning: do not try out this experiment, as the wire gets very hot; there is also a danger of the battery exploding.**

Quite often, in problems, the internal resistance of a supply is assumed to be negligible, so that the potential difference V_R across the load is equal to the e.m.f. of the power supply.

www.iranino.co

IRANINOVIN

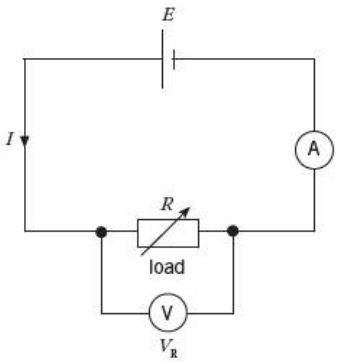


Figure 20.4 Circuit for investigating power transfer to an external load

Effect of internal resistance on power from a battery

The power delivered by a battery to a variable external load resistance can be investigated using the circuit of Figure 20.4. Readings of current I and potential difference V_R across the load are taken for different values of the variable load resistor. The product $V_R I$ gives the power dissipated in the load, and the quotient V_R/I gives the load resistance R.

Figure 20.5 shows the variation with load resistance R of the power $V_R I$ dissipated. The graph indicates that there is a maximum power delivered by the battery at one value of the external resistance. This value is equal to the internal resistance r of the battery.

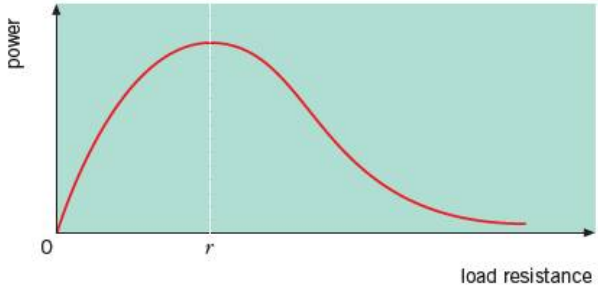


Figure 20.5 Graph of power delivered to external load against load resistance

A battery delivers maximum power to a circuit when the load resistance of the circuit is equal to the internal resistance of the battery.

Example

A high-resistance voltmeter reads 13.0 V when it is connected across the terminals of a battery. The voltmeter reading drops to 12.0 V when the battery delivers a current of 3.0 A to a lamp. State the e.m.f. of the battery. Calculate the potential difference across the internal resistance (the lost volts) when the battery is connected to the lamp. Calculate the internal resistance of the battery.

The e.m.f. is **13.0 V,** since this is the voltmeter reading when the battery is delivering negligible current.

Using $V_r = E - V_R$, lost volts = V_r =13.0 – 12.0 = **1.0 V.**

Using $V_r = Ir$, r = 1.0/3.0 = **0.33 Ω.**

Now it's your turn

2 Three identical cells, each of e.m.f. 1.5 V, are connected in series to a 15 Ω lamp. The current in the circuit is 0.27 A. Calculate the internal resistance of each cell.

3 A cell of e.m.f. 1.5 V has an internal resistance of 0.50 Ω.
   (a) Calculate the maximum current it can deliver. Under what circumstances does it deliver this maximum current?
   (b) Calculate also the maximum power it can deliver to an external load. Under what circumstances does it deliver this maximum power?

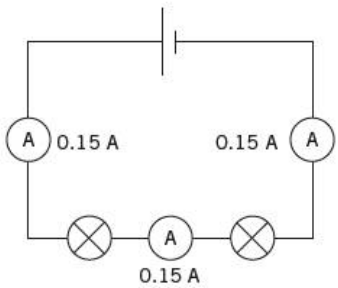


Figure 20.6 The current at each point in a series circuit is the same.

www.iranino.co

20.2 Kirchhoff's first and second laws

Conservation of charge: Kirchhoff's first law

A **series circuit** is one in which the components are connected one after another, forming one complete loop. You have probably connected an ammeter at different points in a series circuit to show that it reads the same current at each point (see Figure 20.6).

IRANINOVIN

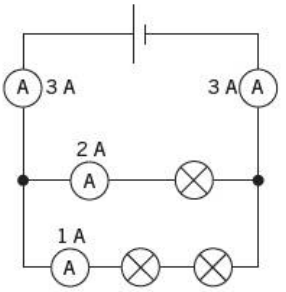


Figure 20.7 The current divides in a parallel circuit.

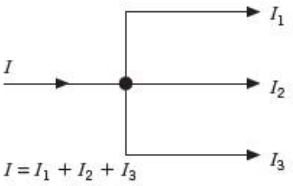


Figure 20.8

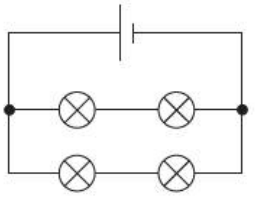


Figure 20.10

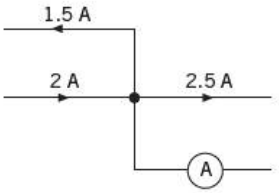


Figure 20.11

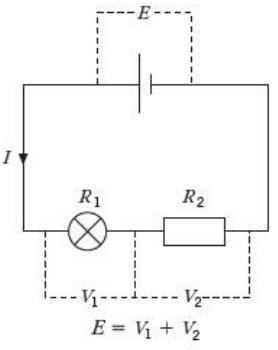


Figure 20.12

www.iranino.co

A **parallel circuit** is one where the current can take alternative routes in different loops. In a parallel circuit, the current divides at a junction, but the current entering the junction is the same as the current leaving it (see Figure 20.7). The fact that the current does not get 'used up' at a junction is because current is the rate of flow of charge, and charges cannot accumulate or get 'used up' at a junction. The consequence of this conservation of electric charge is known as **Kirchhoff's first law**. This law is usually stated as follows.

The sum of the currents entering a junction in a circuit is always equal to the sum of the currents leaving it.

At the junction shown in Figure 20.8,

$$I = I_1 + I_2 + I_3$$

Example

For the circuit of Figure 20.9, state the readings of the ammeters A_1, A_2 and A_3.

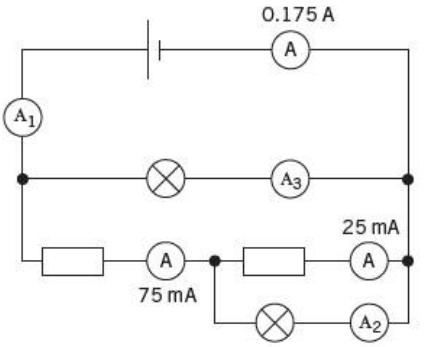


Figure 20.9

A_1 would read **175 mA**, as the current entering the power supply must be the same as the current leaving it.

A_2 would read 75 – 25 = **50 mA**, as the total current entering a junction is the same as the total current leaving it.

A_3 would read 175 – 75 = **100 mA**.

Now it's your turn

4 The lamps in Figure 20.10 are identical. There is a current of 0.50 A through the battery. What is the current in each lamp?

5 Figure 20.11 shows one junction in a circuit. Calculate the ammeter reading.

Conservation of energy: Kirchhoff's second law

Charge flowing round a circuit gains electrical energy on passing through the battery and loses electrical energy on passing through the rest of the circuit. From the law of conservation of energy, we know that the total energy must remain the same. The consequence of this conservation of energy is known as **Kirchhoff's second law**. This law may be stated as follows.

The sum of the electromotive forces in a closed circuit is equal to the sum of the potential differences.

Figure 20.12 shows a circuit containing a battery, lamp and resistor in series. Applying Kirchhoff's second law, the electromotive force in the circuit is the e.m.f. E of the battery. The sum of the potential differences is the p.d. V_1 across the lamp plus the p.d. V_2 across the resistor. Thus, $E = V_1 + V_2$. If the current in the circuit is I and the resistances of the lamp and resistor are R_1 and R_2 respectively, the p.d.s can be written as $V_1 = IR_1$ and $V_2 = IR_2$, so $E = IR_1 + IR_2$.

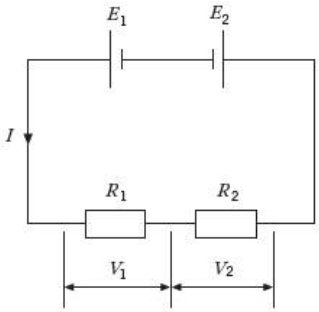


Figure 20.13

It should be remembered that both electromotive force and potential difference have direction. This must be considered when working out the equation for Kirchhoff's second law. For example, in the circuit of Figure 20.13, two cells have been connected in opposition. Here the total electromotive force in the circuit is $E_1 - E_2$, and by Kirchhoff's second law $E_1 - E_2 = V_1 + V_2 = IR_1 + IR_2$.

Resistors in series

Figure 20.14 shows two resistors of resistances R_1 and R_2 connected in series, and a single resistor of resistance R equivalent to them. The current I in the resistors, and in their equivalent single resistor, is the same.

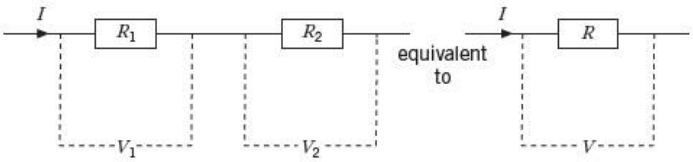


Figure 20.14 Resistors in series

The total potential difference V across the two resistors must be the same as that across the single resistor. If V_1 and V_2 are the potential differences across each resistor,

$$V = V_1 + V_2$$

But since potential difference is given by multiplying the current by the resistance,

$$IR = IR_1 + IR_2$$

Dividing by the current I,

$$R = R_1 + R_2$$

This equation can be extended so that the equivalent resistance R of several resistors connected in series is given by the expression

$$R = R_1 + R_2 + R_3 + \ldots$$

Thus

The combined resistance of resistors in series is the sum of all the individual resistances.

Resistors in parallel

Now consider two resistors of resistance R_1 and R_2 connected in parallel, as shown in Figure 20.15. The current through each will be different, but they will each have the same potential difference. The equivalent single resistor of resistance R will have the same potential difference across it, but the current will be the total current through the separate resistors.

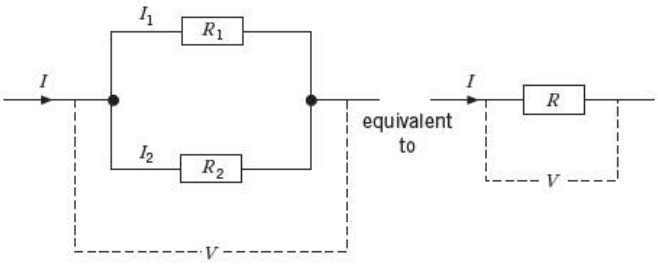


Figure 20.15 Resistors in parallel

www.iranino.co

IRANINOVIN

By Kirchhoff's first law,

$$I = I_1 + I_2$$

and using resistance = p.d./current, so $I = V/R$

$$V/R = V/R_1 + V/R_2$$

Dividing by the potential difference V,

$$1/R = 1/R_1 + 1/R_2$$

This equation can be extended so that the equivalent resistance R of several resistors connected in parallel is given by

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \ldots$$

Thus

The reciprocal of the combined resistance of resistors in parallel is the sum of the reciprocals of all the individual resistances.

Note that:

1 For two identical resistors in parallel, the combined resistance is equal to half of the value of each one.

2 For resistors in parallel, the combined resistance is always less than the value of the smallest individual resistance.

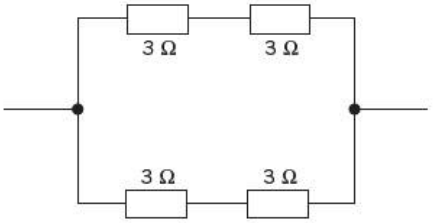


Figure 20.16

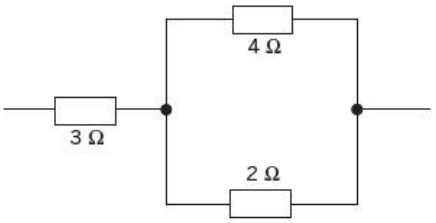
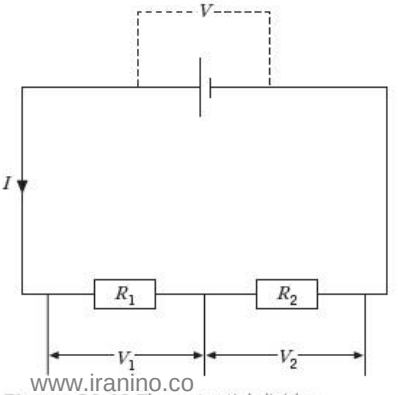


Figure 20.17



www.iranino.co

Figure 20.19 The potential divider

Example

Calculate the equivalent resistance of the arrangement of resistors in Figure 20.16.

The arrangement is equivalent to two 6Ω resistors in parallel, so the combined resistance R is given by $1/R = 1/6 + 1/6 = 2/6$. (Don't forget to find the reciprocal of this value.)

Thus $R = \mathbf{3}\,\Omega$.

Now it's your turn

6 Calculate the equivalent resistance of the arrangement of resistors in Figure 20.17. *Hint*: first find the resistance of the parallel combination.

7 Calculate the effective resistance between the points A and B in the network in Figure 20.18.

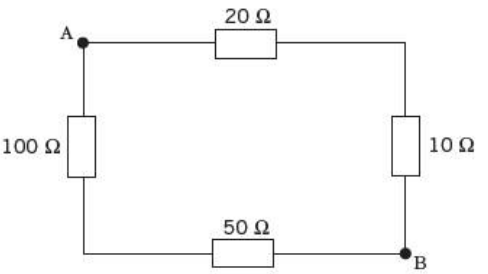


Figure 20.18

20.3 Potential dividers and potentiometers

Two resistors connected in series with a cell each have a potential difference. They may be used to divide the e.m.f. of the cell. This is illustrated in Figure 20.19.

The current in each resistor is the same, because they are in series. Thus $V_1 = IR_1$ and $V_2 = IR_2$. Dividing the first equation by the second gives $V_1/V_2 = R_1/R_2$. The ratio of

IRANINOVIN

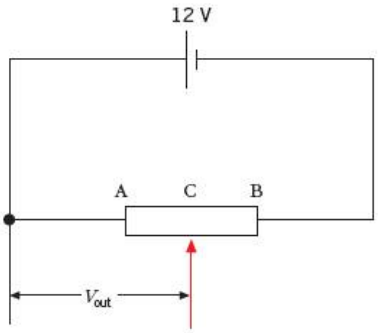


Figure 20.20 Potentiometer circuit

the voltages across the two resistors is the same as the ratio of their resistances. If the potential difference across the combination were 12V and R_1 were equal to R_2, then each resistor would have 6V across it. If R_1 were twice the magnitude of R_2, then V_1 would be 8V and V_2 would be 4V.

A **potentiometer** is a continuously variable potential divider. In Topic 19, a variable voltage supply was used to vary the voltage across different circuit components. A variable resistor, or rheostat, may be used to produce a continuously variable voltage.

Such a variable resistor is shown in Figure 20.20. The fixed ends AB are connected across the battery so that there is the full battery voltage across the whole resistor. As with the potential divider, the ratio of the voltages across AC and CB will be the same as the ratio of the resistances of AC and CB. When the sliding contact C is at the end B, the output voltage V_{out} will be 12V. When the sliding contact is at end A, then the output voltage will be zero. So, as the sliding contact is moved from A to B, the output voltage varies continuously from zero up to the battery voltage. In terms of the terminal p.d. V of the cell, the output V_{out} of the potential divider is given by

$$V_{out} = \frac{VR_1}{(R_1 + R_2)}$$

Where R_1 is the resistance of AC and R_2 is the resistance of CB.

A variable resistor connected in this way is called a potentiometer. A type of potentiometer is shown in Figure 20.21. Note the three connections.

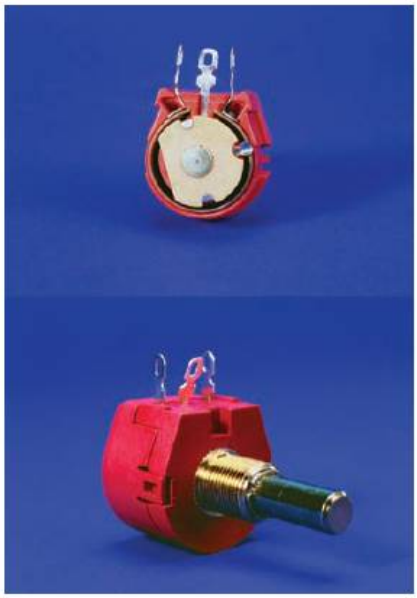


Figure 20.21 Internal and external views of a potentiometer

If a device with a variable resistance is connected in series with a fixed resistor, and the combination is connected to a cell or battery to make a potential divider, then we have the situation of a potential divider that is variable between certain limits. The device of variable resistance could be, for example, a light-dependent resistor or a thermistor. Changes in the illumination or the temperature cause a change in the resistance of one component of the potential divider, so that the potential difference across this component changes. The change in the potential difference can be used to operate control circuitry if, for example, the illumination becomes too low or too high, or the temperature falls outside certain limits. These two components are studied in detail as sensing devices in the A level course.

A potentiometer can also be used as a means of comparing potential differences. The circuit of Figure 20.22 illustrates the principle. In this circuit the variable potentiometer resistor consists of a length of uniform resistance wire, stretched along a metre rule. Contact can be made to any point on this wire using a sliding contact. Suppose that the cell A has a known e.m.f. E_A. This cell is switched into the circuit using the two-way switch. The sliding contact is then moved along the wire until the centre-zero galvanometer reads zero. The length l_A of the wire from the common zero end to the sliding contact is noted. Cell B has an unknown e.m.f. E_B. This cell is then switched into the circuit and the balancing process repeated. Suppose that the position at which the galvanometer reads zero is then a distance l_B from the common zero to the sliding contact. The ratio of the e.m.f.s is the ratio of the balance lengths; that is, $E_B/E_A = l_B/l_A$, and E_B can be determined in terms of the known e.m.f. E_A.

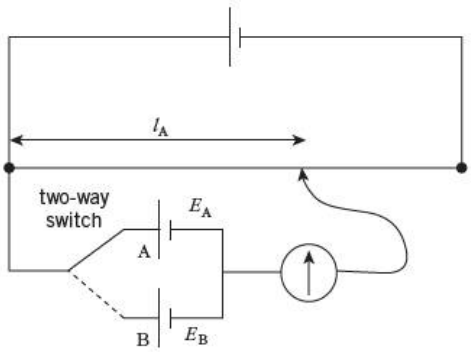


Figure 20.22 Potentiometer used to compare cell e.m.f.s

www.iranino.co

IRANINOVIN

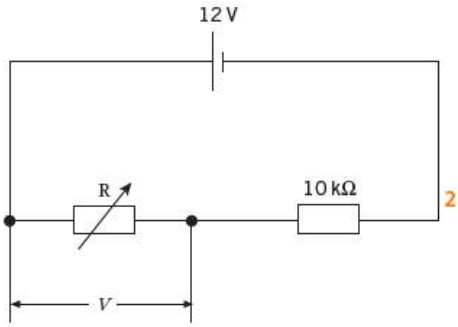


Figure 20.23

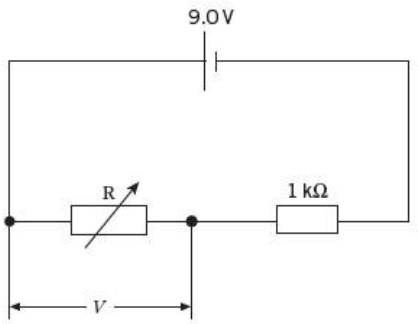


Figure 20.24

Examples

1 A light-emitting diode (LED) is connected in series with a resistor to a 5.0 V supply.
(a) Calculate the resistance of the series resistor required to give a current in the LED of 12 mA, with a voltage across it of 2.0 V.
(b) Calculate the potential difference across the LED when the series resistor has resistance 500 Ω. Assume the resistance of the LED remains constant.

(a) If the supply voltage is 5.0 V and the p.d. across the LED is 2.0 V, the p.d. across the resistor must be 5.0 – 2.0 = 3.0 V. The current through the resistor is 12 mA as it is in series with the LED. Using $R = V/I$, the resistance of the resistor is $3.0/12 \times 10^{-3} = \mathbf{250}\,\Omega$.
(b) The resistance of the LED is given by

$R = V/I = 2.0/12 \times 10^{-3} = 167\,\Omega$. If this resistance is in series with a 500 Ω resistor and a 5.0 V supply, the p.d. across the LED is $5.0 \times 167/(167 + 500) = \mathbf{1.25\,V}$.

2 The e.m.f.s of two cells are compared using the slide-wire circuit of Figure 20.22. Cell A has a known e.m.f. of 1.02 V; using this cell, a balance point is obtained when the slider is 37.6 cm from the zero of the scale. Using cell B, the balance point is at 55.3 cm.
(a) Calculate the e.m.f. of cell B.
(b) State the advantage of using this null method to compare the e.m.f.s.

(a) This is a straightforward application of the formula for the potentiometer, $E_B/E_A = l_B/l_A$. Substituting the values, $E_B = \mathbf{1.50\,V}$.
(b) When comparing the e.m.f.s of cells, it is necessary to arrange for the cells to be on open circuit so that there is no drop in terminal potential difference because of a current passing through the internal resistance. When the potentiometer is balanced, there is no current from the cell under test, which is exactly what is required.

Now it's your turn

8 Figure 20.23 shows a variable resistor R connected in series with a 10 kΩ resistor and a 12 V supply. Calculate:
(a) the p.d. V across R when it has resistance 8.0 MΩ,
(b) the p.d. V across R when it is has resistance 500 Ω,
(c) the resistance of R which makes the p.d. V across R equal to 4.0 V.

9 The resistor R in the potential divider circuit of Figure 20.24 has a resistance which varies between 100 Ω and 6.0 kΩ. Calculate the potential difference V across R when its resistance is
(a) 100 Ω,
(b) 6.0 kΩ.

Summary

- The electromotive force (e.m.f.) of a supply measures the electrical energy gained per unit of charge passing through the supply.
- The potential difference (p.d.) across a resistor measures the electrical energy converted per unit of charge passing through the resistor.
- The voltage across the terminals of a supply (the terminal p.d.) is always less than the e.m.f. of the supply when the supply is delivering a current, because of the lost volts across the internal resistance.
- For a supply of e.m.f. E which has internal resistance r, $E = I(R + r)$ where R is the external circuit resistance and I is the current in the supply.
- A supply delivers maximum power to a load when the load resistance is equal to the internal resistance of the supply.
- At any junction in a circuit, the total current entering the junction is equal to the current leaving it. This is Kirchhoff's first law, and is a consequence of the law of conservation of charge.
- In any closed loop of a circuit, the sum of the electromotive forces is equal to the sum of the potential differences. This is Kirchhoff's second law, and is a consequence of the law of conservation of energy.
- The equivalent resistance R of resistors connected in series is given by: $R = R_1 + R_2 + R_3 + \ldots$
- The equivalent resistance R of resistors connected in parallel is given by: $1/R = 1/R_1 + 1/R_2 + 1/R_3 + \ldots$
- Two resistors in series act as a potential divider, where $V_1/V_2 = R_1/R_2$. If V is the supply voltage: $V_{out} = VR_1/(R_1 + R_2)$
- A potentiometer is a variable resistor connected as a potential divider to give a continuously variable output voltage.

www.iranino.co

IRANINOVIN

Exam style questions

1 The internal resistance of a dry cell increases gradually with age, even if the cell is not being used. However, the e.m.f. remains approximately constant. You can check the age of a cell by connecting a low-resistance ammeter across the cell and measuring the current. For a new 1.5 V cell of a certain type, the short-circuit current should be about 30 A.

- **a** Calculate the internal resistance of a new cell.
- **b** A student carries out this test on an older cell, and finds the short-circuit current to be only 5 A. Calculate the internal resistance of this cell.

2 A torch bulb has a power supply of two 1.5 V cells connected in series. The potential difference across the bulb is 2.2 V, and it dissipates energy at the rate of 550 mW. Calculate:

- **a** the current through the bulb,
- **b** the internal resistance of each cell,
- **c** the heat energy dissipated in each cell in two minutes.

3 Two identical light bulbs are connected first in series, and then in parallel, across the same battery (assumed to have negligible internal resistance). Use Kirchhoff's laws to decide which of these connections will give the greater total light output.

4 You are given three resistors of resistance 22 Ω, 47 Ω and 100 Ω. Calculate:

- **a** the maximum possible resistance,
- **b** the minimum possible resistance,

that can be obtained by combining any or all of these resistors.

5 In the circuit of Fig. 20.25, the currents I_1 and I_2 are equal. Calculate:

- **a** the resistance R of the unknown resistor,
- **b** the total current I_3.

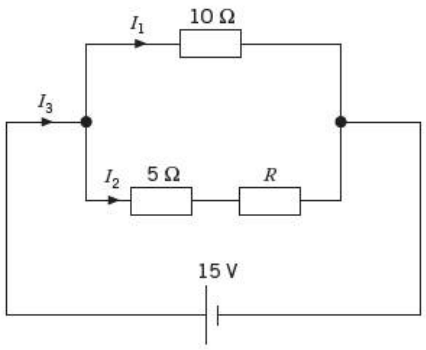


Fig. 20.25

6 Fig. 20.26 shows a potential divider circuit, designed to provide p.d.s of 1.0 V and 4.0 V from a battery of e.m.f. 9.0 V and negligible internal resistance.

- **a** Calculate the value of resistance R.
- **b** State and explain what happens to the voltage at terminal A when an additional 1.0 Ω resistor is connected between terminals B and C in parallel with the 5.0 Ω resistor. No calculations are required.

www.iranino

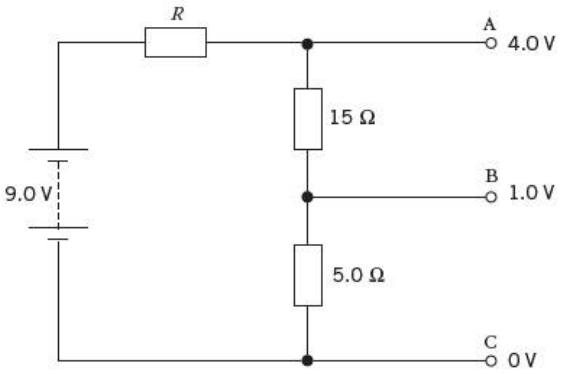


Fig. 20.26

7 A student designs an electrical method to monitor the position of a steel sphere rolling on two parallel rails. Each rail is made from bare wire of length 30 cm and resistance 20 Ω. The position-sensing circuit is shown in Fig. 20.27. The resistance of the steel sphere and the internal resistance of the battery are negligible.

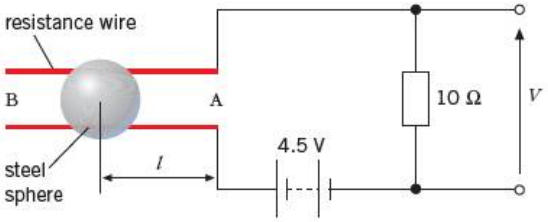


Fig. 20.27

- **a** State the voltage across the 10 Ω resistor when the sphere is at A, where $l = 0$.
- **b** With the sphere at end B of the rails, calculate:
  - **i** the total resistance of the circuit,
  - **ii** the current in the 10 Ω resistor,
  - **iii** the output voltage V.

8 Two equations for the power P dissipated in a resistor are $P = I^2R$ and $P = V^2/R$. The first suggests that the greater the resistance R of the resistor, the more power is dissipated. The second suggests the opposite: the greater the resistance, the less the power. Explain this inconsistency.

9 State the minimum number of resistors, each of the same resistance and power rating of 0.5 W, which must be used to produce an equivalent 1.2 kΩ, 5 W resistor. Calculate the resistance of each, and state how they should be connected.

10 In the circuit shown in Fig. 20.28 the current in the battery is 1.5 A. The battery has internal resistance 1.0 Ω. Calculate:

- **a** the combined resistance of the resistors that are connected in parallel in the circuit of Fig. 20.28,
- **b** the total resistance of the circuit,
- **c** the resistance of resistor Y,
- **d** the current through the 6 Ω resistor.

IRANINOVIN

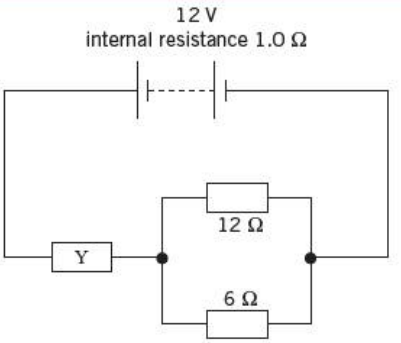


Fig. 20.28

11 The current in the starter motor of a car is 160 A when starting the engine. The connecting cable has total length 1.3 m, and consists of 15 strands of wire, each of diameter 1.2 mm. The resistivity of the metal of the strands is $1.4 \times 10^{-8}\,\Omega\,m$.

a Calculate:

i the resistance of each strand,

ii the total resistance of the cable,

iii the power loss in the cable.

b When the starter motor is used to start the car, 700 C of charge pass through a given cross-section of the cable.

i Assuming that the current is constant at 160 A, calculate for how long the charge flows.

ii Calculate the number of electrons which pass a given cross-section of the cable in this time. The electron charge e is -1.6×10^{-19} C.

c The e.m.f. of the battery is 13.6 V and its internal resistance is 0.012 Ω. Calculate:

i the potential difference across the battery terminals when the current in the battery is 160 A,

ii the rate of production of heat energy in the battery.

12 A copper wire of length 16 m has a resistance of 0.85 Ω. The wire is connected across the terminals of a battery of e.m.f. 1.5 V and internal resistance 0.40 Ω.

a Calculate the potential difference across the wire and the power dissipated in it.

b In an experiment, the length of this wire connected across the terminals of the battery is gradually reduced.

i Sketch a graph to show how the power dissipated in the wire varies with the connected length.

ii Calculate the length of the wire when the power dissipated in the wire is a maximum.

iii Calculate the maximum power dissipated in the wire.

13 a **i** State Kirchhoff's second law. *[1]*

ii Kirchhoff's second law is linked to the conservation of a certain quantity. State this quantity. *[1]*

b The circuit shown in Fig. 20.29 is used to compare potential differences.

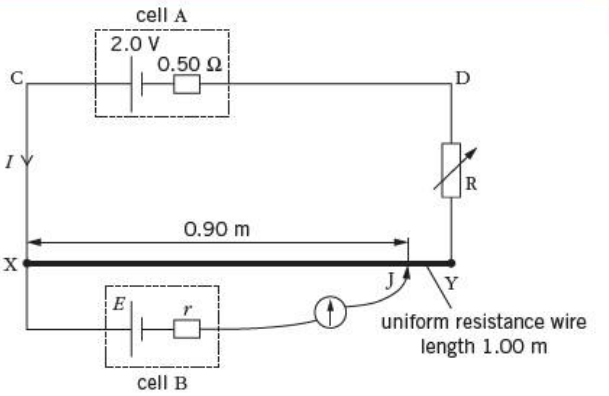


Fig. 20.29

The uniform resistance wire XY has length 1.00 m and resistance 4.0 Ω. Cell A has e.m.f. 2.0 V and internal resistance 0.50 Ω. The current through cell A is I. Cell B has e.m.f. E and internal resistance r.

The current through cell B is made zero when the movable connection J is adjusted so that the length of XJ is 0.90 m. The variable resistor R has resistance 2.5 Ω.

i Apply Kirchhoff's second law to the circuit CXYDC to determine the current I. *[2]*

ii Calculate the potential difference across the length of wire XJ. *[2]*

iii Use your answer in **ii** to state the value of E. *[1]*

iv State why the value of the internal resistance of cell B is not required for the determination of E. *[1]*

Cambridge International AS and A level Physics 9702/21 May/June 2012

14 A thermistor has resistance 3900 Ω at 0°C and resistance 1250 Ω at 30°C. The thermistor is connected into the circuit of Fig. 20.30 in order to monitor temperature changes.

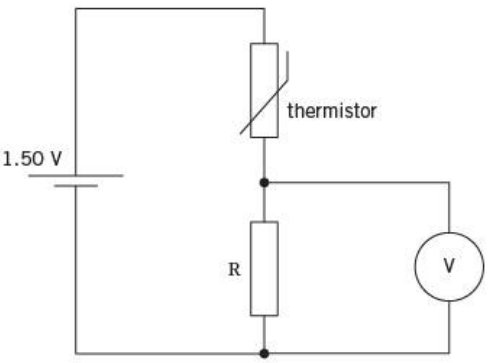


Fig. 20.30

The battery of e.m.f. 1.50 V has negligible internal resistance and the voltmeter has infinite resistance.

a The voltmeter is to read 1.00 V at 0°C. Show that the resistance of resistor R is 7800 Ω. *[2]*

www.iranino.co

IRANINOVIN

b The temperature of the thermistor is increased to 30°C. Determine the reading on the voltmeter. *[2]*

c The voltmeter in Fig. 20.30 is replaced with one having a resistance of 7800 Ω. Calculate the reading on this voltmeter for the thermistor at a temperature of 0°C. *[2]*

Cambridge International AS and A level Physics, 9702/02 May/June 2004 Q 8

15 A car battery has an internal resistance of 0.060 Ω. It is re-charged using a battery charger having an e.m.f. of 14 V and an internal resistance of 0.10 Ω, as shown in Fig. 20.31.

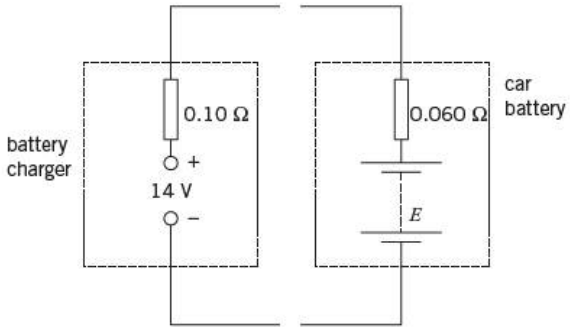


Fig. 20.31

a At the beginning of the re-charging process, the current in the circuit is 42 A and the e.m.f. of the battery is E (measured in volts).

i For the circuit of Fig. 20.31, state

1 the magnitude of the total resistance,

2 the total e.m.f. in the circuit. Give your answer in terms of E. *[2]*

ii Use your answers to **i** and data from the question to determine the e.m.f. of the car battery at the beginning of the re-charging process. *[2]*

b For the majority of the charging time of the car battery, the e.m.f. of the car battery is 12 V and the charging current is 12.5 A. The battery is charged at this current for 4.0 hours. Calculate, for this charging time,

i the charge that passes through the battery, *[2]*

ii the energy supplied from the battery charger, *[2]*

iii the total energy dissipated in the internal resistance of the battery charger and the car battery. *[2]*

c Use your answers in **b** to calculate the percentage efficiency of transfer of energy from the battery charger to stored energy in the car battery. *[2]*

Cambridge International AS and A level Physics, 9702/02 May/June 2007 Q 6

16 A circuit used to measure the power transfer from a battery is shown in Fig. 20.32. The power is transferred to a variable resistor of resistance R.

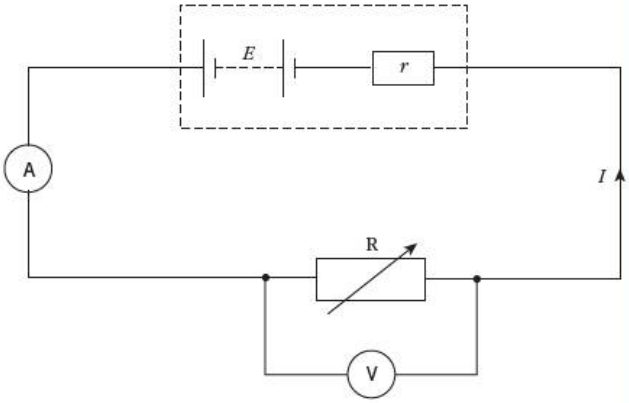


Fig. 20.32

The battery has an electromotive force (e.m.f.) E and an internal resistance r. There is a potential difference (p.d.) V across R. The current in the circuit is I.

a By reference to the circuit shown in Fig. 20.32, distinguish between the definitions of e.m.f. and p.d. *[3]*

b Using Kirchoff's second law, determine an expression for the current I in the circuit. *[1]*

c The variation with current I of the p.d. V across R is shown in Fig. 20.33.

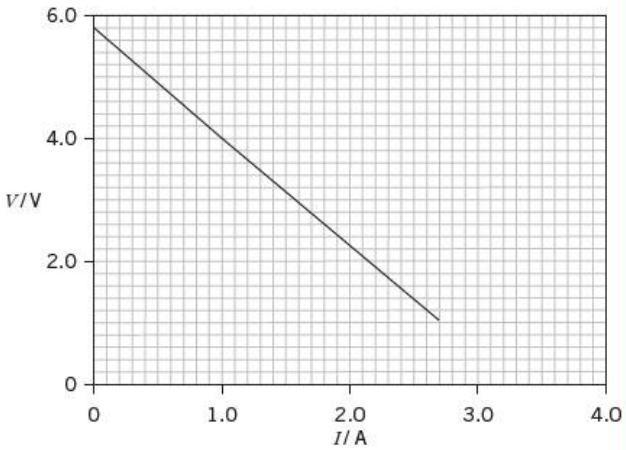


Fig. 20.33

Use Fig. 20.33 to determine

i the e.m.f. E, *[1]*

ii the internal resistance r. *[2]*

d i Using the data from Fig. 20.33, calculate the power transferred to R for a current of 1.6 A. *[2]*

ii Use your answers from **c i** and **d i** to calculate the efficiency of the battery for a current of 1.6 A. *[2]*

Cambridge International AS and A Level Physics, 9702/23 Oct/Nov 2012 Q 4

www.iranino.co

IRANINOVIN

26 Particle physics

By the end of this topic, you will be able to:

26.1 **(a)** infer from the results of the α-particle scattering experiment the existence and small size of the nucleus
(b) describe a simple model for the nuclear atom to include protons, neutrons and orbital electrons
(c) distinguish between nucleon number and proton number
(d) understand that an element can exist in various isotopic forms, each with a different number of neutrons
(e) use the usual notation for the representation of nuclides
(f) appreciate that nucleon number, proton number and mass-energy are all conserved in nuclear processes
(g) show an understanding of the nature and properties of α-, β- and γ-radiations (both β^- and β^+ are included)
(h) state that (electron) antineutrinos and (electron) neutrinos are produced during β^- and β^+ decay

26.2 **(a)** appreciate that protons and neutrons are not fundamental particles since they contain quarks
(b) describe a simple quark model of hadrons in terms of up, down and strange quarks and their respective antiquarks
(c) describe protons and neutrons in terms of a simple quark model
(d) appreciate that there is a weak interaction between quarks, giving rise to β decay
(e) describe β^- and β^+ decay in terms of a simple quark model
(f) appreciate that electrons and neutrinos are leptons

Starting points

- The atom consists of a very small nucleus containing protons and neutrons, surrounded by orbiting electrons.
- The decay of unstable nuclei leads to emissions.
- Appreciate that protons and neutrons are not fundamental particles.

26.1 Atomic structure and radioactivity

The atoms of all elements are made up of three particles called **protons**, **neutrons**, and **electrons**. The protons and neutrons are at the centre or nucleus of the atom. The electrons orbit the nucleus.

We shall see later that the diameter of the nucleus is only about 1/10000 of the diameter of an atom.

Figure 26.1 illustrates very simple models of a helium atom and a lithium atom.

The protons and neutrons both have a mass of about one atomic mass unit u ($1\,u = 1.66 \times 10^{-27}\,kg$). The atomic mass unit is defined and used in the A level course in Topic 10. By comparison, the mass of an electron is very small, about 1/2000 of 1 u. The vast majority of the mass of the atom is therefore in the nucleus.

www.iranino.co

IRANINOVIN

The basic properties of the proton, neutron and electron are summarised in Table 26.1.

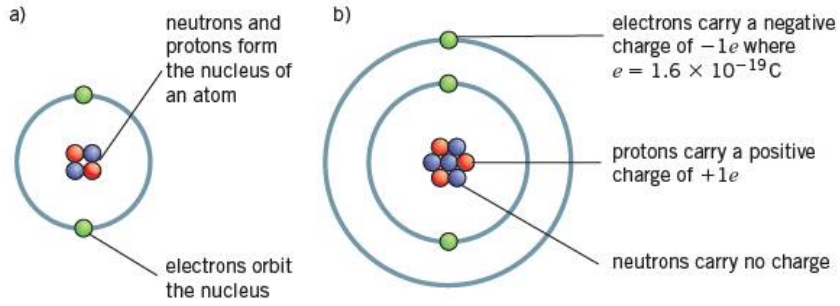


Figure 26.1 Structures of a) a helium atom and b) a lithium atom

Table 26.1

	approximate mass	charge	position
proton	u	$+e$	in nucleus
neutron	u	0	in nucleus
electron	u/2000	$-e$	orbiting nucleus

Atoms and ions

Atoms are uncharged because they contain equal numbers of protons and electrons and the charge on an electron is equal and opposite to the charge on a proton. If an atom loses one or more electrons, so that it does not contain an equal number of protons and electrons, it becomes charged and is called an **ion**.

For example, if a sodium atom loses one of its electrons, it becomes a positive sodium ion.

$$\underset{\text{sodium atom}}{\text{Na}} \rightarrow \underset{\text{sodium ion}}{\text{Na}^+} + \underset{\text{electron}}{\text{e}^-}$$

If an atom gains an electron, it becomes a negative ion.

Proton number and nucleon number

The number of protons in the nucleus of an atom is called the **proton number** (or **atomic number**) Z.

The number of protons together with the number of neutrons in the nucleus is called the **nucleon number** (or **mass number**) A.

A **nucleon** is the name given to either a proton or a neutron in the nucleus.

The difference between the nucleon number (A) and the proton number (Z) gives the number of neutrons in the nucleus.

Representation of nuclides

If the chemical symbol of an element is X, a particular atom of this element, a nuclide, is represented by the notation

$$\overset{\text{nucleon number}}{\underset{\text{proton number}}{\text{X}}} = {}^{A}_{Z}\text{X}.$$

The element changes for every Z number and the symbol X changes. A nuclide is the name given to a class of atoms whose nuclei contain a specified number of protons and a specified number of neutrons. The nucleus of one form of sodium contains 11 protons and 12 neutrons. Therefore its proton number Z is 11 and the nucleon number A is 11 + 12 = 23. This nuclide can be shown as ${}^{23}_{11}\text{Na}$. All atoms with nuclei that contain 11 protons and 12 neutrons belong to this class and are the same nuclide.

www.iranino.co

IRANINOVIN

Nuclides with nuclei that have the same atomic number Z but a different mass number are **isotopes** of the same element.

Example

An oxygen nucleus is represented by $^{16}_{8}O$. Describe its atomic structure.

The nucleus has a proton number of 8 and a nucleon number of 16. Thus, its nucleus contains **8 protons** and 16 – 8 = **8 neutrons**. There are also **8 electrons** (equal to the number of protons) orbiting the nucleus.

Now it's your turn

1 Write down the proton number and the nucleon number for the potassium nucleus $^{40}_{19}K$. Deduce the number of neutrons in the nucleus.

Isotopes

Sometimes atoms of the same element have different numbers of neutrons in their nuclei. The most abundant form of chlorine contains 17 protons and 18 neutrons in its nucleus, giving it a nucleon number of 17 + 18 = 35. This is often called chlorine-35. Another form of chlorine contains 17 protons and 20 neutrons in the nucleus, giving it a nucleon number of 37. This is chlorine-37. Chlorine-35 and chlorine-37 are said to be **isotopes** of chlorine.

Isotopes are different forms of the same element which have the same number of protons but different numbers of neutrons in their nuclei.

Some elements have many isotopes, but others have very few. For hydrogen, the most common isotope is hydrogen-1. Its nucleus is a single proton. Hydrogen-2 is called deuterium; its nucleus contains one proton and one neutron. Hydrogen-3, with one proton and two neutrons, is called tritium.

Note that the term isotope is also used to describe nuclei with the same proton number (that is, nuclei of the same element) but with different nucleon numbers. You may also come across the term **nuclide**.

A nuclide is one type of nucleus with a particular nucleon number and a particular proton number.

α-particles, β-particles and γ-radiation

Some elements have nuclei which are unstable. That is, the combination of protons and neutrons in the nucleus is such that the forces acting on the nucleons do not balance. In order to become more stable, they emit particles and/or electromagnetic radiation. The nuclei are said to be **radioactive**, and the emission is called **radioactivity**. The emissions are invisible to the eye, but their tracks were first made visible in a device called a cloud chamber. The photograph in Figure 26.2 shows tracks created by one type of emission, α-particles.

Investigations of the nature and properties of the emitted particles or radiation show that there are three different types of emission. The three types are α-particles (alpha-particles), β-particles (beta-particles) and γ-radiation (gamma radiation). All three emissions originate from the nucleus.

www.iranino.co

IRANINOVIN

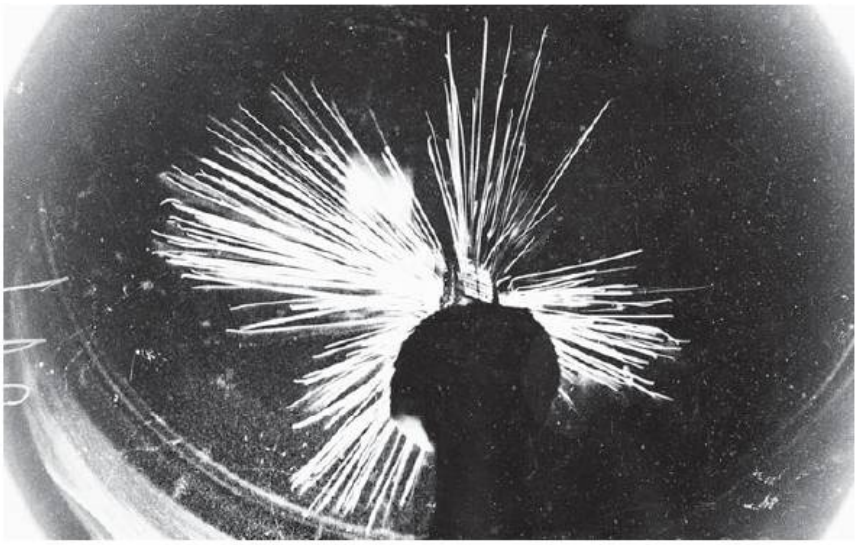


Figure 26.2 Tracks of α-particles

α-particles

Like a helium nucleus, an α-particle contains two protons and two neutrons and, hence, carries a charge of $+2e$. α-particles travel at speeds of up to about 10^7 m s^{-1} (about 5% of the speed of light). α-particle emission is the least penetrating of the three types of emission. It can pass through very thin paper, but is unable to penetrate thin card. Its range in air is a few centimetres. Because α-particles are charged, they can be deflected by electric and magnetic fields.

An α-particle is identical to the nucleus of a helium atom.

In terms of symbols:

An α-particle is written as ^4_2He.

As α-particles travel through matter, they interact with nearby atoms causing them to lose one or more electrons. The ionised atom and the dislodged electron are called an ion pair. The production of an ion pair requires the separation of unlike charges, and this process requires energy. α-particles have a relatively large mass and charge, and consequently they are efficient ionisers. They may produce as many as 10^5 ion pairs for every centimetre of air through which they travel. Thus, they lose energy relatively quickly, and have low penetrating power.

When the nucleus of an atom emits an α-particle, it is said to undergo α-decay. The nucleus loses two protons and two neutrons in this emission.

In α-decay, the proton number of the nucleus decreases by two, and the nucleon number decreases by four.

Each element has a particular proton number, and therefore α-decay causes one element to change into another. (This process is sometimes called **transmutation**.) The original nuclide is called the **parent nuclide**, and the new one the **daughter nuclide**.

For example, uranium-234 (the parent nuclide) may emit an α-particle. The daughter nuclide is thorium-230. In addition, energy is released. This emission is represented by the nuclear equation

$$^{234}_{92}\text{U} \rightarrow {}^{230}_{90}\text{Th} + {}^4_2\text{He} + \text{energy}$$

The atomic mass of the decay products is less than the mass of the parent nuclide ($^{234}_{92}$U). The energy equivalent of the difference in the mass appears as kinetic energy of the α-particle and the recoiling daughter nuclide ($^{230}_{90}$Th) and a γ-photon. Therefore,

www.iranino.co

IRANINOVIN

mass–energy is conserved. Linear momentum is also conserved in this type of nuclear reaction. The same amount of energy is released in the decay of each nucleus of $^{234}_{92}U$. The α-particles emitted from a particular radioactive nuclide have the same kinetic energy.

β-particles

A radioactive nucleus that decays by β decay may emit a negative (β^-) or positive (β^+) electron. The positive electron (β^+) is also known as a positron or an antielectron ($\bar{e}$).

β-particles are fast moving electrons, β^-, or positrons, β^+.

β-particles have speeds in excess of 99% of the speed of light. These particles have half the charge and very much less mass than α-particles. Consequently, they are much less efficient than α-particles in producing ion pairs. They are, thus, far more penetrating than α-particles, being able to travel up to about a metre in air. They can penetrate card and sheets of aluminium up to a few millimetres thick. Their charge means that they are affected by electric and magnetic fields. However, there are important differences between the behaviour of α- and β-particles in these fields. β-particles may carry negative charge or positive charge, and thus may be deflected in the same direction or opposite direction to the positively charged α-particles. β-particles experience a much larger deflection when moving at the same speed as α-particles, because the mass of a β-particle is much less than that of an α-particle.

A β^- particle may be emitted from a lead-214 nucleus (the parent nuclide). The daughter nuclide is bismuth-214 and, in addition, energy is released. The emission is represented by the nuclear equation

$$^{214}_{82}Pb \rightarrow {}^{214}_{83}Bi + {}^{0}_{-1}e + {}^{0}_{0}\bar{\nu} + \text{energy}$$

A β^+ particle may be emitted from a phosphorus-30 nucleus (the parent nuclide). The daughter nuclide is silicon-30 and energy is also released. The emission is represented by the nuclear equation:

$$^{30}_{15}P \rightarrow {}^{30}_{14}Si + {}^{0}_{1}e + {}^{0}_{0}\nu + \text{energy}$$

The symbols ${}^{0}_{0}\nu$ and ${}^{0}_{0}\bar{\nu}$ represent a neutrino and an antineutrino respectively. These particles have no electrical charge and little or no mass and are emitted from the nucleus at the same time as the β-particle.

It was stated on page 168 that the nucleus contains protons and neutrons. What, then, is the origin of β-particle emission? Each β-particle certainly comes from a nucleus, not from the electrons outside the nucleus. The process for this type of decay is that, just prior to β^- emission, a neutron in the nucleus forms a proton, a negative electron and an antineutrino. The ratio of protons to neutrons in the nucleus is changed and this makes the daughter nucleus more stable.

In fact, *free* neutrons are known to decay as follows:

$$^{1}_{0}n \rightarrow {}^{1}_{1}p + {}^{0}_{-1}e + {}^{0}_{0}\bar{\nu} + \text{energy}$$

A similar process happens in the nucleus. In β^- decay, a negative electron and antineutrino $\bar{\nu}$ are emitted from the nucleus. This leaves the nucleus with the same number of nucleons as before, but with one extra proton and one fewer neutron.

In β^+ emission, a proton in the nucleus forms a neutron, a positive electron and a neutrino. This process again changes the ratio of protons to neutrons in the nucleus and makes the daughter nucleus more stable.

In β^+ decay the proton is considered to transform itself as follows:

$$^{1}_{1}p \rightarrow {}^{1}_{0}n + {}^{0}_{1}e + {}^{0}_{0}\nu + \text{energy}$$

In β^+ decay, the positive electron and a neutrino are emitted from the nucleus. This leaves the nucleus with the same number of nucleons as before, but with one extra neutron and one fewer proton.

In β^- decay (negative electron), a daughter nuclide is formed with the proton number increased by one, but with the same nucleon number.

www.iranino.co

IRANINOVIN

In β^+ decay (positive electron), a daughter nuclide is formed with the proton number decreased by one, but with the same nucleon number.

The antimatter particle, the positive electron, very quickly meets its equivalent matter particle, the negative electron. The two particles annihilate each other to produce γ-radiation. This makes the positive electron difficult to detect.

The atomic mass of the decay products is less than the mass of the parent nucleus. The energy equivalent of the difference in the mass is shared between the kinetic energy of the β-particle and the recoiling daughter nucleus and the energy of the neutrino or antineutrino. Therefore mass–energy is conserved. The same amount of energy is released in the decay of each particular parent nucleus. However, the electrons emitted from a particular radioactive nuclide have varying amounts of kinetic energy. The amount depends on the way the total energy available is shared between the electron and the neutrino. The sum of the electron's energy and the neutrino's energy is constant for the decay of a particular nuclide.

Kinetic energy of the subatomic particles

The SI unit of energy is the joule. The energies in nuclear reactions are very small compared to the joule. A more convenient unit to use is the electron-volt (eV). This is the work done (energy gained) by an electron when accelerated through a potential difference of one volt. Since work done equals potential difference × charge, one eV is equivalent to 1.602×10^{-19} J. One mega electron-volt (MeV) is 10^6 eV or 1.602×10^{-13} J.

Examples

1 A strontium-90 atom (the parent nuclide) may decay with the emission of a β-particle to form the daughter nuclide yttrium-90. The decay is represented by the nuclear equation

$$^{90}_{38}\text{Sr} \rightarrow {}^{90}_{39}\text{Y} + \beta + x + \text{energy}$$

State and explain whether the β-particle is a negative or positive electron. State the type of particle represented by x.

The proton number has increased by one, hence a **negative electron** is emitted. The x is an **antineutrino** as this particle is emitted with a negative electron.

2 Calculate the energy in joules of 1 GeV.

potential difference = energy transformed / charge

energy = 10^9 (V) × charge on electron (e) = 1.6×10^{-10} J

Hence 1 GeV = $\mathbf{1.6 \times 10^{-10}\ J}$

Now it's your turn

2 The mass of strontium-90 in the example above does not equal the total mass of particles formed after the reaction. Explain why mass seems not to be conserved.

3 The energy released in the example above is 3.6 MeV. Calculate the energy in joules.

γ-radiation

γ-radiation is part of the electromagnetic spectrum with wavelengths between 10^{-11} m and 10^{-13} m.

Since γ-radiation has no charge, its ionising power is much less than that of either α- or β-particles. γ-radiation penetrates almost unlimited thicknesses of air, several metres of concrete or several centimetres of lead.

α- and β-particles are emitted by unstable nuclei which have excess energy. The emission of these particles results in changes in the ratio of protons to neutrons, but the nuclei may still have excess energy. The nucleus may return to its unexcited (or ground) state by emitting energy in the form of γ-radiation.

In γ-emission, no particles are emitted and there is, therefore, no change to the proton number or nucleon number of the parent nuclide.

www.iranino.co

IRANINOVIN

For example, when uranium-238 decays by emitting an α-particle, the resulting nucleus of thorium-234 contains excess energy (it is in an excited state) and emits a photon of γ-radiation to return to the ground state. This process is represented by the nuclear equation

$$^{234}_{90}\text{Th}^* \rightarrow {}^{234}_{90}\text{Th} + \gamma$$

The * next to the symbol Th on the left-hand side of the equation shows that the thorium nucleus is in an excited state.

Note that in all radioactive decay processes (and, in fact, in all processes of nuclear reactions) nucleon number and proton number are conserved. Hence, for all equations representing nuclear reactions, the sum of the numbers at the top of the symbols on the left-hand side of the equation (the sum of the nucleon numbers) is equal to the sum of the nucleon numbers on the right-hand side. Similarly, the sum of the numbers at the bottom of the symbols on the left-hand side (the sum of the proton numbers) is equal to the sum of the proton numbers on the right-hand side. Energy and mass, taken together, are also conserved in all nuclear processes.

Summary of the properties of radioactive emissions

Table 26.2 summarises the properties of α-particles, β-particles and γ-radiation.

Table 26.2

property	α-particle	β-particle	γ-radiation
mass	4 u	about u/2000	0
charge	$+2e$	$-e$ or $+e$	0
nature	helium nucleus (2 protons + 2 neutrons)	negative or positive electron	short-wavelength electromagnetic waves
speed	up to 0.05c	more than 0.99c	c
penetrating power	few cm of air	few mm of aluminium	few cm of lead
relative ionising power	10^4	10^2	1
affects photographic film?	yes	yes	yes
deflected by electric, magnetic fields?	yes, see Figure 26.3	yes, see Figure 26.3	no

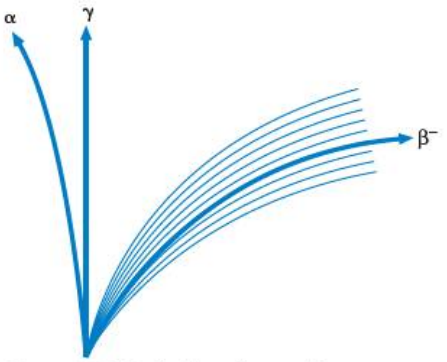


Figure 26.3 Deflection of α- and β^--particles and γ-radiation by a magnetic field into the page

Figure 26.3 illustrates a hypothetical demonstration of the effect of a magnetic field on α, β^- and γ emissions. The direction of the magnetic field is perpendicularly into the page. γ-radiation is uncharged, and is not deflected by the magnetic field. Because α- and β^--particles have opposite charges, they are deflected in opposite directions. Note that the deviation of α-particles is, in general, much less than that of β-particles and the relative deflections are not shown to scale. The α-particles from a particular nuclide all deviate by the same amount. The β-particles from a particular nuclide, by contrast, have a range of deflections indicating that they have a range of energies. Deflections can be observed with an electric field. (Make sure that you can confirm that an electric field should be in the plane of the paper and with a direction horizontal and to the left in order to obtain the deflections in the same direction as those obtained with the magnetic field.) Note, β^+-particles would be deflected in the same direction as the α-particles, with a similar pattern to that of the β^--particles.

Radioactive decay series

The daughter nuclide of a radioactive decay may, itself, be unstable and so may emit radiation to give another different nuclide. This sequence of radioactive decay from parent nuclide through succeeding daughter nuclides is called a **radioactive decay series**. The series ends when a stable nuclide is reached.

www.iranino.co

IRANINOVIN

Table 26.3 Part of the decay series of uranium-238

decay	radiation emitted
$^{238}_{92}U \rightarrow {}^{234}_{90}Th + {}^{4}_{2}He + \gamma$	α, γ
$^{234}_{90}Th \rightarrow {}^{234}_{91}Pa + {}^{0}_{-1}e + \gamma$	β^-, γ
$^{334}_{91}Pa \rightarrow {}^{234}_{92}U + {}^{0}_{-1}e + \gamma$	β^-, γ
$^{234}_{92}U \rightarrow {}^{230}_{90}Th + {}^{4}_{2}He + \gamma$	α, γ
$^{230}_{90}Th \rightarrow {}^{226}_{88}Ra + {}^{4}_{2}He + \gamma$	α, γ
$^{226}_{88}Ra \rightarrow {}^{226}_{86}Rn + {}^{4}_{2}He$	α
$^{222}_{86}Rn \rightarrow {}^{218}_{84}Po + {}^{4}_{2}He$	α

Part of such a radioactive decay series, the uranium series, is shown in Table 26.3.

Detecting radioactivity

Some of the methods used to detect radioactive emissions are based on the ionising properties of the particles or radiation.

The Geiger counter

Figure 26.4 illustrates a Geiger-Müller tube with a scaler connected to it. When radiation enters the window, it creates ion pairs in the gas in the tube. These charged particles, and particularly the electrons, are accelerated by the potential difference between the central wire anode and the cylindrical cathode. These accelerated particles then cause further ionisation. The result of this continuous process is described as an **avalanche effect**. That is, the entry of one particle into the tube and the production of one ion pair results in very large numbers of electrons and ions arriving at the anode and cathode respectively. This gives a pulse of charge which is amplified and counted by the scaler or ratemeter. (A scaler measures the total count of pulses in the tube during the time that the scaler is operating. A ratemeter continuously monitors the number of counts per second.) Once the pulse has been registered, the charges are removed from the gas in readiness for further radiation entering the tube.

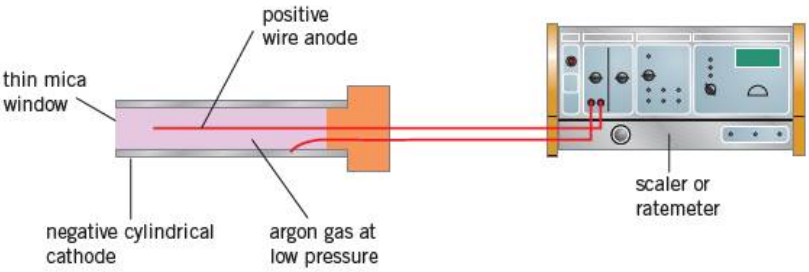


Figure 26.4 Geiger-Müller tube and scaler



Figure 26.5 Film badge dosimeter

Photographic plates

When a radioactive emission strikes a photographic film, the film reacts as if it had been exposed to a small amount of visible light. When the film is developed, fogging or blackening is seen. This fogging can be used to detect, not only the presence of radioactivity, but also the dose of the radiation.

Figure 26.5 shows a film badge dosimeter. It contains a piece of photographic film which becomes fogged when exposed to radiation. Workers who are at risk from radiation wear such badges to gauge the type and dose of radiation to which they have been exposed. The radiation passes through different filters before reaching the film. Consequently, the type of radiation, as well as the quantity, can be assessed.

The scintillation counter

Early workers with radioactive materials used glass screens coated with zinc sulfide to detect radiation. When radiation is incident on the zinc sulfide, it emits a tiny pulse of light called a **scintillation**. The rate at which these pulses are emitted indicates the intensity of the radiation.

The early researchers worked in darkened rooms, observing the zinc sulfide screen by eye through a microscope and counting the number of flashes of light occurring in a certain time. Now a scintillation counter is used (Figure 26.6).

Often a scintillator crystal is used instead of a zinc sulfide screen. The crystal is mounted close to a device known as a photomultiplier, a vacuum-tube device which uses the principle of photoelectric emission (see Topic 25). Flashes of light cause the emission of photoelectrons from the negative electrode of the photomultiplier. The photoelectric current is amplified inside the tube. The output electrode is connected to a scaler or ratemeter, as with the Geiger-Müller tube.

www.iranino.co

IRANINOVIN

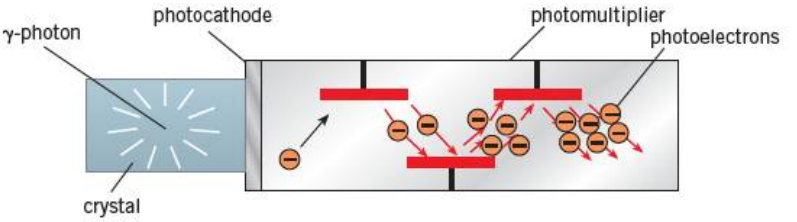


Figure 26.6 Scintillation counter

Background radiation

Radioactivity is a natural phenomenon. Rocks such as granite contain small amounts of radioactive nuclides, some foods we eat emit radiation, and even our bodies are naturally radioactive. Although the atmosphere provides life on Earth with some shielding, there is, nevertheless, some radiation from outer space (cosmic radiation). In addition to this natural radioactivity, we are exposed to radiation from man-made sources. These are found in medicine, in fallout from nuclear explosions, and in leaks from nuclear power stations. The sum of all this radiation is known as **background radiation**. Figure 26.7 indicates the relative proportions of background radiation coming from various sources.

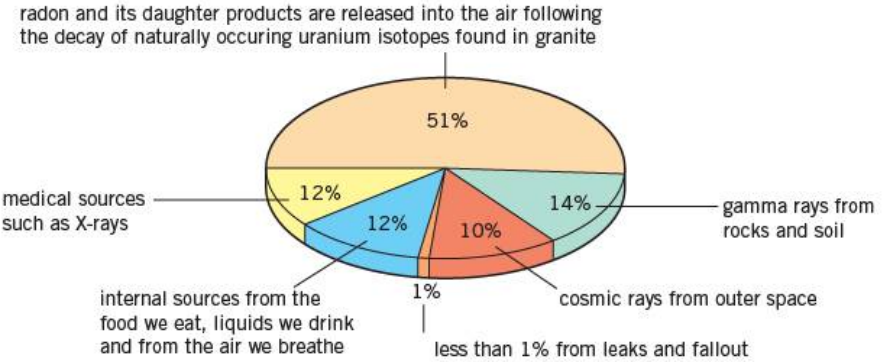


Figure 26.7 Sources of background radiation

In carrying out experiments with radioactive sources, it is important to take account of background radiation. In order to determine the count-rate due to the radioactive source, the background count-rate must be subtracted from the total measured count-rate. Allowance for background radiation gives the corrected count-rate.

The spontaneous and random nature of radioactive decay

Detection of the count-rate of radioactive sources shows that the emission of radiation is both *spontaneous* and *random*. It is a spontaneous process because it is not affected by any external factors, such as temperature or pressure. Decay is random in that it is not possible to predict which nucleus in a sample will decay next. There is, however, a constant probability (or chance) that a nucleus will decay in any fixed period of time. We will look at this in more detail in the A level course Topic 26.

Probing matter

Figure 26.8 shows a photograph taken with an ion microscope, a device which makes use of the de Broglie wavelength of gas ions (see Topic 25). It shows a sample of iridium at a magnification of about five million. The positions of individual iridium atoms can be seen.

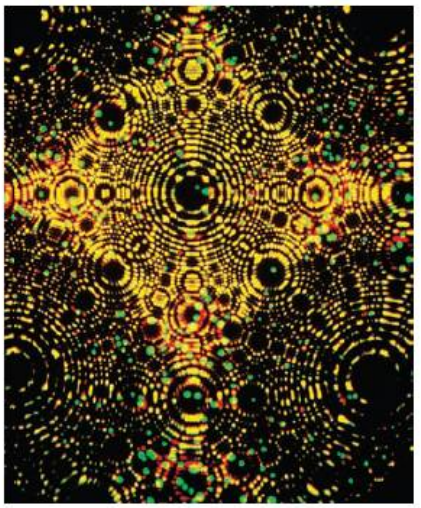


Figure 26.8 Ion microscope photograph of iridium

Photographs like this reinforce the idea that all matter is made of very small particles that we call atoms. Advances in science at the end of the nineteenth and the beginning of the twentieth century led most physicists to believe that atoms themselves are

www.iranino.co

IRANINOVIN

made from even smaller particles, some of which have positive or negative charges. Unfortunately, even the most powerful microscopes cannot show us the internal structure of the atom. Many theories were put forward about the structure of the atom, but it was a series of experiments carried out by Ernest Rutherford and his colleagues around 1910 that led to the birth of the model we now know as the **nuclear atom**.

Probing matter using α-particles

In 1911, Rutherford and two of his associates, Geiger and Marsden, fired a beam of α-particles at a very thin piece of gold foil. A zinc sulfide detector was moved around the foil to detect the directions in which the α-particles travelled after striking the foil (Figure 26.9).

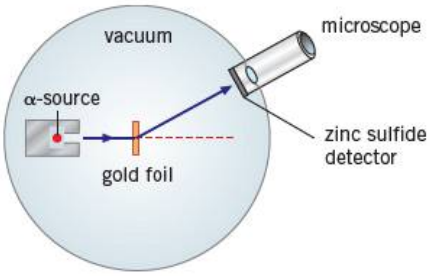


Figure 26.9 α-scattering experiment

They discovered that:

- the vast majority of the α-particles passed through the foil with very little or no deviation from their original path
- a small number of particles were deviated through an angle of more than about 10°
- an extremely small number of particles (one in ten thousand) were deflected through an angle greater than 90°.

From these observations, the following conclusions could be drawn.

- The majority of the mass of an atom is concentrated in a very small volume at the centre of the atom. Most α-particles would, therefore, pass through the foil undeviated.
- The centre (or nucleus) of an atom is charged. α-particles, which are also charged, passing close to the nucleus will experience a repulsive force causing them to deviate.
- Only α-particles that pass very close to the nucleus, almost striking it head-on, will experience large enough repulsive forces to cause them to deviate through angles greater than 90°. The fact that so few particles did so confirms that the nucleus is very small, and that most of the atom is empty space.

Figure 26.10 shows some of the possible trajectories of the α-particles. Using the nuclear model of the atom and equations to describe the force between charged particles, Rutherford calculated the fraction of α-particles that he would expect to be deviated through various angles. The calculations agreed with the results from the experiment. This confirmed the nuclear model of the atom. Rutherford calculated that the diameter of the nucleus is about 10^{-15} m, and the diameter of the whole atom about 10^{-10} m. Figure 26.11 shows the features of the nuclear model of a nitrogen atom.

Some years later, the α-particle scattering experiment was repeated using α-particles with higher energies. Some discrepancies between the experimental results and Rutherford's scattering formula were observed. These seemed to be occurring because the high-energy α-particles were passing very close to the nucleus, and were experiencing, not only the repulsive electrostatic force, but also a strong attractive force which appears to act over only a very short range. This became known as the **strong nuclear force**. This is the force that holds the nucleus together.

www.iranino.co

IRANINOVIN

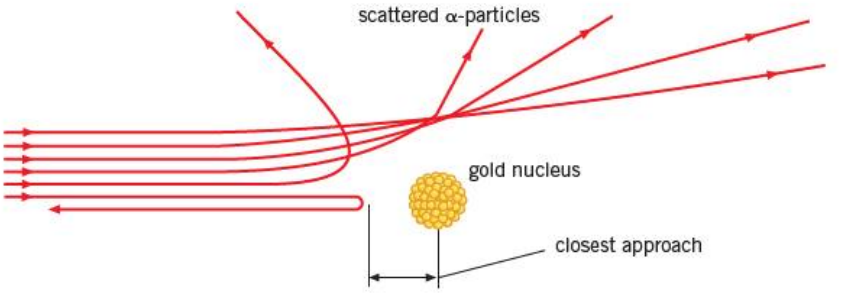


Figure 26.10

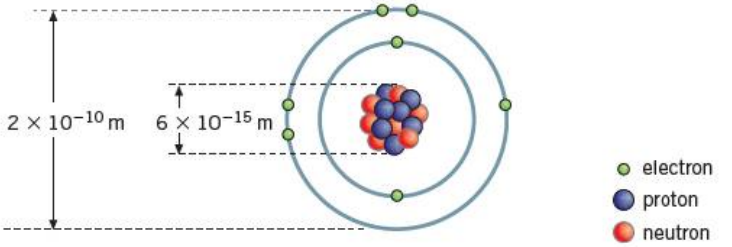


Figure 26.11 The diameter of a nitrogen atom is more than 30 000 times bigger than the diameter of its nucleus.

Probing matter using electrons

Electrons are not affected by the strong nuclear force. It was suggested that they might, therefore, be a more effective tool with which to investigate the structure of the atom. The A level course in Topic 25 shows that moving electrons have a wave-like property, and can be diffracted.

If a beam of electrons is directed at a sample of powdered crystal and the electron wavelength is comparable with the interatomic spacing in the crystal, the electron waves are scattered from planes of atoms in the tiny crystals, creating a diffraction pattern (Figure 26.12). The fact that a diffraction pattern is obtained confirms the regular arrangement of the atoms in a crystalline solid. Measurements of the angles at which strong scattering is obtained can be used to calculate the distances between planes of atoms.

If the energy of the electron beam is increased, the wavelength decreases. Eventually, the electron wavelength may be of the same order of magnitude as the diameter of the nucleus. Probing the nucleus with high-energy electrons, rather than α-particles, gives a further insight into the dimensions of the nucleus, and also gives information about the distribution of charge in the nucleus itself.

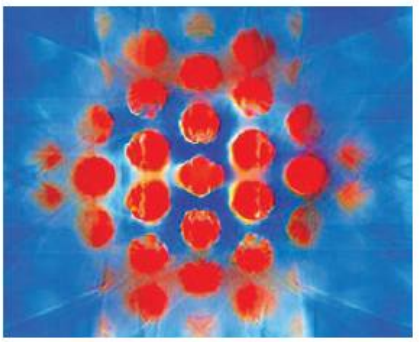


Figure 26.12 Electron diffraction pattern of a sample of pure titanium

26.2 Fundamental particles

In the nineteenth century, the atom was considered to be the fundamental particle from which all matter was composed. This idea was used to explain the basic structure of all elements. Experiments performed at the end of the nineteenth century and beginning of the twentieth century provided evidence for the structure of an atom. The conclusions were that all atoms have a nucleus containing protons which is surrounded by electrons and that the nucleus was very small compared with the size of the atom. The neutron was introduced to explain the discrepancy between the mass of the atom and the mass from the number of protons (number of positive charges). In 1932 Chadwick discovered the neutron and the fundamental particles were then considered to be the proton, the neutron and the electron. The structure of the atom was then considered to be similar to that shown in Figure 26.1.

www.iranino.co

IRANINOVIN

The particles in an atom must experience forces in order to maintain its structure. The forces were the gravitational force that acts between all masses (see Topic 8) and the electrostatic force that acts between charged objects (see Topic 17). The electrostatic force of repulsion is approximately 10^{36} times greater than the gravitational force of attraction between protons. Another attractive force must keep the protons together in the nucleus. This force is known as the **strong force** and acts between nucleons. The force does not seem to have any effect outside the nucleus and is, therefore, considered to be very short range (a little more than the diameter of nuclei, 10^{-14} m). There appears to be a limiting spacing between nucleons which is similar in different nuclei and this suggests that the force is repulsive as soon as the nucleons come close together. The strong force does not act on electrons.

The strong force acts on protons and neutrons but **not** on electrons.

Example

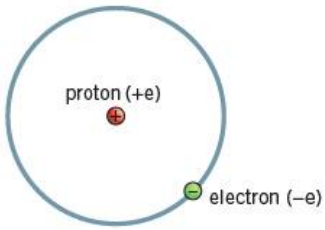


Figure 26.13 Hydrogen atom

Figure 26.13 illustrates a hydrogen atom with an electron orbiting the nucleus.

(a) State, for the forces acting on the electron and the proton,
- **(i)** their nature,
- **(ii)** their direction.

(b) Explain why a strong force does not act on the electron or proton.

(a)
- **(i)** gravitational force (due to the mass of the electron and proton), electrostatic force (due to the charge on the electron and proton)
- **(ii)** both forces are attractive and, therefore, directed from the one particle towards the other particle.

(b) The electron is not a nucleon and, hence, is not affected by the strong force. There is only one nucleon and the strong force acts between nucleons.

Now it's your turn

4 State the forces acting on the nucleons of a helium nucleus.

Hadrons and leptons

The discovery of antimatter in cosmic radiation supported the theory developed from the special theory of relativity and quantum theory that all fundamental particles have a corresponding antimatter particle. The matter and antimatter particles have the same mass but opposite charge. The following particles were required to support the theory: the antiproton, the antineutron and the antielectron. The symbols used for the antiparticle are $\bar{p}$ for the antiproton, $\bar{n}$ for the antineutron and $\bar{e}$ for the antielectron.

The antielectron or positive electron was introduced in β-particle decay on page 172. It is also known as the positron.

Many other particles were discovered in cosmic radiation throughout the twentieth century, giving support for the idea that the electron, proton and neutron were not the only fundamental particles.

The numerous types of particles are placed into two main categories. Those affected by the strong force are called **hadrons**, for example protons and neutrons, and those not affected by the strong force are called **leptons**, for example electrons and positrons.

The many different particles discovered in cosmic radiation have been reproduced in high-energy collisions using accelerators such as those at Stanford in California and CERN in Switzerland during the second half of the twentieth century. A vast number of collisions were carried out and a large number of hadrons were produced. Two of the conclusions to these reactions were:

- the total electrical charge remains constant
- the total number of nucleons normally remains constant.

www.iranino.co

IRANINOVIN

The quark model of hadrons

The problem of what were considered to be fundamental particles was resolved by the quark model for hadrons. In this model, the hadrons are made up of three smaller particles called **quarks**. The types of quark, called flavours of quark, are **up** (**u**), **down** (**d**) and **strange** (**s**). The quark flavours have charge and strangeness as shown in Table 26.4.

Table 26.4 Charge and strangeness values for the three quarks

flavour	charge	strangeness
up (u)	$+\frac{2}{3}$	0
down (d)	$-\frac{1}{3}$	0
strange (s)	$-\frac{1}{3}$	−1

There are three **antiquarks**, $\bar{u}$, $\bar{d}$ and $\bar{s}$: these have the opposite values of charge and strangeness.

Protons and neutrons consist of three quarks.

proton:		u	u	d	**neutron**:		u	d	d
charge	+1	$+\frac{2}{3}$	$+\frac{2}{3}$	$-\frac{1}{3}$	charge	0	$+\frac{2}{3}$	$-\frac{1}{3}$	$-\frac{1}{3}$
strangeness	0	0	0	0	strangeness	0	0	0	0

In strong interactions, the quark flavour is conserved.

Example

State the values of charge and strangeness for the antiquarks $\bar{u}$, and $\bar{d}$

$\bar{u}$ charge $-\frac{2}{3}$ strangeness 0

and $\bar{d}$ charge $+\frac{1}{3}$ strangeness 0

Now it's your turn

5 Show whether the following reaction can occur.

$$p + p \rightarrow p + \bar{p} + n$$

Leptons

Leptons are particles that are not affected by the strong force. The electron and neutrino and their antimatter partners, the positron and antineutrino, are examples of leptons. These types of particle do not appear to be composed of any smaller particles and are, therefore, considered to be fundamental particles.

The emission of electrons or positrons from nuclei was discussed earlier in this topic (β-decay page 172). During the decay of a neutron in the nucleus, a proton is formed and an electron and antineutrino emitted. In terms of the fundamental particles, quarks, the reaction can be shown as follows:

$${}^{1}_{0}n \rightarrow {}^{1}_{1}p + {}^{0}_{-1}e + {}^{0}_{0}\bar{\nu}$$

u u

d u

d d

The quark flavour is not conserved as a down quark has changed to an up quark. The reaction cannot be due to the strong force. The β-decay must be due to another force. This force is called the **weak force** or **weak interaction**.

www.iranino.co

IRANINOVIN

The total lepton number before a reaction is equal to the total lepton number after the reaction.

The lepton number is +1 for the particle and –1 for the antiparticle.

The total lepton number before the reaction is zero in the β⁻-decay above. The lepton numbers for the particles after the reaction are +1 for the electron and –1 for the antineutrino, giving a total of zero.

Example

Describe the reaction where a proton in the nucleus turns into a neutron and emits a β-particle in terms of the quark model.

$${}^{1}_{1}\mathrm{p} \rightarrow {}^{1}_{0}\mathrm{n} + {}^{0}_{+1}\mathrm{e} + {}^{0}_{0}\nu$$

u u

u d

d d

An up quark changes into a down quark.

Now it's your turn

6 What is the difference between a hadron and a lepton?

Summary

- An atom consists of a nucleus containing protons and neutrons surrounded by orbiting electrons.
- Most of the mass of an atom is contained in its nucleus.
- An atom is neutral as it contains an equal number of protons and electrons.
- Atoms which have gained or lost electrons are charged, and are called ions.
- The nucleon number A of a nucleus is the number of nucleons (protons and neutrons) in the nucleus.
- The proton number Z of a nucleus is the number of protons in the nucleus; hence the number of neutrons in the nucleus is $A - Z$.
- A nucleus (chemical symbol X) may be represented by: $^{\text{nucleon number}}_{\text{proton number}}\mathrm{X}$
- Isotopes are different forms of the same element, that is, nuclei with the same proton number but with different nucleon numbers.
- An α-particle is a helium nucleus (two protons and two neutrons).
- A β-particle is a fast-moving electron.
- γ-radiation consists of short-wavelength electromagnetic waves.
- In nuclear notation the emissions are represented as: α-particle ${}^{4}_{2}\mathrm{He}$; β⁻-particle ${}^{0}_{-1}\mathrm{e}$ or β⁺-particle ${}^{0}_{+1}\mathrm{e}$; γ-radiation ${}^{0}_{0}\gamma$.
- α-emission reduces the nucleon number of the parent nucleus by 4, and reduces the proton number by 2.
- β-emission causes no change to the nucleon number of the parent nucleus, and increases or decreases the proton number by 1.
- γ-emission causes no change to nucleon number or proton number of the parent nucleus.
- Radioactive decay is a spontaneous, random process.
- The Rutherford α-particle experiment confirmed the nuclear model of the atom: the atom consists of a small, positively-charged nucleus, surrounded by negatively-charged electrons in orbit about the nucleus and that the vast majority of the mass of the atom is in the nucleus.
- The diameter of the nucleus is about 10^{-15} m; the diameter of the atom is about 10^{-10} m.
- Electron diffraction gives evidence for the regular arrangement of atoms in crystals, and allows the measurement of the distance between planes of atoms in solids.

www.iranino.co

IRANINOVIN

- For every type of subatomic particle there is an antimatter particle which has the same mass but opposite electrical charge.
- The antiparticle of the electron is called the positron.
- Protons and neutrons are hadrons and are affected by the strong force.
- During hadron reactions, charge and strangeness are conserved.
- The simple quark model has three flavours of quark (up, down and strange) together with their antiquarks.
- Protons are composed of quarks up, up and down, and neutrons of quarks up, down and down.
- Electrons and neutrinos are leptons which are fundamental particles and are affected by a weak interaction.
- During β^- decay: ${}^{1}_{0}n \rightarrow {}^{1}_{1}p + {}^{0}_{-1}e$ (electron) + ${}^{0}_{0}\bar{\nu}$ (antineutrino)
- During β^+ decay: ${}^{1}_{1}p \rightarrow {}^{1}_{0}n + {}^{0}_{1}e$ (positron) + ${}^{0}_{0}\nu$ (neutrino)
- The energy of subatomic particles is often measured in eV or MeV.

Examination style questions

1 You are provided with a radioactive source which could be emitting α- or β-particles, or γ-radiation, or any combination of these. Describe a simple experiment, based on their relative penetrating qualities, you might carry out to determine the nature of the radiation(s) being emitted.

2 Explain the changes that take place to the nucleus of an atom when it emits:

a an α-particle,
b a β-particle,
c γ-radiation.

3 Complete the following radioactive series.

${}^{235}_{92}U \rightarrow {}^{?}_{?}Y + {}^{?}_{?}He$

${}^{?}_{?}Y \rightarrow {}^{?}_{?}Z + {}^{0}_{-1}e$

${}^{?}_{?}Z \rightarrow {}^{?}_{?}Z + {}^{?}_{?}?$

4 Calculate the speed of

a an electron with kinetic energy of 1.5 keV,
b an α-particle with kinetic energy of 1.5 keV.

5 A stationary radium nucleus (${}^{224}_{88}Ra$) of mass 224 u spontaneously emits an α-particle. The α-particle is emitted with an energy of 9.2×10^{-13} J, and the reaction gives rise to a nucleus of radon (Rn).

a Write down a nuclear equation to represent α-decay of the radium nucleus.
b Show that the speed with which the α-particle is ejected from the radium nucleus is 1.7×10^{7} m s^{-1}.
c Calculate the speed of the radon nucleus on emission of the α-particle. Explain how the principle of conservation of momentum is applied in your calculation.

6 When an α-particle travels through air, it loses energy by ionisation of air molecules. For every air molecule ionised, approximately 5.6×10^{-18} J of energy is lost by the α-particle.

a Suggest a typical value for the range of an α-particle in air. Hence estimate the number of air molecules ionised per millimetre of the path of the α-particle, given that the α-particle has initial energy 9.2×10^{-13} J.
b It has been discovered that the number of ionisations per unit length of the path of an α-particle suddenly increases just before the α-particle stops. State, with a reason, the effect that this observation will have on your estimate.

7 **a** The radioactive decay of some nuclei gives rise to the emission of α-particles.
State
i what is meant by an α-particle, *[1]*
ii two properties of α-particles. *[2]*

b One possible nuclear reaction involves the bombardment of a stationary nitrogen-14 nucleus by an α-particle to form oxygen-17 and another particle.
i Copy and complete the nuclear equation for this reaction. *[2]*

$${}^{14}_{7}N + {}^{\cdots\cdots}_{\cdots\cdots}\alpha \rightarrow {}^{17}_{8}O + \cdots\cdots\cdots\cdots$$

ii The total mass–energy of the nitrogen-14 nucleus and the α-particle is less than that of the particles resulting from the reaction. This mass–energy difference is 1.1 MeV.
1 Suggest how it is possible for mass–energy to be conserved in this reaction. *[1]*
2 Calculate the speed of an α-particle having kinetic energy of 1.1 MeV. *[4]*

Cambridge International AS and A level Physics, 9702/22 May/June 2010 Q 7

8 **a** Evidence for the nuclear atom was provided by the α-particle scattering experiment. State the results of this experiment. *[2]*
b Give estimates for the diameter of
i an atom, *[1]*
ii a nucleus. *[1]*

Cambridge International AS and A level Physics, 9702/02 Oct/Nov 2007 Q 7

www.iranino.co

IRANINOVIN

9 **a** Uranium (U) has at least fourteen isotopes. Explain what is meant by *isotopes*. *[2]*

b One possible nuclear reaction involving uranium is

$$^{235}_{92}U + {}^{1}_{0}n \longrightarrow {}^{141}_{56}Ba + {}^{92}_{Z}Kr + x\,{}^{1}_{0}n + \text{energy}.$$

i State three quantities that are conserved in a nuclear reaction. *[3]*

ii For this reaction, determine the value of

1 Z, *[1]*

2 x. *[1]*

Cambridge International AS and A level Physics, 9702/21 Oct/Nov 2010 Q 7

10 **a** β-radiation is emitted during the spontaneous radioactive decay of an unstable nucleus.

i State the nature of a β-particle. *[1]*

ii State two properties of β-radiation. *[2]*

iii Explain the meaning of *spontaneous radioactive decay*. *[1]*

b The following equation represents the decay of a nucleus of hydrogen-3 by the emission of a β-particle. Copy and complete the equation.

$$^{3}_{1}H \longrightarrow {}^{\cdots\cdots}_{\cdots\cdots}He + {}^{\cdots\cdots}_{\cdots\cdots}\beta \quad [2]$$

c The β-particle is emitted with an energy of 5.7×10^3 eV. Calculate the speed of the β-particle. *[3]*

d A different isotope of hydrogen is hydrogen-2 (deuterium). Describe the similarities and differences between the atoms of hydrogen-2 and hydrogen-3. *[2]*

Cambridge International AS and A level Physics, 9702/23 Oct/Nov 2012 Q 6

11 Uranium-236 ($^{236}_{92}U$) and uranium-237 ($^{237}_{92}U$) are both radioactive. Uranium-236 is an α-emitter and uranium-237 is a β-emitter.

a Distinguish between an α-particle and a β-particle. *[4]*

b The grid of Fig. 26.13 shows some proton numbers Z on the x-axis and the number N of neutrons in the nucleus on the y-axis.

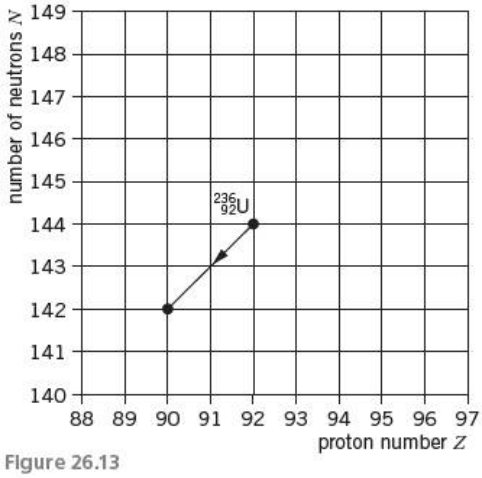


Figure 26.13

The α-decay of uranium-236 ($^{236}_{92}U$) is represented on the grid. This decay produces a nucleus of thorium (Th).

i Write down the nuclear equation for this α-decay. *[2]*

ii Copy Fig. 26.13, mark the position for a nucleus of

1 Uranium-237 (mark this position with the letter U),

2 Neptunium, the nucleus produced by the β-decay of uranium-237 (mark this position with the letters Np). *[2]*

Cambridge International AS and A level Physics, 9702/02 May/June 2008 Q 7

www.iranino.co

IRANINOVIN

7 Motion in a circle

By the end of this topic, you will be able to:

7.1
- **(a)** define the radian and express angular displacement in radians
- **(b)** understand and use the concept of angular speed to solve problems
- **(c)** recall and use $v = r\omega$ to solve problems

7.2
- **(a)** describe qualitatively motion in a curved path due to a perpendicular force, and understand the centripetal acceleration in the case of uniform motion in a circle
- **(b)** recall and use centripetal acceleration equations $a = r\omega^2$ and $a = v^2/r$
- **(c)** recall and use centripetal force equations $F = mr\omega^2$ and $F = mv^2/r$

Starting points

- Velocity is instantaneous speed in a given direction.
- An acceleration is change in velocity brought about by a resultant force.
- Newton's laws of motion.

7.1 Radian measure and angular displacement

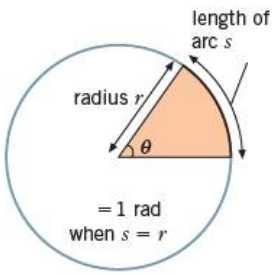


Figure 7.1 θ in radians = arc/radius

In circular motion, it is convenient to measure angles in **radians** rather than degrees. One degree is, by tradition, equal to the angle of a complete circle divided by 360.

One radian (rad) is defined as the angle subtended at the centre of a circle by an arc equal in length to the radius of the circle.

Thus, to obtain an angle in radians, we divide the length of the arc by the radius of the circle (see Figure 7.1).

$$\theta = \frac{\text{length of arc}}{\text{radius of circle}} \left(= \frac{s}{r}\right)$$

The angle in radians in a complete circle would be

$$\theta = \frac{\text{circumference of the circle}}{\text{radius}} = \frac{2\pi r}{r} = 2\pi$$

Since the angle of a complete circle is 360°, then

$$2\pi \text{ rad} = 360°$$

or

$$1 \text{ rad} = 57.3°$$

Angular speed

For an object moving in a circle:

The **angular speed** *is defined as the angle swept out by the radius per unit time.*

www.iranino.co

IRANINOVIN

The **angular velocity** is the angular speed in a given direction (for example, clockwise). The unit of angular speed and angular velocity is the radian per second (rad s^{-1}).

$$\text{angular speed } \omega = \frac{\Delta\theta}{\Delta t}$$

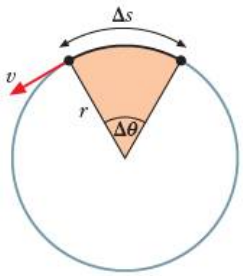


Figure 7.2 Angular velocity $\omega = v/r$

Figure 7.2 shows an object travelling at constant speed v in a circle of radius r.

In a time Δt the object moves along an arc of length Δs and sweeps out an angle $\Delta\theta$. From the definition of the radian,

$$\Delta\theta = \Delta s/r \text{ or } \Delta s = r\Delta\theta$$

Dividing both sides of this equation by Δt,

$$\Delta s/\Delta t = r\Delta\theta/\Delta t$$

By definition, for small angles, $\Delta s/\Delta t$ is the linear speed v of the object, and $\Delta\theta/\Delta t$ is the angular speed ω. Hence,

$$v = r\omega$$

Example

An aircraft in a display team makes a turn in a horizontal circle of radius 500 m. It is travelling at a speed of 100 m s^{-1}.

Calculate the angular speed of the aircraft.

From $v = r\omega$, the angular speed ω is 100/500 = **0.2 rad s^{-1}**

Now it's your turn

1 A car is travelling along a circular path with linear speed 15 m s^{-1} and angular speed 0.36 rad s^{-1}. What is the radius of curvature of the track?

7.2 Centripetal acceleration and centripetal force

Newton's first law of motion (see Topic 4) tells us that an object with a resultant force of zero acting on it will either not be moving at all, or it will be moving in a straight line at constant speed (that is, its velocity does not change). The object is said to be in equilibrium. (The full conditions of equilibrium require there to be no resultant force and no resultant moment acting on the body.)

An object travelling in a circle may have a constant speed, but it is not travelling in a straight line. The velocity is changing as velocity is a vector (has magnitude and direction) and its direction is changing. A change in velocity means the object is accelerating.

This acceleration is towards the centre of the circle. It is called the **centripetal acceleration**. In order to make an object accelerate, there must be a resultant force acting on it. This force is called the **centripetal force**. The centripetal force acts towards the centre of the circle, in the same direction as the acceleration. This means that the centripetal force always acts at right angles to the instantaneous velocity of the object.

Consider a ball on a string which is being swung in a horizontal circle. The tension in the string provides the centripetal force.

At any instant, the direction of the ball's velocity is along the tangent to the circle, as shown in Figure 7.3. If the string breaks or is released, there is no longer any tension in the string and hence no centripetal force. The ball will travel in the direction of the tangent to the circle at the moment of release.

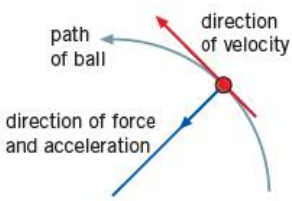


Figure 7.3 A ball swung in a circle on the end of a string

www.iranino.co

IRANINOVIN

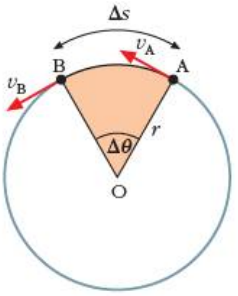


Figure 7.4 Diagram for proof of $a = v^2/r$

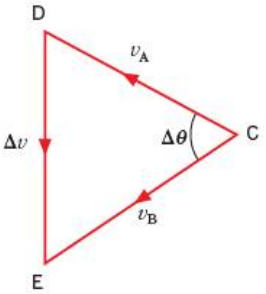


Figure 7.5 Vector diagram for proof of $a = v^2/r$

Figure 7.4 shows an object which has travelled at constant speed v in a circular path from A to B in time Δt. At A, its velocity is v_A, and at B the velocity is v_B. Both v_A and v_B are vectors.

The change in velocity Δv may be seen in the vector diagram of Figure 7.5. A vector Δv must be added to v_A in order to give the new velocity v_B.

The angle between the two radii OA and OB is $\Delta\theta$. This angle is also equal to the angle between the vectors v_A and v_B, because triangles OAB and CDE are similar. Consider angle $\Delta\theta$ to be so small that the arc AB may be approximated to a straight line. Then, using similar triangles, DE/CD = AB/OA, and $\Delta v/v_A = \Delta s/r$ or

$$\Delta v = \Delta s(v_A/r)$$

The time to travel either the distance Δs or the angle $\Delta\theta$ is Δt. Dividing both sides of the equation by Δt,

$$\Delta v/\Delta t = (\Delta s/\Delta t)(v_A/r).$$

and from the definitions of acceleration ($a = \Delta v/\Delta t$) and speed ($v = \Delta s/\Delta t = v_A = v_B$) we have $a = v(v/r)$ or $a = v^2/r$.

This expression can be written in terms of angular speed ω. Since $v = r\omega$,

$$\text{centripetal acceleration} = \frac{v^2}{r} = r\omega^2$$

Now, force F is related to acceleration a by the expression $F = ma$, where m is the mass.

$$\text{centripetal force} = \frac{mv^2}{r} = mr\omega^2$$

Example

The drum of a spin dryer has a radius of 20 cm and rotates at 600 revolutions per minute.

(a) Show that the angular speed of the drum is about 63 rad s^{-1}.

(b) Calculate, for a point on the edge of the drum:

(i) its linear speed,

(ii) its acceleration towards the centre of the drum.

(a) 600 revolutions per minute is 10 revolutions per second. The time for one revolution is thus 0.10 s. Each revolution is 2π rad, so the angular speed $\omega = \theta/t = 2\pi/0.10 = \mathbf{63\ rad\ s^{-1}}$.

(b) (i) Using $v = r\omega$, $v = 0.20 \times 63 = \mathbf{13\ m\ s^{-1}}$ ($\mathbf{12.6\ m\ s^{-1}}$).

(ii) Using $a = v^2/r$, $a = (12.6)^2/0.20 = \mathbf{800\ m\ s^{-2}}$.

Now it's your turn

2 A toy train moves round a circular track of diameter 0.70 m, completing one revolution in 10 seconds. Calculate, for this train:

(a) the linear speed,

(b) the angular speed,

(c) the centripetal acceleration.

3 A stone attached to a string is moving in a horizontal circle of radius 90 cm. The stone has mass 65 g and completes one revolution in 0.70 s. Calculate the tension in the string, keeping the stone in its circular path.

Examples of circular motion

When a ball is whirled round on the end of a string, you can see clearly that the tension in the string is making the ball accelerate towards the centre of the circle. However, in other examples it is not always so easy to see what force is providing the centripetal acceleration.

A satellite in Earth orbit experiences gravitational attraction towards the centre of the Earth. This attractive force provides the centripetal force and causes the satellite to

www.iranino.co

IRANINOVIN

accelerate towards the centre of the Earth, and so it moves in a circle. We shall return to this in detail in Topic 8.

A charged particle moving at right angles to a magnetic field experiences a force at right angles to its direction of motion, and therefore moves in the arc of a circle. This will be considered in more detail in Topic 22.

For a car travelling in a curved path, the frictional force between the tyres and the road surface provides the centripetal force. If this frictional force is not large enough, for example if the road is oily or slippery, then the car carries on moving in a straight line – it skids.

A passenger in a car that is cornering appears to be flung away from the centre of the circle. The centripetal force required to maintain the passenger in circular motion is provided through the seat of the car. This force is below the centre of mass, M, of the passenger, causing rotation about the centre of mass (Figure 7.6). The effect is that the upper part of the passenger moves outwards unless another force acts on the upper part of the body, preventing rotation.

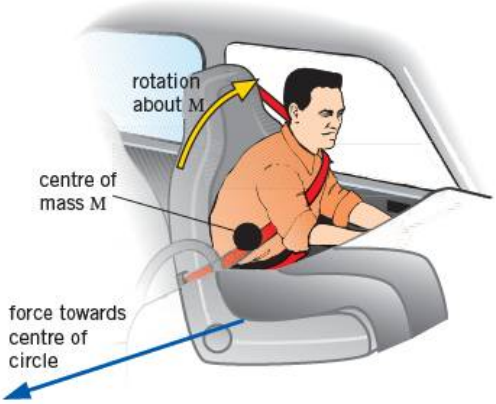


Figure 7.6 Passenger in a car rounding a corner

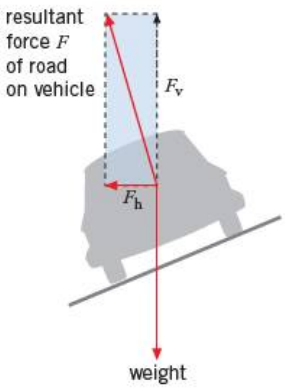


Figure 7.7 Cornering on a banked track

For cornering which does not rely only on friction, the road can be banked (Figure 7.7). The road provides a resultant force normal to its surface through contact between the tyres and the road. This resultant force F is at an angle to the vertical, and can be resolved into a vertical component F_v and a horizontal component F_h, as shown in Figure 7.7. F_v is equal to the weight of the vehicle, thus maintaining equilibrium in the vertical direction. The horizontal component F_h provides the centripetal force towards the centre of the circle. Many roads are banked for greater road safety, so as to reduce the chance of loss of control of vehicles due to skidding outwards on the corner, and for greater passenger comfort.

An aircraft has a lift force caused by the different rates of flow of air above and below the wings. The lift force balances the weight of the aircraft when it flies on a straight, level path (Figure 7.8a). In order to change direction, the aircraft is banked so that the wings are at an angle to the horizontal (Figure 7.8b). The lift force now has a horizontal component which provides a centripetal force to change the aircraft's direction.

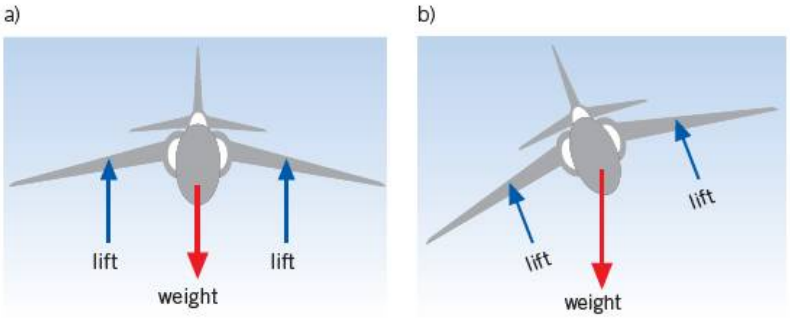


Figure 7.8 An aircraft a) in straight, level flight and b) banking

www.iranino.co

IRANINOVIN

A centrifuge (Figure 7.9) is a device that is used to spin objects at high speed about an axis. It is used to separate particles in mixtures. More massive particles require larger centripetal forces in order to maintain circular motion than do less massive ones. As a result, the more massive particles tend to separate from less massive particles, collecting further away from the axis of rotation. Space research centres, such as NASA, use centrifuges which are large enough to rotate a person (Figure 7.10). Their purpose is to investigate the effects of large accelerations on the human body.



Figure 7.9 Separation of a solid from a liquid in a laboratory centrifuge



Figure 7.10 Centrifuge testing the effect of acceleration on the human body

Motion in a vertical circle

Some theme park rides involve rotation in a vertical circle. A person on such a ride must have a resultant force acting towards the centre of the circle.

The forces acting on the person are the person's weight, which always acts vertically downwards, and the reaction from the seat, which acts at right angles to the seat.



Figure 7.11 Ferris wheel at a theme park

www.iranino.co

IRANINOVIN

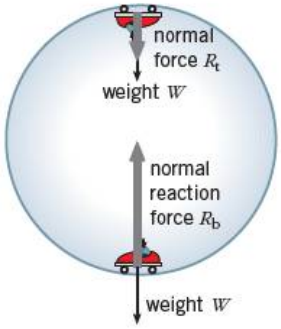


Figure 7.12 Forces on a person on a circular ride

Consider a person moving round a vertical circle at speed v.

At the bottom of the ride, the reaction R_b from the seat must provide the centripetal force, as well as overcoming the weight W of the person. Figure 7.12 illustrates the situation.

The centripetal force is given by

$$mv^2/r = R_b - W$$

At the top of the ride, the weight W and the reaction force R_t both act downwards towards the centre of the circle. The centripetal force is now given by

$$mv^2/r = R_t + W$$

This means that the force R_t from the seat at the top of the ride is less than the force R_b at the bottom. If the speed v is not large, then at the top of the circle the weight may be greater than the centripetal force. The person would lose contact with the seat and fall inwards.

Example

A rope is tied to a bucket of water, and the bucket is swung in a vertical circle of radius 1.2 m. What must be the minimum speed of the bucket at the highest point of the circle if the water is to stay in the bucket throughout the motion?

This example is similar to the problem of the theme park ride. Water will fall out of the bucket if its weight is greater than the centripetal force. The critical speed v is given by $mv^2/r = mg$ or $v^2 = gr$.

Here, $v = \sqrt{(9.8 \times 1.2)}$ $= \mathbf{3.4\,m\,s^{-1}}$.

Now it's your turn

4 At an air show, an aircraft diving at a speed of 170 m s^{-1} pulls out of the dive by moving in the arc of a circle at the bottom of the dive.
   (a) Calculate the minimum radius of this circle if the centripetal acceleration of the aircraft is not to exceed five times the acceleration of free fall.
   (b) The pilot has mass 85 kg. What is the resultant force upwards on him at the instant when the aircraft is at its lowest point?

Summary

- Angles may be measured in radians (rad). One radian is the angle subtended at the centre of a circle by an arc of the circle equal in length to its radius.
- Angular speed ω is the angle swept out per unit time by a line rotating about a point.
- A particle moving along a circle of radius r with linear speed v has angular velocity ω given by $v = r\omega$.
- A resultant force acting towards the centre of the circle, called the centripetal force, is required to make an object move in a circle.
- An object moving along a circle of radius r with linear speed v and angular speed ω has an acceleration a towards the centre (the centripetal acceleration) given by $a = v^2/r = r\omega^2$.
- For an object of mass m moving along a circle of radius r with linear speed v and angular speed ω, the centripetal force F is given by $F = mv^2/r = mr\omega^2$.

www.iranino.co

IRANINOVIN

Examination style questions

1 State how the centripetal force is provided in the following examples:

- **a** a planet orbiting the Sun,
- **b** a child on a playground roundabout,
- **c** a train on a curved track,
- **d** a passenger in a car going round a corner.

2 NASA's 20-G centrifuge is used for testing space equipment and the effect of acceleration on humans. The centrifuge consists of an arm of length 17.8 m, rotating at constant speed and producing an acceleration equal to 20 times the acceleration of free fall. Calculate

- **a** the angular speed required to produce a centripetal acceleration of 20 g,
- **b** the rate of rotation of the arm (g = 9.8 m s^{-2}).

3 A satellite orbits the Earth 200 km above its surface. The acceleration towards the centre of the Earth is 9.2 m s^{-2}. The radius of the Earth is 6400 km. Calculate:

- **a** the speed of the satellite,
- **b** the time to complete one orbit.

4 A light rigid rod is pivoted at one end C so that the rod rotates in a vertical circle at constant speed as shown in Fig. 7.13.

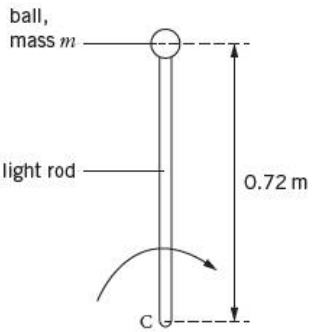


Fig. 7.13

A small ball of mass m is fixed to the free end of the rod so that the ball moves in a vertical circle. When the ball is vertically above point C, the tension T in the rod is given by

$$T = 2mg$$

where g is the acceleration of free fall.

a **i** Explain why the centripetal force on the ball is **not** equal to $2mg$. *[2]*

ii State, in terms of mg,

1 the magnitude of the centripetal force, *[1]*

2 the tension in the rod when the ball is vertically below point C. *[1]*

b The distance from point C to the centre of the ball is 0.72 m.
Determine, for the ball *[3]*

i the angular speed,

ii the linear speed.

Cambridge International AS and A Level Physics 9646/03 Oct/Nov 2010 Q 2

www.iranino.co

IRANINOVIN

8 Gravitational fields

By the end of this topic, you will be able to:

8.1 (a) understand the concept of a gravitational field as an example of a field of force and define gravitational field strength as force per unit mass

8.2 (a) understand that, for a point outside a uniform sphere, the mass of the sphere may be considered to be a point mass at its centre

(b) recall and use Newton's law of gravitation in the form $F = Gm_1m_2/r^2$

(c) analyse circular orbits in inverse square law fields, including geostationary orbits, by relating the gravitational force to the centripetal acceleration it causes

8.3 (a) derive, from Newton's law of gravitation and the definition of gravitational field strength, the equation $g = GM/r^2$ for the gravitational field strength of a point mass

(b) recall and solve problems using the equation $g = GM/r^2$ for the gravitational field strength of a point mass

(c) show an appreciation that on the surface of the Earth, g is approximately constant

8.4 (a) define potential at a point as the work done per unit mass in bringing a small test mass from infinity to the point

(b) solve problems using the equation $\phi = -GM/r$ for the potential in the field of a point mass

Starting points

- There is a force of attraction between all masses. On Earth, the force attracting objects to the Earth is referred to as 'force due to gravity' and is called weight.
- At the surface of the Earth, all objects have the same acceleration when in a vacuum.

8.1 Gravitational field

We are familiar with the fact that the Earth's force of gravity is responsible for our weight, the force which pulls us towards the Earth. Isaac Newton concluded that the Earth's force of gravity is also responsible for keeping the Moon in orbit.

We now know that every mass attracts every other mass. This force is known as the force due to gravity – the gravitational force. The region around a mass where this gravitational force is felt is known as a **gravitational field**.

A gravitational field is a region of space where a mass experiences a force.

Force between masses

Isaac Newton showed that the Earth's force of gravity extends into space, but weakens with distance according to an inverse square law. That is, the Earth's force of gravity varies inversely with the square of the distance from the centre of the Earth. If you go one Earth radius above the Earth's surface, the force is a quarter of the force on the Earth's surface. This is part of **Newton's law of gravitation**.

www.iranino.co

IRANINOVIN

Newton's law of gravitation states that two point masses attract each other with a force that is proportional to the product of their masses and inversely proportional to the square of their separation.

Hence, if F is the force of attraction between two bodies of mass m_1 and m_2 respectively with distance r between their centres, then

$$F \propto m_1 m_2 / r^2$$

or

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

where the constant of proportionality G is called the **gravitational constant**.

The value of G is 6.67×10^{-11} N m^2 kg^{-2}.

Note: Newton's law specifies that the two masses are point masses. However, the law still holds where the diameter/size of the masses is small compared to their separation. For example, the Sun and the Earth are not point masses. Newton's law still applies because their separation (1.5×10^8 km) is very large in comparison to their diameters (Earth, 1.3×10^4 km; Sun, 1.4×10^6 km).

Notice that this equation has a similar form to that for Coulomb's law between two charges (page 275). Both Newton's law and Coulomb's law are inverse square laws of force. We say that the equations are analogous, or that there is an analogy between this aspect of electric fields and gravitational fields. However, there are important differences.

- The electric force acts on charges, whereas the gravitational force acts on masses.
- The electric force can be attractive or repulsive, depending on the signs of the interacting charges, whereas two masses always attract each other.

It is possible to measure the gravitational constant G in a school laboratory, but the force of gravity between laboratory-sized masses is so small that it is not easy to obtain a reliable result.

8.2 & 8.3 Gravitational field strength

Any mass in a gravitational field experiences a force. The magnitude of this force depends on the gravitational field strength.

The **gravitational field strength** *at a point is defined as the force per unit mass acting on a small mass placed at that point.*

This continues the analogy between electric and gravitational fields. Remember that electric field strength is defined as the force per unit positive charge.

The force per unit mass is also a measure of the acceleration of free fall. By Newton's second law of motion $F = ma$ (see page 56). Thus, the gravitational field strength is given by $F/m = g$, where F is in newtons and m is in kilograms. This means that the gravitational field strength at the Earth's surface is about 9.8 N kg^{-1}. The unit N kg^{-1} is equivalent to the unit of acceleration, m s^{-2}. (Remember that we had two equivalent units for electric field strength, N C^{-1} and V m^{-1}. Although there is a clear *analogy* between N kg^{-1} and N C^{-1}, there is no *direct link* between m s^{-2} and V m^{-1}.

The gravitational field outside a spherical uniform mass is radial, as shown in Figure 8.1. This means that, from outside the sphere, all the lines of gravitational force appear to radiate from the centre of the sphere.

www.iranino.co

IRANINOVIN

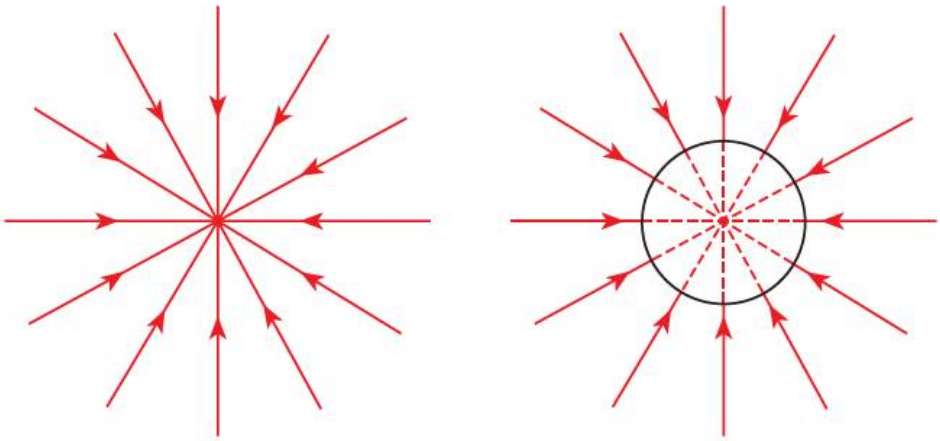


Figure 8.1 Gravitational fields

For a point outside a sphere whose mass is uniformly distributed, the sphere behaves as a point mass with its mass concentrated at its centre.

Example

The masses of the Sun and the Earth are 2.0×10^{30} kg and 6.0×10^{24} kg respectively. The separation of their centres is 1.5×10^{8} km.

Calculate the force of attraction between the Sun and the Earth.

The separation is large in comparison to their radii so, using Newton's law,

$F = Gm_1m_2/r^2$

$F = (6.67 \times 10^{-11} \times 2.0 \times 10^{30} \times 6.0 \times 10^{24})/(1.5 \times 10^{8} \times 10^{3})^2$

$= \mathbf{3.6 \times 10^{22}\,N}$

Now it's your turn

1 The mass of the planet Pluto is 1.2×10^{22} kg. Calculate the force of attraction between Pluto and the Sun when their separation is 5.9×10^{9} km.

Circular orbits

Most planets in the solar system have orbits which are nearly circular. We now bring together the idea of a gravitational force and that of a centripetal force (see page 186) to derive a relation between the period and the radius of the orbit of a planet describing a circular path about the Sun, or a satellite moving round the Earth or another planet.

Consider a planet of mass m in circular orbit about the Sun, of mass M, as shown in Figure 8.2.

If the radius of the orbit is r, the gravitational force F_{grav} between the Sun and the planet is, by Newton's law of gravitation,

$$F_{grav} = GMm/r^2.$$

It is this force that provides the centripetal force as the planet moves in its orbit. Note that the planet is changing direction continuously and is, therefore, not in equilibrium. The gravitational force provides the accelerating force – the centripetal force.

The centripetal force F_{circ} is given by

$$F_{circ} = mv^2/r$$

where v is the linear speed of the planet. As has just been stated,

$$F_{grav} = F_{circ}$$

www.iranino.co

IRANINOVIN

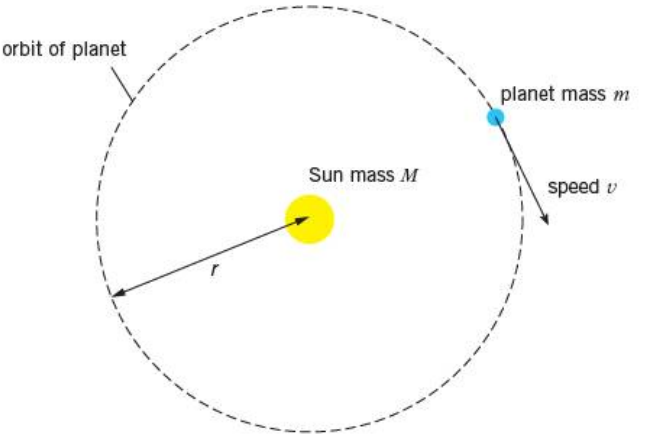


Figure 8.2 Circular orbit of a planet about the Sun

Thus

$$GMm/r^2 = mv^2/r$$

The period T of the planet in its orbit is the time required for the planet to travel a distance $2\pi r$. It is moving at speed v, so

$$v = 2\pi r/T$$

Putting this into the equation above, we have

$$GMm/r^2 = m(4\pi^2r^2/T^2)/r$$

or, simplifying,

$$T^2 = (4\pi^2/GM)r^3$$

Another way of writing this is

$$\frac{T^2}{r^3} = \frac{4\pi^2}{GM}$$

Look at the right-hand side of this equation. The quantities π and G are constants. If we are considering the relation between T and r for planets in the Solar System, then M is the same for each planet because it is the mass of the Sun.

This equation shows that:

For planets or satellites describing circular orbits about the same central body, the square of the period is proportional to the cube of the radius of the orbit.

This relation is known as **Kepler's third law of planetary motion**. Johannes Kepler (1571–1630) analysed data collected by Tycho Brahe (1546–1601) on planetary observations. He showed that the observations fitted a law of the form $T^2 \propto r^3$. Fifty years later Newton showed that an inverse square law of gravitation, the idea of centripetal acceleration, and the second law of motion, gave an expression of exactly the same form. Newton cited Kepler's law in support of his law of gravitation. In fact, the orbits of the planets are not circular, but elliptical, a fact recognised by Kepler. The derivation is simpler for the case of a circular orbit.

Table 8.1 gives information about T and r for planets of the solar system. The last column shows that the value of T^2/r^3 is indeed a constant. Moreover, the value of T^2/r^3 agrees very well with the value of $4\pi^2/GM$ which is 2.97×10^{-25} yr^2 m^{-3}.

www.iranino.co

IRANINOVIN

Table 8.1

planet	T/(Earth years)	r/km	T^2/r^3(yr^2 m^{-3})
Mercury	0.241	57.9×10^6	2.99×10^{-25}
Venus	0.615	108.0×10^6	3.00×10^{-25}
Earth	1.00	150.0×10^6	2.96×10^{-25}
Mars	1.88	228.0×10^6	2.98×10^{-25}
Jupiter	11.9	778.0×10^6	3.01×10^{-25}
Saturn	29.5	1.43×10^9	2.98×10^{-25}
Uranus	84.0	2.87×10^9	2.98×10^{-25}
Neptune	165	4.50×10^9	2.99×10^{-25}
Pluto*	248	5.90×10^9	2.99×10^{-25}
			(average 2.99×10^{-25})

*Since 2006, Pluto has been classified as a 'dwarf planet'.

Satellites are widely used in telecommunication. Many communication satellites are placed in what is called **geostationary orbit**. That is, they are in equatorial orbits with exactly the same period of rotation as the Earth (24 hours), and move in the same direction as the Earth (west to east) so that they are always above the same point on the Equator. Such satellites are called **geostationary satellites**. Details of the orbit of such a satellite are worked out in the example which follows.



Figure 8.3 Communication satellite

Example

For a geostationary satellite, calculate:

(a) the height above the Earth's surface,

(b) the speed in orbit.

(radius of Earth = 6.38×10^6 m; mass of Earth = 5.98×10^{24} kg)

(a) The period of the satellite is 24 hours = 8.64×10^4 s.
Using $T^2/r^3 = 4\pi^2/GM$, we have
$r^3 = 6.67 \times 10^{-11} \times 5.98 \times 10^{24} \times (8.64 \times 10^4)^2/4\pi^2$, giving
$r^3 = 7.54 \times 10^{22}$ m^3. Taking the cube root, the radius r of the orbit is 4.23×10^7 m.
The distance above the Earth's surface is $(4.23 \times 10^7 - 6.38 \times 10^6) =$ **3.59×10^7 m**.

(b) Since $v = 2\pi r/T$, the speed is given by
$v = 2\pi \times 4.23 \times 10^7/8.64 \times 10^4 =$ **3070 m s^{-1}**.

Now it's your turn

2 The radius of the Moon's orbit is 3.84×10^8 m, and its period is 27.4 days. Use Kepler's law to calculate the period of the orbit of a satellite orbiting the Earth just above the Earth's surface (radius of Earth = 6.38×10^6 m).

www.iranino.co

IRANINOVIN

Weightlessness

Suppose that you are carrying out an experiment involving the use of a newton balance in a lift. When the lift is stationary, an object of mass 10 kg, suspended from the balance, will give a weight reading of $10g$ N. If the lift accelerates upwards with acceleration $0.1g$, the reading on the balance increases to $11g$ N. If the lift accelerates downwards with acceleration $0.1g$, the apparent weight of the object decreases to $9g$ N. If, by an unfortunate accident, the lift cable were to break and there were no safety restraints, the lift would accelerate downwards with acceleration g. The reading on the newton balance would be zero. If, during the fall, you were to drop the pencil with which you are recording the balance readings, it would not fall to the floor of the lift but would remain stationary with respect to you. Both you and the pencil are in free fall. You are experiencing **weightlessness**.

Figure 8.4 illustrates your predicament. It might be more correct to refer to this situation as *apparent* weightlessness, as you can only be truly weightless in the absence of a gravitational field – that is, at an infinite distance from the Earth or any other attracting object.

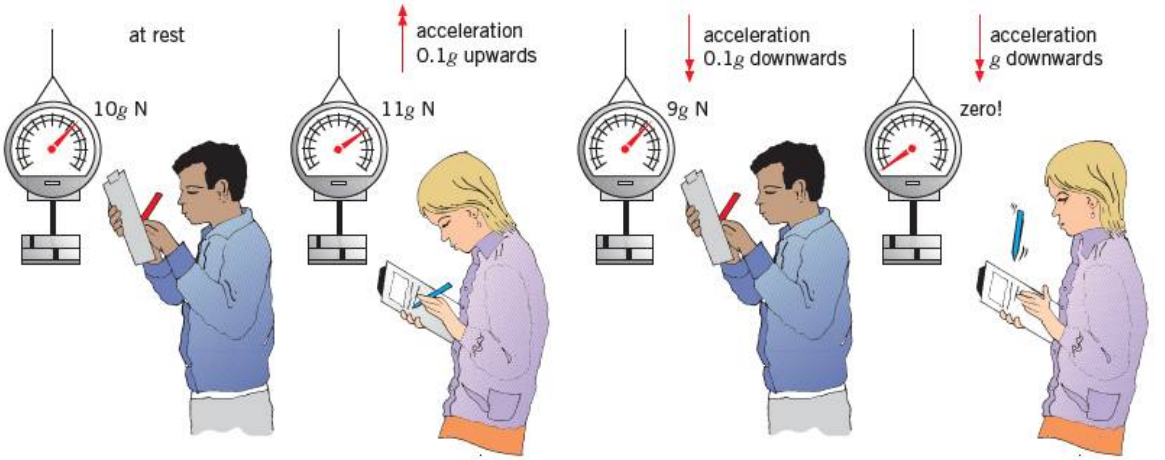


Figure 8.4 Lift experiment on weightlessness

A similar situation arises in a satellite orbiting the Earth. The force of gravity, which provides the centripetal force, is causing the body to fall out of its expected straight-line path. People and objects inside the satellite are experiencing a free fall situation and apparent weightlessness.

Gravitational field of a point mass

We have already met the idea of a point mass in talking about the motion of a particle (see page 40). Of course, all the masses we come across in the laboratory have a finite size. But for calculations involving gravitational forces, it is fortunate that a spherical mass behaves as if it were a point mass at the centre of the sphere, with all the mass of the sphere concentrated at that point. This is very similar to the idea that the charge on a conducting sphere can be considered to be concentrated at the centre of the sphere.

From Newton's law, the attractive force on a mass m caused by another mass M, with a distance of r between their centres, is given by

$$F = \frac{GMm}{r^2}$$

This means that the force per unit mass or gravitational field strength g is given by

$$g = \frac{F}{m} = \frac{GM}{r^2}$$

www.iranino.co

IRANINOVIN

The field strengths due to masses that you find in a laboratory are tiny. For example, the field strength one metre away from an isolated mass of one kilogram is only 7×10^{-11} N kg^{-1}. However, field strengths due to the masses of objects such as the Earth or Moon are much larger. We already know that the field strength due to the Earth at the surface of the Earth is about 10 N kg^{-1}. We can use this information to deduce information about the Earth, for example the mass of the Earth. Look at the example that follows.

Example

The radius of the Earth is 6.4×10^6 m and the gravitational field strength at its surface is 9.8 N kg^{-1}.

(a) Assuming that the field is radial, calculate the mass of the Earth.

(b) The radius of the Moon's orbit about the Earth is 3.8×10^8 m. Calculate the strength of the Earth's gravitational field at this distance.

(c) The mass of the Moon is 7.4×10^{22} kg. Calculate the gravitational attraction between the Earth and the Moon. (gravitational constant $G = 6.7 \times 10^{-11}$ N m^2 kg^{-2})

(a) Using $g = GM/r^2$, we have
$M = gr^2/G = 9.8 \times (6.4 \times 10^6)^2/6.7 \times 10^{-11} = \mathbf{6.0 \times 10^{24}\ kg}$

(b) Using $g = GM/r^2$, we have
$g = 6.7 \times 10^{-11} \times 6.0 \times 10^{24}/(3.8 \times 10^8)^2 = \mathbf{2.8 \times 10^{-3}\ N\,kg^{-1}}$

(c) Using $F = GMm/r^2$, we have
$F = 6.7 \times 10^{-11} \times 6.0 \times 10^{24} \times 7.4 \times 10^{22}/(3.8 \times 10^8)^2 = \mathbf{2.1 \times 10^{20}\ N}$

Now it's your turn

3 The mass of Jupiter is 1.9×10^{27} kg and its radius is 7.1×10^7 m. Calculate the gravitational field strength at the surface of Jupiter.
(gravitational constant $G = 6.7 \times 10^{-11}$ N m^2 kg^{-2})

Acceleration of free fall

For a body falling freely near the Earth's surface, the acceleration is given by *force/mass* where the force is the gravitational attraction.

We have already seen that gravitational field strength is defined as force per unit mass. Thus,

gravitational field strength g is equal to the acceleration of free fall.

For a uniform sphere,

$$g = \frac{F}{m} = \frac{GM}{r^2}$$

At the surface of Earth, we can assume that the Earth is approximately homogeneous and, therefore, the equation applies.

The radius r of Earth is approximately 6.4×10^3 km. If we move a few kilometres h above the Earth's surface, then the acceleration of free fall becomes

$$g_h = GM/(r + h)^2$$

Now, h is much smaller than r and so

$$r^2 \approx (r + h)^2$$

and

$$g \approx g_h$$

For small distances above the Earth's surface, g is approximately constant and is called the acceleration of free fall.

www.iranino.co

IRANINOVIN

Mass and weight

In Topic 4, mass was said to be a measure of the inertia of a body to changes in velocity. Unless the body is travelling at speeds close to that of light, its mass is constant.

In a gravitational field, by definition, there is a force acting on the mass equal to the product of mass and gravitational field strength. This force is called the **weight**.

For an object of mass m in a gravitational field of strength g, the weight W is given by

$$weight = mass \times gravitational\ field\ strength$$

or

$$W = mg$$

Although mass is invariant, weight depends on gravitational field strength. For example, a person of mass 60 kg has a weight of 600 N on Earth, but only 100 N on the Moon, although the mass is still 60 kg.

8.4 Gravitational potential and gravitational potential energy

In Topic 17, we will meet ideas about electric potential energy and electric potential. There is a very strong analogy between gravitational and electric fields, and this will help when we talk about electrical potential energy and electrical potential.

Gravitational potential *at a point in a gravitational field is defined as the work done per unit mass in bringing a small test mass from infinity to the point.*

The symbol for gravitational potential is ϕ and its unit is the joule per kilogram (J kg^{-1}). For a field produced by a point mass, the equation for the potential at a point in the field is

$$\phi = -\frac{GM}{r}$$

Note the minus sign.

The gravitational potential at infinity is defined as being zero. The gravitational force is always attractive and so, as the test mass moves from infinity, work can be done by the test mass and as a result its potential decreases. The gravitational potential is negative.

This is entirely consistent with the electrical case. We have a negative electric potential when the field is produced by a negative charge, so that the force between the negative field-producing charge and the positive test charge is attractive. Here the attractive gravitational force between field-producing mass and test mass also gives a negative potential.

Gravitational potential is work done per unit mass. For a body of mass m, then the gravitational potential energy of the body will be m times as large as for a body of unit mass.

$$Gravitational\ potential\ energy = mass \times gravitational\ potential$$

$$Gravitational\ potential\ energy = m\phi = -\frac{GMm}{r}$$

www.iranino.co

IRANINOVIN

Example

How much work is done by the gravitational field in moving a mass m of 2.0kg from infinity to a point A, 0.40m from a mass M of 30kg?

The work which would have to be done by an external force is simply the change in gravitational potential energy. The potential energy at infinity is zero, so

$W = m\phi_A = -mGM/r = -2.0 \times 6.7 \times 10^{-11} \times 30/0.40 = -1.0 \times 10^{-8}$J

This is the work which would be done by an external force and is negative, so the work done by the field is positive and is equal to **1.0×10^{-8}J**.

Note the similarity in the method of calculation with the electric potential energy calculation on page 277.

Now it's your turn

4 The Earth has mass 6.0×10^{24}kg and radius 6.4×10^{6}m. Find the change in gravitational potential energy of a meteorite of mass 150kg as it moves from an infinite distance to the Earth's surface. The meteorite starts from rest. With what speed does it strike the Earth? (gravitational constant $G = 6.67 \times 10^{-11}$ N m^2 kg^{-2})

Summary

- The attractive force between two point masses is proportional to the product of the masses and inversely proportional to the square of the distance between them. This is Newton's law of gravitation: $F = Gm_1m_2/r^2$
- A gravitational field is a region round a mass where another mass feels a force. Gravitational field strength g is the force per unit mass: $g = F/m$
  g is also the acceleration of free fall.
- The gravitational field strength at a point in the gravitational field of a point mass M is $g = GM/r^2$
- The gravitational potential at a point in a gravitational field is the work done per unit mass in bringing a small test mass from infinity to the point.
- The potential at a point in a field produced by a point mass is: $\phi = -GM/r$
- Gravitational potential energy at a point in a field produced by a point mass is: $E_p = \phi m = -GMm/r$
- For a planet or satellite in circular orbit about a body, the square of the period is proportional to the cube of the radius of the orbit. This is Kepler's third law of planetary motion.

Examination style questions

1 The Earth's radius is about 6.4×10^{6}m. The acceleration of free fall at the Earth's surface is 9.8m s^{-2}. The gravitational constant G is 6.7×10^{-11} N m^2 kg^{-2}. Use this information to estimate the mean density of the Earth.

2 The times for Mars and Jupiter to orbit the Sun are 687 days and 4330 days respectively. The radius of the orbit of Mars is 228×10^{6}km. Calculate the radius of the orbit of Jupiter.

3 The weight of a passenger in an aircraft on the runway is W. His weight when the aircraft is flying at an altitude of 16km above the Earth's surface is W_a. The percentage change F in his weight is given by $F = 100(W_a - W)/W\%$. Taking the radius of the Earth as 6.378×10^{3}km, calculate F. Calculate also the percentage change P in his gravitational potential energy.

4 Fig. 8.5 illustrates the apparatus used by Cavendish in 1798 to find a value for the gravitational constant G. In a school experiment using similar apparatus, two lead spheres are attached to a light horizontal beam which is suspended by a wire. When a flask of mercury is brought close to each sphere, the gravitational attraction causes the beam to twist through a small angle. From measurements of the twisting (torsional) oscillations of the beam, a value can be found for the force producing a measured deflection. G can then be calculated if the large and small masses are known.

www.iranino.co

IRANINOVIN

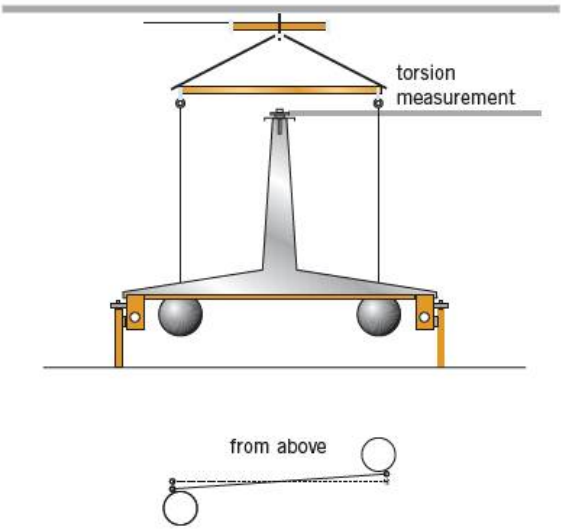


Fig. 8.5 Cavendish's experiment for G

a In such an experiment, one lead sphere has mass 6.22×10^{-3} kg and the mass of the mercury flask is 0.713 kg. Calculate the force between them when they are 72.0 mm apart. (Gravitational constant $G = 6.67 \times 10^{-11}$ N m^2 kg^{-2})

b Comment on the size of the force.

5 The gravitational field strength at the surface of the Moon is 1.62 N kg^{-1}. The radius of the Moon is 1740 km.

a Show that the mass of the Moon is 7.32×10^{22} kg.

b The Moon rotates about its axis (as well as moving in orbit about the Earth). In the future, scientists may wish to put a satellite into an orbit about the Moon, such that the satellite remains stationary above one point on the Moon's surface.

i Explain why this orbit must be an equatorial orbit.

ii The period of rotation of the Moon about its axis is 27.4 days. Calculate the radius of the required orbit. ($G = 6.7 \times 10^{-11}$ N m^2 kg^{-2})

6 **a** Define *gravitational field strength*. *[1]*

b An isolated star has radius R. The mass of the star may be considered to be a point mass at the centre of the star.

The gravitational field strength at the surface of the star is g_s.

On a copy of Fig. 8.6, sketch a graph to show the variation of the gravitational field strength of the star with distance from its centre. You should consider distances in the range R to $4R$. *[2]*

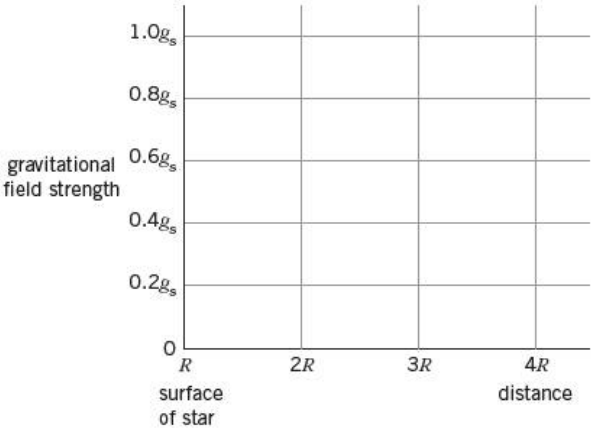


Fig. 8.6

c The Earth and the Moon may be considered to be spheres that are isolated in space with their masses concentrated at their centres.

The masses of the Earth and the Moon are 6.00×10^{24} kg and 7.40×10^{22} kg respectively.

The radius of the Earth is R_E and the separation of the centres of the Earth and the Moon is $60\,R_E$, as illustrated in Fig. 8.7.

i Explain why there is a point between the Earth and the Moon at which the gravitational field strength is zero. *[2]*

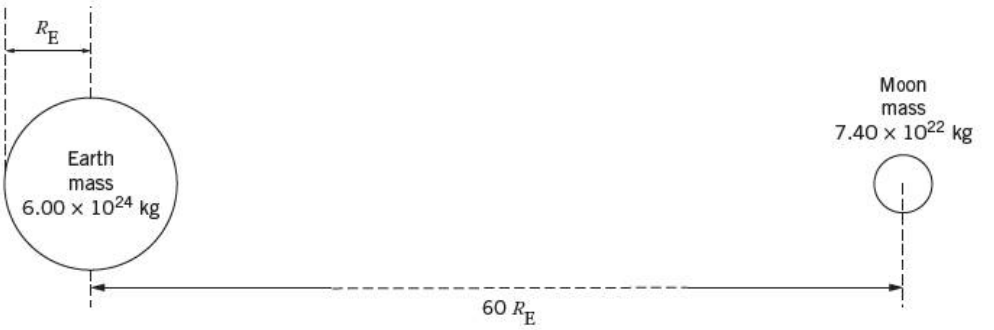


Fig. 8.7

www.iranino.co

IRANINOVIN

ii Determine the distance, in terms of R_E, from the centre of the Earth at which the gravitational field strength is zero. *[3]*

iii On the axes of a copy of Fig. 8.8, sketch a graph to show the variation of the gravitational field strength with position between the surface of the Earth and the surface of the Moon. *[3]*

Cambridge International AS and A Level Physics, 9702/42 Oct/Nov 2010 Q 1

7 a A moon is in a circular orbit of radius r about a planet. The angular speed of the moon in its orbit is ω. The planet and its moon may be considered to be point masses that are isolated in space.

Show that r and ω are related by the expression

$$r^3\omega^2 = \text{constant}$$

Explain your working. *[3]*

b Phobos and Deimos are moons that are in circular orbits about the planet Mars. Data for Phobos and Deimos are shown in Table 8.2.

Table 8.2

moon	radius of orbit/m	period of rotation about Mars/hours
Phobos	9.39×10^6	7.65
Deimos	1.99×10^7	

i Use data from Table 8.2 to determine

- the mass of Mars. *[3]*
- the period of Deimos in its orbit about Mars. *[3]*

ii The period of rotation of Mars about its axis is 24.6 hours.

Deimos is in an equatorial orbit, orbiting in the same direction as the spin of Mars about its axis.

Use your answer in **i** to comment on the orbit of Deimos. *[1]*

Cambridge International AS and A Level Physics, 9702/42 Oct/Nov 2011 Q 1

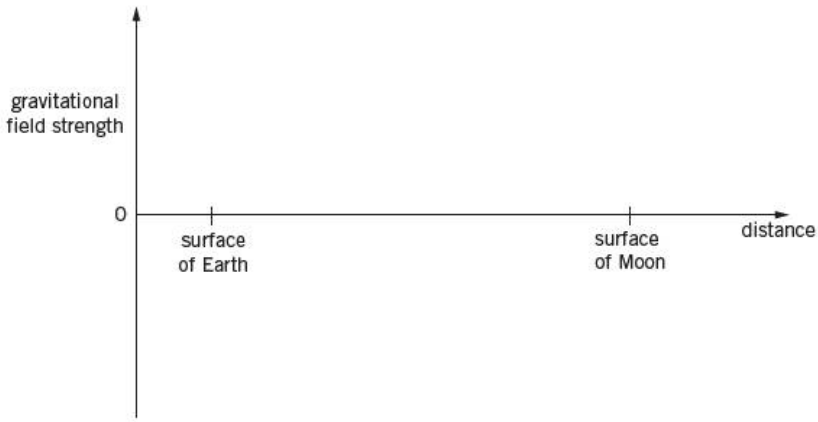


Fig. 8.8

www.iranino.co

IRANINOVIN

10 Ideal gases

By the end of this topic, you will be able to:

10.1 **(a)** Recall and solve problems using the equation of state for an ideal gas expressed as $pV = nRT$, where n = number of moles

10.2 **(a)** infer from a Brownian motion experiment the evidence for the movement of molecules

(b) state the basic assumptions of the kinetic theory of gases

(c) explain how molecular movement causes the pressure exerted by a gas and hence deduce the relationship $pV = \frac{1}{3}Nm\langle c^2\rangle$, where N = number of molecules [a rigorous derivation is not required]

10.3 **(a)** recall that the Boltzmann constant k is given by the expression $k = R/N_A$

(b) compare $pV = \frac{1}{3}Nm\langle c^2\rangle$ with $pV = NkT$ and hence deduce that the average translational kinetic energy of a molecule is proportional to T

Starting points

- All matter consists of atoms or groups of atoms called molecules.
- The spacing of these atoms/molecules and the forces between them determine whether the matter is solid, liquid or gas.
- Gases have no fixed volume and no fixed shape.



Figure 10.1 Robert Boyle

10.1 & 10.3 Equation of state of an ideal gas

Experiments in the seventeenth and eighteenth centuries showed that the volume, pressure and temperature of a given sample of gas are all related.

For a given mass of gas, Robert Boyle (1627–91) found that

> The volume V of a gas is inversely proportional to its pressure p, provided that the temperature is held constant.

This relation is known as **Boyle's law**.

Expressed mathematically, if V is the volume of the gas at pressure p, this is

$$V \propto 1/p$$

Or

$$pV = \text{constant}$$

Another way of writing this equation is

$$p_1V_1 = p_2V_2$$

where p_1 and V_1 are the initial pressure and volume of the gas, and p_2 and V_2 are the final values after a change of pressure and volume carried out at constant temperature.

The effect of temperature on the volume of a gas was investigated by the French scientist Jacques Charles (1746–1823). Charles found that the graph of volume V against temperature θ is a straight line (Figure 10.2). Because gases liquefy when

www.iranino.co

IRANINOVIN

the temperature is reduced, experimental points could not be obtained below the liquefaction temperature. But if the graph was projected backwards, it was found that it cut the temperature axis at about –273 °C.

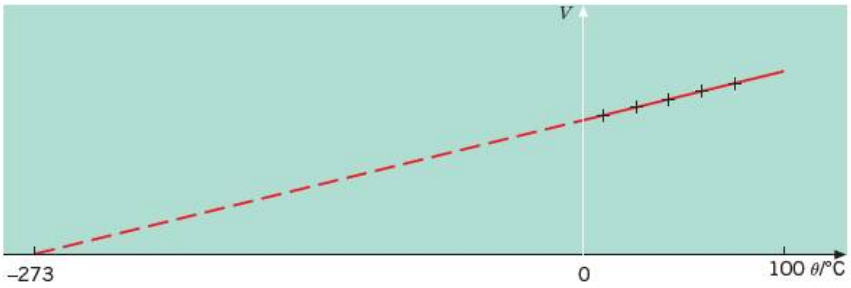


Figure 10.2 Graph of V against θ

The effect of temperature on the pressure of a gas was investigated by another Frenchman, Joseph Gay-Lussac (1778–1850).

The graph of pressure p against temperature θ is a straight line, which, if projected like the volume–temperature graph, also meets the temperature axis at about –273 °C (Figure 10.3). This fact will be used in Topic 11 to introduce the thermodynamic scale of temperature and the idea of the absolute zero of temperature.

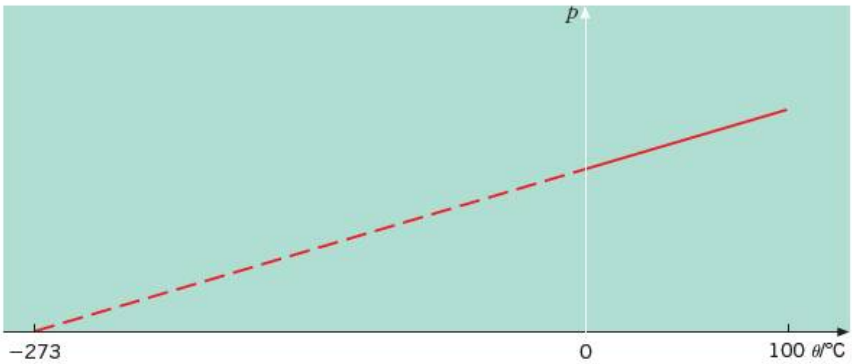


Figure 10.3 Graph of p against θ

If Celsius temperatures are converted to thermodynamic temperatures T, Charles' results are expressed as

$$V \propto T$$

This is **Charles' law**. Another way of writing the equation is

$$\frac{V_1}{T_1} = \frac{V_2}{T_2}$$

where T is the thermodynamic temperature and where V_1 and T_1 are the initial volume and temperature and V_2 and T_2 are the final values.

The corresponding relation between pressure and temperature is $p \propto T$, where T is the thermodynamic temperature.

This is **Gay-Lussac's law**, or the **law of pressures**. Another way of writing the equation is

$$\frac{p_1}{T_1} = \frac{p_2}{T_2}$$

where T is the thermodynamic temperature. We can combine the three gas laws into a single relation between pressure p, volume V and thermodynamic temperature T for a fixed mass of gas. This relation is

www.iranino.co

IRANINOVIN

$$pV \propto T$$

or

$$\frac{p_1V_1}{T_1} = \frac{p_2V_2}{T_2}$$

The laws relate to a fixed mass of gas. Another series of experiments could be carried out to find out how the volume of a gas, held at constant pressure and temperature, depends on the mass of gas present. It would be found that the volume is proportional to the mass. This would give the combined relation

$$pV \propto mT$$

where m is the mass of gas, or

$$pV = AmT$$

where A is a constant of proportionality. However, this is not a very useful way of expressing the relation, as the constant A has different values for different gases. We need to find a way of including the quantity of gas. The way to do this is to express the fixed mass of gas in Boyle's, Charles' and Gay-Lussac's laws in terms of the number of moles of gas present.

We met the mole in Topic 1, as one of the base SI units, and the Avogadro constant N_A. The **mole** (abbreviated mol) is the amount of substance which contains as many elementary entities as there are atoms in 0.012 kg of carbon-12. The **Avogadro constant** N_A is the number of atoms in 0.012 kg of carbon-12. That is, it is the number of atoms in a mole of atoms. It has the value $6.02 \times 10^{23}\,\text{mol}^{-1}$.

The **relative atomic mass** A_r is the ratio of the mass of an atom to one-twelfth of the mass of an atom of carbon-12. Since the definitions of the mole and the Avogadro constant relate to 0.012 kg (or 12 g) of carbon-12, the relative atomic mass is numerically equal to the mass in grams of a mole of atoms. Similarly, the **relative molecular mass** M_r is the ratio of the mass of a molecule to one-twelfth of the mass of an atom of carbon-12, and is numerically equal to the mass in grams of a mole of molecules.

Thus, a simple way of finding the mass of one mole of a particular element is to take its nucleon number (see Topic 26) expressed in grams. For example, the nucleon number of argon ($^{40}_{18}$Ar) is 40. One mole of argon has mass 40 g.

Returning to the combination of the three gas laws and using the number n of moles of the gas, we have

$$pV \propto nT$$

or, putting in a new constant of proportionality R,

$$pV = nRT$$

R is called the **molar gas constant** (sometimes the **universal gas constant**, because it has the same value for all gases). It has the value $8.3\,\text{J K}^{-1}\,\text{mol}^{-1}$.

Sometimes the equation $pV = nRT$ is expressed in the form

$$pV_m = RT$$

where the molar volume V_m is the volume occupied by one mole of the gas.

Another version is

$$pV = NkT$$

where N is the number of molecules in the gas and k is a constant called the **Boltzmann constant**. It has the value $1.38 \times 10^{-23}\,\text{J K}^{-1}$. The molar gas constant R and the Boltzmann constant k are connected through the Avogadro constant N_A.

$$k = R/N_A$$

www.iranino.co

IRANINOVIN

Strictly speaking, the laws of Boyle, Charles and Gay-Lussac are not really laws, as their validity is restricted. They are accurate for real gases only if the pressure of the sample is not too great, and if the gas is well above its liquefaction temperature. But they can be used to define an 'ideal' gas.

An **ideal gas** *is one which obeys the equation of state $pV = nRT$ at all pressures, volumes and temperatures.*

For approximate calculations, the ideal gas equation can be used with real gases if the gas is well above the temperature at which it would liquefy and the pressure is not high.

Examples

1 Find the volume occupied by 1 mole of air at standard temperature and pressure (273 K and 1.01×10^5 Pa), taking R as 8.3 J K^{-1} mol^{-1} for air.
Since $pV = nRT$, $V = nRT/p$. Substituting the values $n = 1$ mol, $R = 8.3$ J K^{-1} mol^{-1}, $T = 273$ K and $p = 1.01 \times 10^5$ Pa, $V = (1 \times 8.3 \times 273)/1.01 \times 10^5 =$ **2.24×10^{-2} m^3**.

2 Find the number of molecules per cubic metre of air at standard temperature and pressure.
We have just shown that the volume occupied by one mole of air at standard temperature and pressure is 2.24×10^{-2} m^3. One mole of air contains N_A molecules, where N_A is the Avogadro constant (6.02×10^{23} mol^{-1}). Thus the number of molecules per cubic metre of air is $6.02 \times 10^{23}/2.24 \times 10^{-2} =$ **2.69×10^{25} m^{-3}**.
Note: It is useful to remember these two quantities, the *molar volume* of a gas and the *number density* of molecules in it. They give an idea of the relatively small volume occupied by a mole of gas at standard temperature and pressure (a cube of side about 28 cm), and the enormous number of molecules in every cubic metre of a gas under these conditions.

3 A syringe contains 25×10^{-6} m^3 of helium gas at a temperature of 20 °C and a pressure of 5.0×10^4 Pa. The temperature is increased to 400 °C and the pressure on the syringe is increased to 2.4×10^5 Pa. Find the new volume of gas in the syringe.
Boyle's law, $pV =$ constant, and Charles' law, $V/T =$ constant, may be combined to give $pV/T =$ constant. This is written in the form $p_1V_1/T_1 = p_2V_2/T_2$ and re-arranged as $V_2 = p_1V_1T_2/p_2T_1$. Substituting the values $p_1 = 5.0 \times 10^4$ Pa, $V_1 = 25 \times 10^{-6}$ m^3, $T_2 = 673$ K, $p_2 = 2.4 \times 10^5$ Pa, $T_1 = 293$ K (again note that temperatures are converted from °C to K), $V_2 = 5.0 \times 10^4 \times 25 \times 10^{-6} \times 673/2.4 \times 10^5 \times 293 =$ **12×10^{-6} m^3**.

Now it's your turn

1 The number of molecules per cubic metre of air at standard temperature and pressure is about 2.7×10^{25} m^{-3}. What is the average separation of these molecules?
2 The mean mass of one mole of air (which is made up mainly of nitrogen, oxygen and argon) is 0.029 kg. What is the density of air at standard temperature and pressure (273 K and 1.01×10^5 Pa)?
3 The volume of a sample of gas is 3.2×10^{-2} m^3 when the pressure is 8.6×10^4 Pa and the temperature is 27 °C. How many moles of gas are there in the sample? How many molecules? What is the number of molecules per cubic metre?
4 A sample of air has volume 15×10^{-6} m^3 when the pressure is 5.0×10^5 Pa. The pressure is reduced to 3.0×10^5 Pa, without changing the temperature. What is the new volume?
5 A sample of gas, originally at standard temperature and pressure (273 K and 1.01×10^5 Pa), has volume 3.0×10^{-5} m^3 under these conditions. When the pressure is increased to 4.0×10^5 Pa, the temperature rises to 40 °C. What is the new volume?

10.2 A microscopic model of a gas

One of the aims of physics is to describe and explain the behaviour of various systems. For mechanical systems, this involves calculating the motion of the parts of the system in detail. For example, we have already seen how to predict the motion of a stone thrown in a uniform gravitational field. Using the equations of uniformly accelerated motion, it is not too difficult to calculate the position and velocity of the stone at any time (Topic 3).

www.iranino.co

IRANINOVIN

However, there are some cases in which it is quite impossible to describe what happens to each component of the system.

This sort of problem arises if we try to describe the properties of a gas in terms of the motion of each of its molecules. The difficulty is that the numbers are so large. One cubic metre of atmospheric air contains about 3×10^{25} molecules. There is no practical method of determining the position and velocity of every single molecule at a given time. Even the most advanced computer would be unable to handle the calculation of the motions of such a very large number of molecules.

In some ways, the fact that the gas is made up of such an enormous number of molecules is an advantage. It means that we can give a large-scale description of the gas in terms of only a few variables. These variables are quantities such as pressure p, volume V and thermodynamic temperature T. They tell us about average conditions in the gas, instead of describing the behaviour of each molecule. We have already met the experimental laws relating to these quantities. Our aim now is to relate the ideal gas equation, which deals with the large-scale (macroscopic) quantities p, V and T, to the small-scale (microscopic) behaviour of the particles of the gas. We shall do this by taking averages over the very large numbers of molecules involved. We shall find that we can derive the equation for Boyle's law when we make very simple assumptions about the atoms or molecules which make up the gas. This is the **kinetic theory of an ideal gas**. We shall also see that temperature can be related to the kinetic energy of the molecules of the gas.

Before we go too far with this model of a gas, is there any experimental evidence for believing that its molecules are moving around all the time? In 1827 the biologist Robert Brown was observing, under a microscope, tiny pollen grains suspended in water. He saw that the grains were always in a jerky, haphazard motion, even though the water was completely still. This movement is now called **Brownian motion**. If we make the assumption that the water molecules are in rapid, random motion, then it is easy to see how the pollen grains could move so jerkily under the random bombardment, from all sides, of the water molecules. We can also reproduce Brownian motion by observing the motion of tiny soot particles in smoke. These particles, too, move in the jerky manner of Brown's pollen grains. Thus it seems that the molecules of both gases and liquids are in the rapid, random motion as required by the molecular model. Nearly a century later, Einstein made a theoretical analysis of Brownian motion and, from the masses and distances moved by the suspended particles, was able to estimate that the diameter of a typical atom is of the order of 10^{-10} m.

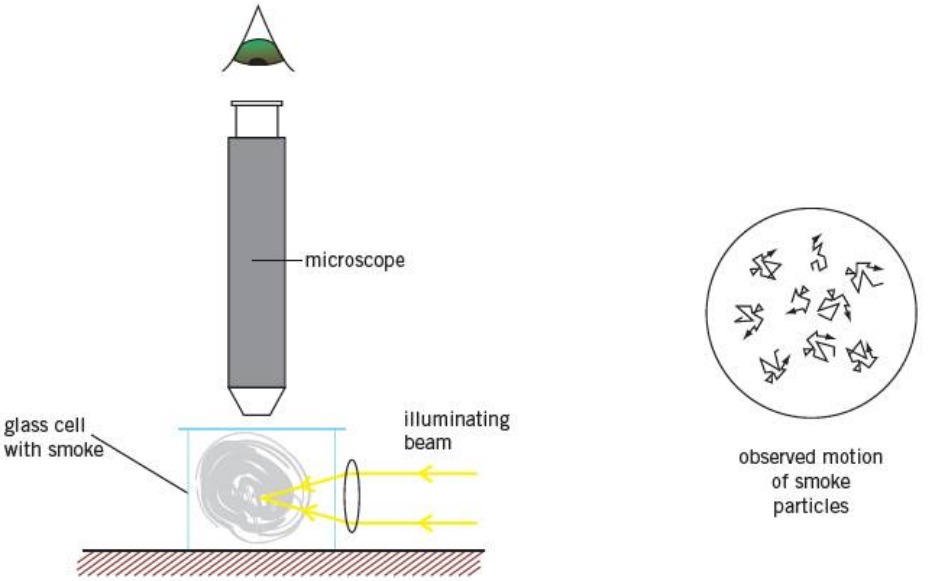


Figure 10.4 Observing Brownian motion

www.iranino.co

IRANINOVIN

If we are able to think of gases and liquids as consisting of large numbers of atoms or molecules in rapid, random motion, what about solids? The obvious difference between a solid and a liquid or gas is that the solid has rigidity. However, the microscopic model explains this by suggesting that the atoms are held in more or less fixed positions by much stronger interatomic forces than exist in liquids or gases. This is because the atoms are arranged quite close together in a solid, whereas in a gas there is a great deal of empty space. Atoms in a solid can still move, but their motion is restricted to vibration about their equilibrium positions. In the change of state of a solid to a liquid, work must be done to break the rigid interatomic forces so that the atoms can move freely. This work is the latent heat of fusion of the solid, which we will meet in Topic 12.

The kinetic theory

An explanation of how a gas exerts a pressure was developed by Robert Boyle in the seventeenth century and, in greater detail, by Daniel Bernoulli in the eighteenth century. The basic idea was that the gas consists of atoms or molecules randomly moving about at great speed (later visualised by Robert Brown).

When a gas molecule hits the wall of its containing vessel, it rebounds. There is a change in momentum of the molecule and it experiences an impulse (see page 63). By Newton's third law, the wall of the vessel also experiences an impulse. There are many collisions per unit time of molecules with the wall and all these collisions and the associated impulses average out to give a constant force on the wall. Force per unit area is pressure.

We shall make some simplifying assumptions about the molecules of the gas. The assumptions of the kinetic theory of an ideal gas are:

- All molecules behave as identical, hard, perfectly elastic spheres.
- The volume of the molecules is negligible compared with the volume of the containing vessel.
- There are no forces of attraction or repulsion between molecules.
- There are many molecules, all moving randomly.

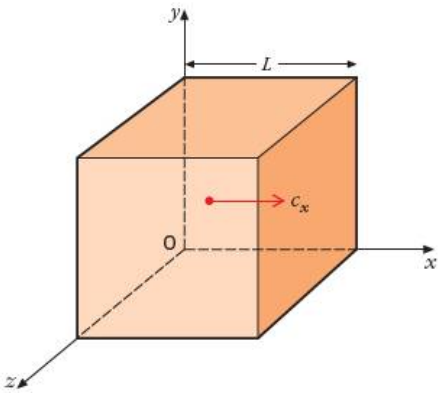


Figure 10.5 A gas molecule in a cubic container

The number of molecules must be very large, so that average behaviour can be considered.

Suppose that the container is a cube of side L (Figure 10.5). The motion of each molecule can be resolved into x-, y- and z-components. For convenience we shall take the x-, y- and z-directions to be parallel to the edges of the cube.

Consider one molecule. Let the x-component of its velocity be c_x. When this molecule collides with the wall of the container perpendicular to the x-axis, the x-component of velocity will be exactly reversed, because (from our assumptions) the collision of the molecule with the wall is perfectly elastic. The time taken for the molecule to move between the two walls perpendicular to the x-axis is L/c_x. To make the round trip from one wall to the opposite one and back again takes $2L/c_x$. This is the time between one collision of the molecule with a wall and its next collision with the same wall.

When the molecule strikes a wall, the component of velocity is reversed in direction, from c_x to $-c_x$. Thus, during each collision with a wall, the x-component of momentum changes by

$$\Delta p_x = 2mc_x$$

where m is the mass of the molecule. The rate at which this molecule transfers momentum to the wall is

$$\textit{change of momentum/time between collisions} = 2mc_x/(2L/c_x) = mc_x^2/L$$

From Newton's second law, this rate of change of momentum is the average force exerted by this particular molecule on the wall through its collisions with the wall. If there are N molecules in the container, the total force is Nmc_x^2/L. Pressure is force divided by area, and the area of the wall is L^2, so the pressure exerted by all N molecules on this wall is Nmc_x^2/L^3.

www.iranino.co

IRANINOVIN

The volume V of the container is L^3, giving us

$$p = Nmc_x^2/V$$

as an expression for the pressure. This expression relates only to the x-component of velocity of the molecules. For a molecule moving with velocity c, an extension of Pythagoras' theorem to three dimensions gives the relation between c and the three components of velocity c_x, c_y and c_z as $c^2 = c_x^2 + c_y^2 + c_z^2$. Because we are dealing with a large number of molecules in random motion, the average value of the component in the x-direction will be the same as for those in the y-direction or in the z-direction. Therefore, taking the averages, $\langle c_x^2\rangle = \langle c_y^2\rangle = \langle c_z^2\rangle$ and $\langle c_x^2\rangle = \frac{1}{3}\langle c^2\rangle$. The notation $\langle c_x^2\rangle$ means the average value of c_x^2. Our expression for the pressure can now be written

$$pV = \tfrac{1}{3}Nm\langle c^2\rangle$$

Since N is the total number of molecules in the container, then Nm is the total mass of gas and Nm/V is the density ρ of the gas. So,

$$p = \tfrac{1}{3}\rho\langle c^2\rangle.$$

Since the average kinetic energy of a molecule is

$$\langle E_k\rangle = \tfrac{1}{2}m\langle c^2\rangle$$

there seems to be a link between our kinetic theory equation for pV and energy. We can find this relation by re-writing the pV equation as

$$pV = \tfrac{2}{3}N\left(\tfrac{1}{2}m\left\langle c^2\right\rangle\right) = \tfrac{2}{3}N\left\langle E_k\right\rangle$$

We have already stated the ideal gas equation in the forms $pV = nRT = NkT$. Indeed, even real gases obey this law reasonably well under normal conditions. If our kinetic theory model and the subsequent theoretical derivation of the equation are correct, we can bring together the two equations for pV. This will allow us to relate the temperature of a gas to the average kinetic energy of its molecules.

$$pV = \tfrac{2}{3}N\left\langle E_k\right\rangle = NkT$$

and

$$\left\langle E_k\right\rangle = \tfrac{3}{2}kT$$

This is an important result. We have obtained a relation between the average kinetic energy of a molecule in a gas and the temperature of the gas. This will allow us to obtain an idea of the average speed of the molecules. Since

$$\left\langle E_k\right\rangle = \tfrac{1}{2}m\langle c^2\rangle = \tfrac{3}{2}kT$$

we have

$$\langle c^2\rangle = 3kT/m$$

and

$$\sqrt{\langle c^2\rangle} = \sqrt{3kT/m}$$

The quantity $\sqrt{\langle c^2\rangle}$ is called the **root-mean-square speed** or **r.m.s. speed** (c_{rms}) of the molecules. It is not exactly equal to the average speed of the molecules, but is often taken as being so. The average speed is about 0.92 of the root-mean-square speed. The difference between the r.m.s. speed and the average speed is highlighted in Example 1 on page 209.

www.iranino.co

IRANINOVIN

Note that the r.m.s. speed is proportional to the square root of the thermodynamic temperature of the gas, and inversely proportional to the square root of the mass of the molecule. Thus, at a given temperature, less massive molecules move faster, on average, than more massive molecules. For a given gas, the higher the temperature, the faster the molecules move.

Examples

1 The speeds of seven molecules in a gas are numerically equal to 2, 4, 6, 8, 10, 12 and 14 units. Find the numerical values of
   (a) the mean speed $\langle c \rangle$,
   (b) the mean speed squared $\langle c \rangle^2$
   (c) the mean-square speed $\langle c^2 \rangle$,
   (d) the r.m.s. speed.

   (a) $\langle c \rangle = (2 + 4 + 6 + 8 + 10 + 12 + 14)/7 =$ **8.0 units**
   (b) $\langle c \rangle^2 = 8^2 =$ **64 units^2**
   (c) $\langle c^2 \rangle = (4 + 16 + 36 + 64 + 100 + 144 + 196)/7 =$ **80 units^2**
   (d) r.m.s. speed $= \sqrt{80} =$ **8.9 units**

2 Find the total kinetic energy of the molecules in one mole of an ideal gas at standard temperature (273 K).
   We know that the average kinetic energy of one molecule is $\frac{3}{2}kT$.
   For one mole of molecules, that is N_A molecules, the energy is

$$\frac{3}{2}N_A kT = \frac{3}{2}RT$$

   Substituting the values $R = 8.3\,\text{J K}^{-1}\,\text{mol}^{-1}$ and $T = 273\,\text{K}$, we have

$$E_k = \frac{3}{2} \times 8.3 \times 273 = \textbf{3400 J mol}^{-1}$$

3 Find the root-mean-square speed of the molecules in nitrogen gas at 27 °C. The mass of a nitrogen molecule is 4.6×10^{-26} kg.
   Since $c_{rms} = \sqrt{3kT/m}$, we have

$$c_{rms} = \sqrt{3 \times 1.38 \times 10^{-23} \times 300 / 4.6 \times 10^{-26}} = \textbf{520 m s}^{-1}$$

Now it's your turn

6 Find the average kinetic energy of a molecule in an ideal gas at a temperature of 500 K.
7 Find the root-mean-square speed of the molecules in hydrogen gas at 500 °C. The relative molecular mass of hydrogen is 2.

Summary

- The equation of state for an ideal gas relates the pressure p, volume V and thermodynamic temperature T of n moles of gas: $pV = nRT$ where $R = 8.3\,\text{J K}^{-1}\,\text{mol}^{-1}$, the molar gas constant.
- For N molecules of gas: $pV = NkT$ where $k = 1.38 \times 10^{-23}\,\text{J K}^{-1}$, the Boltzmann constant.
- The relation between R and k is $k = R/N_A$ where $N_A = 6.02 \times 10^{23}\,\text{mol}^{-1}$, the Avogadro constant (the number of atoms in 0.012 kg of carbon-12).
- The assumptions of the kinetic theory of gases are
  1 Molecules behave as identical, hard, perfectly elastic spheres.
  2 Volume of the molecules is negligible compared with the volume of the containing vessel.
  3 There are no forces of attraction or repulsion between atoms.
  4 There are many molecules, all moving randomly.
- Kinetic theory equation: $pV = \frac{1}{3}Nm\langle c^2 \rangle$, $p = \frac{1}{3}\rho\langle c^2 \rangle$
- Average kinetic energy of a molecule: $\langle E_k \rangle = \frac{1}{2}m\langle c^2 \rangle = \frac{3}{2}kT$
- Root-mean-square speed of molecules: $c_{rms} = \sqrt{3kT/m}$

www.iranino.co

IRANINOVIN

Examination style questions

1 On a day when the atmospheric pressure is 102 kPa and the temperature is 8 °C, the pressure in a car tyre is 190 kPa above atmospheric pressure. After a long journey the temperature of the air in the tyre rises to 29 °C. Calculate the pressure above atmospheric of the air in the tyre at 29 °C. Assume that the volume of the tyre remains constant.

2 A helium-filled balloon is released at ground level, where the temperature is 17 °C and the pressure is 1.0 atmosphere. The balloon rises to a height of 2.5 km, where the pressure is 0.75 atmospheres and the temperature is 5 °C. Calculate the ratio of the volume of the balloon at 2.5 km to that at ground level.

3 Estimate the root-mean-square speed of helium atoms near the surface of the Sun, where the temperature is about 6000 K. (mass of helium atom = 6.6×10^{-27} kg)

4 **a** State what is meant by a *mole*. *[2]*

b Two containers A and B are joined by a tube of negligible volume, as illustrated in Fig. 10.6.

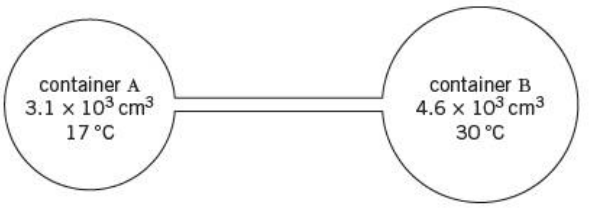


Fig. 10.6

The containers are filled with an ideal gas at a pressure of 2.3×10^5 Pa.

The gas in container A has volume 3.1×10^3 cm^3 and is at a temperature of 17 °C.

The gas in container B has volume 4.6×10^3 cm^3 and is at a temperature of 30 °C.

Calculate the total amount of gas, in mol, in the containers. *[4]*

Cambridge International AS and A Level Physics, 9702/42 May/June 2011 Q 2

5 **a** The kinetic theory of gases is based on some simplifying assumptions.
The molecules of the gas are assumed to behave as hard elastic identical spheres.

State the assumption about idea gas molecules based on

i the nature of their movement, *[1]*

ii their volume. *[2]*

b A cube of volume V contains N molecules of an ideal gas. Each molecule has a component c_x of velocity normal to one side S of the cube, as shown in Fig. 10.7.

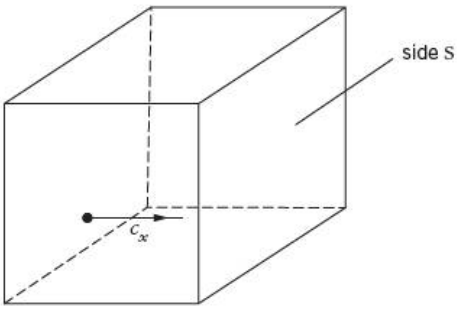


Fig. 10.7

The pressure p of the gas due to the component c_x of velocity is given by the expression

$$pV = Nmc_x^2$$

where m is the mass of a molecule.

Explain how the expression leads to the relation

$$pV = \tfrac{1}{3} Nm\langle c^2 \rangle$$

where $\langle c^2 \rangle$ is the mean square speed of the molecules. *[3]*

c The molecules of an ideal gas have a root-mean-square (r.m.s.) speed of 520 m s^{-1} at a temperature of 27 °C.

Calculate the r.m.s. speed of the molecules at a temperature of 100 °C. *[3]*

Cambridge International AS and A Level Physics, 9702/41 May/June 2012 Q 2

www.iranino.co

IRANINOVIN

11 Temperature

By the end of this topic, you will be able to:

11.1 **(a)** appreciate that thermal energy is transferred from a region of higher temperature to a region of lower temperature

(b) understand that regions of equal temperature are in thermal equilibrium

11.2 **(a)** understand that a physical property that varies with temperature may be used for the measurement of temperature and state examples of such properties

(b) understand that there is an absolute scale of temperature that does not depend on the property of any particular substance (i.e. the thermodynamic scale and the concept of absolute zero)

(c) convert temperatures measured in kelvin to degrees Celsius and recall that $T/\mathrm{K} = T/^{\circ}\mathrm{C} + 273.15$

11.3 **(a)** compare the relative advantages and disadvantages of thermistor and thermocouple thermometers as previously calibrated instruments

Starting points

- Temperature measures the degree of hotness of an object, not the amount of thermal energy in it.
- Temperature is measured using a thermometer.

11.1 Temperature

Our everyday idea of temperature is based on our sense of touch. Putting your hand into a bowl containing ice immediately gives a sense of cold; putting the other hand into a bowl of warm water gives the sense of something that is hot ([Figure 11.1). Intuitively, we would say that the water is at a higher temperature than the ice.



Figure 11.1

Think about a thick metal bar with one end at a higher temperature than the other. For example, one end can be heated by pouring hot water over it, and the other end cooled by holding it under a cold-water tap. The effect of the temperature difference between the ends is that thermal energy is transferred along the bar from the high

www.iranino.co

IRANINOVIN

temperature end to the low temperature end. We can think of this in terms of the vibrations of the atoms of the metal. One atom passes on some of its vibrational energy to its neighbour, which originally had less. If the bar is removed from the arrangement for keeping the ends at different temperatures, eventually the whole of the bar will end up in equilibrium at the same temperature. When different regions in thermal contact are at the same temperature, they are said to be in **thermal equilibrium**.

In Physics, we look for ways of defining and measuring quantities. In the case of temperature, we will first look at ways of measuring this quantity, and then think about the definition.

Many physical properties change with temperature. Most materials (solids, liquids and gases) expand as their temperature is increased. The electrical resistance of a metal wire increases as the temperature of the wire is increased. If two wires of different metals are twisted together at one end, and the other ends are connected to a voltmeter, the reading on the voltmeter depends on the temperature of the junction of the wires. All these properties may be used in different types of thermometer.

A thermometer is an instrument for measuring temperature. The physical property on which a particular thermometer is based is called the thermometric property, and the working material of the thermometer, the property of which varies with temperature, is called the thermometric substance. Thus, in the familiar mercury-in-glass thermometer, the thermometric substance is mercury and the thermometric property is the length of the mercury thread in the capillary tube of the thermometer.

Remember that temperature measures the degree of 'hotness' of a body. It does not measure the *amount* of heat energy.

11.2 Temperature scales

Each type of thermometer can be used to establish its own temperature scale. To do this, the fact that substances change state (from solid to liquid, or from liquid to gas) at fixed temperatures is used to define reference temperatures, which are called **fixed points**. By taking the value of the thermometric property at two fixed points, and dividing the range of values into a number of equal steps (or degrees), we can set up what is called an **empirical scale of temperature** for that thermometer. ('Empirical' means 'derived by experiment'.) If the fixed points are the melting point of ice (the ice point) and the temperature of steam above water boiling at normal atmospheric pressure (the steam point), and if we choose to have one hundred equal degrees between the temperatures corresponding to these fixed points, taken as 0 degrees and 100 degrees respectively, we arrive at the empirical centigrade scale of temperature for that thermometer. If the values of the thermometric property P are P_i and P_s at the ice- and steam-points respectively, and if the property has the value P_θ at an unknown temperature θ, the unknown temperature is given by

$$\theta = \frac{100(P_\theta - P_i)}{(P_s - P_i)}$$

on the empirical centigrade scale of this particular thermometer. This equation is illustrated in graphical form in Figure 11.2.

It is important to realise that the choice of a different thermometric substance and thermometric property would lead to a different centigrade scale. Agreement between scales occurs only at the two fixed points. This happens because the property may not vary linearly with temperature.

This situation, with temperature values depending on the type of thermometer on which they are measured, is clearly unsatisfactory for scientific purposes. It is found that the differences between empirical scales are small in the case of thermometers based on gases as thermometric substances. In the constant-volume gas thermometer (Figure 11.3), the pressure of a fixed volume of gas (measured by the height difference h) is used as the thermometric property.

www.iranino.co

IRANINOVIN

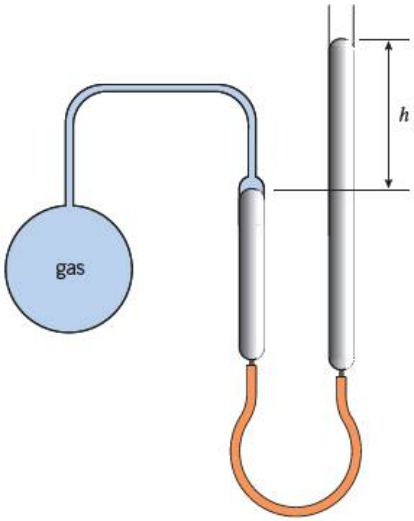


Figure 11.3 Constant-volume gas thermometer

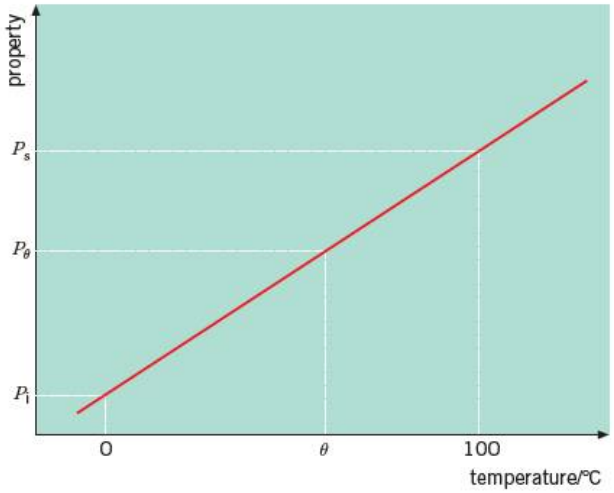


Figure 11.2 Empirical centigrade scale

The differences between the scales of different gas thermometers become even less as the pressures used are reduced. This is because the lower the pressure of a real gas, the more linear is the variation of pressure or volume. If we set up an empirical centigrade scale for a real gas in a constant-volume thermometer by obtaining the pressures of the gas at the ice-point (0 °C) and the steam-point (100 °C) we can extrapolate the graph of pressure p against the centigrade temperature θ to find the temperature at which the pressure of the gas would become zero. This is shown in Figure 11.4.

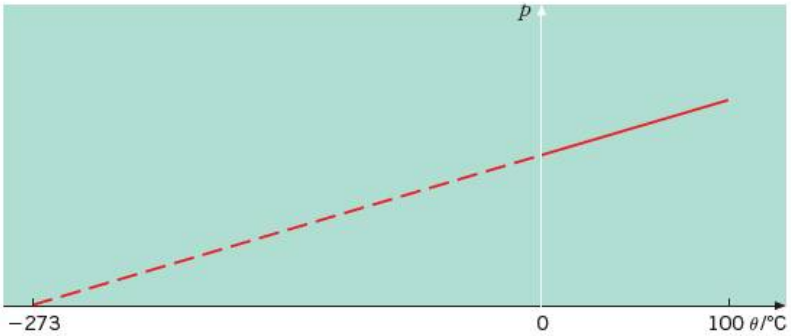


Figure 11.4 Graph of p against θ for constant-volume gas thermometer

The extrapolated temperature will be found to be close to –273 degrees on the empirical centigrade constant-volume gas thermometer scale. If the experiment is repeated with lower and lower pressures of gas in the thermometer, the extrapolated temperature tends to a value of –273.15 degrees. This temperature is known as **absolute zero** on the thermodynamic scale of temperature. It does not depend on the properties of any particular substance.

Example

The length of the mercury column in a mercury-in-glass thermometer is 25 mm at the ice-point and 180 mm at the steam-point. What is the temperature when the length of the column is 55 mm?

Using the equation $\theta = 100(P_\theta - P_i)/(P_s - P_i)$, the thermometric property P is the length of the mercury column.

Thus, $\theta = 100(55 - 25)/(180 - 25) =$ **19.4 °C** on the centigrade scale of this mercury-in-glass thermometer.

www.iranino.co

IRANINOVIN

Now it's your turn

1 A resistance thermometer has resistance 95.2Ω at the ice-point and 138.6Ω at the steam-point. What resistance would be obtained at a temperature of 19.4°C on the centigrade scale of this resistance thermometer?

Thermodynamic temperature

As mentioned above, decreasing the pressure of a real gas makes it behave more and more like an ideal gas. For an ideal gas, the relation between pressure p, volume V and temperature T is

$$\frac{pV}{T} = \text{constant}$$

Here T is the **thermodynamic temperature**. We now use the property of an ideal gas to define the thermodynamic scale of temperature. We have already seen from the idea of extrapolating the p/θ graph for a constant-volume gas thermometer that there seems to be a natural zero of temperature, absolute zero. This is used as one of the fixed points of the thermodynamic scale. The upper fixed point is taken as the triple point of water – the temperature at which ice, water and water vapour are in equilibrium. This is found to be less dependent on environmental conditions, such as pressure, than the ice-point. The thermodynamic temperature of the triple point of water is taken as 273.16 units, by international agreement. This defines the kelvin (symbol K), the unit of thermodynamic temperature.

One kelvin is the fraction 1/273.16 of the thermodynamic temperature of the triple point of water.

Thus, if a constant-volume gas thermometer gives a pressure reading of p_{tr} at the triple point, and a pressure reading of p at an unknown temperature T, the unknown temperature (in K) is given by

$$T = 273.16(p/p_{tr})$$

The Celsius scale

Why choose 273.16 as the number of units between the two fixed points of this scale? The reason is that this number will give 100K between the ice- and steam-points, allowing agreement between the thermodynamic temperature scale and a centigrade scale based on the pressure of an ideal gas. The ideal-gas centigrade scale is based on experiments with real gases at decreasing pressures. This agreement is based on a slightly awkward linking up of the theoretical thermodynamic scale and the empirical constant-volume gas thermometer scale. To avoid this complication, a new scale, the Celsius scale, was defined by international agreement.

The unit of temperature on the Celsius scale is the degree Celsius (°C), which is exactly equal to the kelvin.

The equation linking temperature θ on the Celsius scale and thermodynamic temperature T is

$$\theta/°C = T/K - 273.15 \text{ or } T/K = \theta/°C + 273.15$$

In this equation, θ is measured in °C and T in K. Note that the degree sign ° always appears with the Celsius symbol, but it is never used with the kelvin symbol K.

Be careful to avoid confusion between the numbers 273.15 and 273.16. Absolute zero on the ideal gas constant-volume scale is –273.15 degrees. The fact that 273.16 occurs in the definition of the kelvin means that the temperature of the triple point of water is 0.01 °C. For many purposes where three-significant figure data is provided, in particular in calculations involving conversions of temperatures from the Celsius to the thermodynamic scale or the other way, it is sufficient to work with the number 273.

www.iranino.co

IRANINOVIN

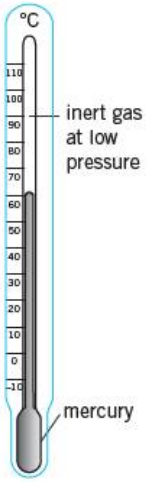


Figure 11.5 Mercury-in-glass thermometer

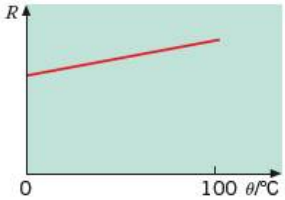


Figure 11.6 Resistance R of a metal wire over a small range of temperatures



Figure 11.7 Thermistors are useful temperature measuring devices

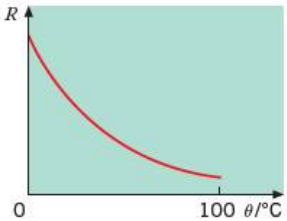


Figure 11.8 Resistance R of a thermistor over a small range of temperatures

www.iranino.co

Example

The pressure reading of a constant-volume gas thermometer is 1.50×10^4 Pa at the triple point of water. When the bulb is placed in a certain liquid, the pressure is 4.28×10^3 Pa. Find the temperature of the liquid.

We use the equation $T = 273.16(p/p_{tr})$. Because the pressure readings are given to three significant figures only, we can approximate the 273.16 to 273.

Thus, $T = 273 \times (4.28 \times 10^3/1.50 \times 10^4) =$ **77.9 K**. (The liquid is liquid nitrogen at its boiling-point.)

Now it's your turn

2 The pressure reading of a constant-volume gas thermometer is 1.367×10^3 Pa at the steam-point. What is the pressure reading at the triple point of water?

11.3 Thermometers

We have already mentioned some of the many different types of thermometer. In choosing a thermometer for particular application, we need to consider a number of aspects. These include accuracy, sensitivity (the distance between divisions on its scale), and the range of temperatures it is able to measure. In some cases the thermometer has to measure rapidly varying temperatures, in which case its speed of response will be important, and whether it can be read directly, or requires time-consuming adjustments by the operator. For other applications it is important that the sensitive part of the thermometer is small and does not absorb much heat, so that it does not change the temperature of the object during measurement.

Gas thermometers (Figure 11.3) are important because they provide a link with the thermodynamic scale. However, they are bulky and inconvenient. Because of the adjustments necessary, they are certainly not suitable for the measurement of rapidly varying temperatures. The bulb containing the gas may have a volume of as much as a litre (1000 cm^3), so a gas thermometer could not be used for measuring the temperature of a small body.

The most familiar type of thermometer is probably the liquid-in-glass type (Figure 11.5), based on the expansion of the liquid. Such thermometers are convenient, sensitive and moderately quick-acting.

The range of the thermometer is limited by the liquid used and the glass containing it. Mercury-in-glass thermometers cover the range from about –40 °C to 350 °C; ethanol-in-glass (the ethanol is normally coloured with a red dye) from about –120 °C to 80 °C. An advantage of liquid-in-glass thermometers is that, over their temperature range, their empirical centigrade scale is very close to the thermodynamic scale.

An important class of thermometers is based on the temperature variation of electrical resistance. The resistance of a metal wire increases with increasing temperature. The range of metal resistance thermometers is very wide, from about –260 °C to 1700 °C. The temperature sensor is a coil of fine wire, often platinum. The variation with temperature of the resistance of a metal wire is not exactly linear with thermodynamic temperature, and for accurate work the resistance thermometer needs to be calibrated at a number of standard temperatures. However, over a small range of temperature, the variation of resistance is linear, as shown in Figure 11.6. Semiconducting materials are also used as resistance thermometers: these are called **thermistors** (see Figure 11.7). The electrical resistance of such devices decreases very rapidly with increasing temperature, as shown in Figure 11.8. Such thermistors are referred to as 'negative temperature coefficient' (NTC) thermistors. Positive temperature coefficient (PTC) thermistors also exist, where resistance increases rapidly with temperature rise. In this syllabus thermistors will always be assumed to be NTC. (Note that the resistance of NTC thermistors is in the opposite sense to the behaviour of a metal wire.) Because of the rapid change of resistance, thermistors are very sensitive devices, but their scales are very non-linear compared with thermodynamic

IRANINOVIN

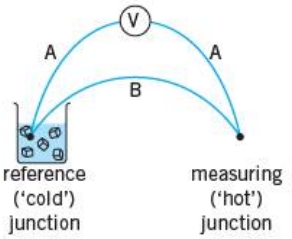


Figure 11.9 Thermocouple

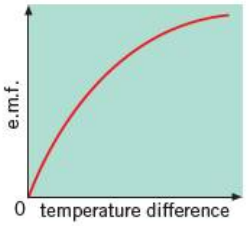


Figure 11.10 Graph of thermocouple e.m.f. against temperature



Figure 11.11 Thermocouple thermometer.

temperature, and calibration is essential. By selection of suitable semiconducting materials, a wide range of temperatures can be covered. Their very small size (some smaller than a pin-head) can be used to measure the temperature of small objects, and also varying temperatures. A simple circuit allows the resistance of the thermistor, or a quantity related to the resistance, to be displayed on a meter, which can then be calibrated in terms of temperature. A common use of a thermistor thermometer is as a temperature sensor in car radiators.

Another type of electrical thermometer makes use of the **thermoelectric effect**. When the junctions of two different conductors (Figure 11.9) are at different temperatures, an e.m.f. (voltage) is developed. The device is known as a **thermocouple**.

The relation between the e.m.f. and the temperature difference θ is not linear with thermodynamic temperature (see Figure 11.10), and calibration is essential.

The e.m.f. generated is rather small (for a thermocouple made of a copper wire and a constantan wire, it is about 5 mV for a temperature difference of 100 °C), but it is easy to arrange a circuit to amplify this e.m.f. and provide a direct reading, which can then be calibrated. Thermocouples cover a large temperature range. Because the sensing part (the junction between the two wires) is very small, the thermocouple can be used to measure rapidly varying temperatures, and the temperatures of small objects. One is shown in Figure 11.11.

Both the thermistor thermometer and the thermocouple thermometer can be read at a distance from where the temperature is being monitored unlike, for example, a liquid-in-glass thermometer.

Summary

- Two regions that are at the same temperature are said to be in thermal equilibrium.
- Empirical centigrade scale of temperature: $\theta = 100(P_\theta - P_i)/(P_s - P_i)$, where θ is the centigrade temperature on the scale of a thermometer based on a property P. P_s, P_i and P_θ are the values of the property at the steam-point, ice-point and at temperature θ respectively.
- The kelvin (K), unit of thermodynamic temperature, is the fraction 1/273.16 of the thermodynamic temperature of the triple point of water.
- Thermodynamic scale of temperature: for a constant-volume gas thermometer, $T = 273.16(p/p_{tr})$, where T is the thermodynamic temperature, p is the pressure reading at temperature T and p_{tr} is the reading at the triple point of water.
- Celsius scale: $\theta = T - 273.15$, where θ is the Celsius temperature (in °C) and T is the thermodynamic temperature (in K).
- The thermometric properties of common thermometers include the length of a column of liquid (in liquid-in-glass thermometers), the pressure of a gas (in constant-volume gas thermometers), the resistance of a coil of wire or of a sample of semiconductor (in thermistors), and the thermoelectric e.m.f. (in thermocouples).

Examination style questions

1 The e.m.f. of a certain thermocouple is 5.60 mV when one junction is placed in melting ice and the other in steam. With one junction in melting ice and the other in a boiling liquid, the e.m.f. is –2.46 mV. Calculate the boiling point of the liquid on the centigrade scale of this thermocouple.

2 The gain (amplification) of an electrical circuit is found to depend on temperature. At 5 °C, the gain is 110 and at 18 °C, it is 125. Assuming that the gain depends linearly on Celsius temperature, estimate the gain when the circuit is used on a hot day, at a temperature of 25 °C.

3 The resistance of a thermistor is 3500 Ω at 2.0 °C. At 20 °C its resistance is 1800 Ω. The thermistor is used as a thermometer. What will be the recorded temperature when the thermistor resistance is 2150 Ω, if the resistance of the thermistor is assumed to vary linearly with temperature?

Suggest why the temperature, measured using a mercury-in-glass thermometer, may give a different value for the temperature.

www.iranino.co

IRANINOVIN

12 Thermal properties of materials

By the end of this topic, you will be able to:

12.1 **(a)** explain using a simple kinetic model for matter
- the structure of solids, liquids and gases
- why melting and boiling take place without a change of temperature
- why the specific latent heat of vaporisation is higher than the specific latent heat of fusion for the same substance
- why a cooling effect accompanies evaporation

(b) define and use the concept of specific heat capacity, and identify the main principles of its determination by electrical methods

(c) define and use the concept of specific latent heat, and identify the main principles of its determination by electrical methods

12.2 **(a)** understand that internal energy is determined by the state of the system and that it can be expressed as the sum of a random distribution of kinetic and potential energies associated with the molecules of a system

(b) relate a rise in temperature of a body to an increase in its internal energy

(c) recall and use the first law of thermodynamics $\Delta U = q + w$ expressed in terms of the increase in internal energy, the heating of the system and the work done on the system

Starting points

- All matter consists of atoms or groups of atoms called molecules.
- Solid, liquid and gas are three different states of matter.
- A substance can change state when energy is involved.

12.1 Solids, liquids and gases, and thermal (heat) energy

A simple kinetic model of matter can be used to describe solids, liquids and gases in terms of the particles (ions, atoms or molecules) making up the particular sample we are considering.

Let us now imagine we are heating a solid. We think of a solid as having molecules, atoms or ions that have definite equilibrium positions, but vibrating about these positions with kinetic energy that depends on temperature.

A solid has fixed volume and fixed shape.

As we come to the melting point of the solid, the energy supplied does not increase the kinetic energy, and hence the temperature, of the solid. Instead, it is used to overcome the forces between the atoms or molecules. This means that the potential energy of the molecules is increased. The increase in potential energy is the latent heat of fusion of the solid. The solid changes to the liquid state.

As heating continues, the atoms can now move about relatively freely in the liquid phase, but they are still close enough to experience interatomic forces, and still have

www.iranino.co

IRANINOVIN

potential energy associated with these forces as well as kinetic energy corresponding to the temperature. The rigid forces between molecules have been broken but the mean separation of molecules has not increased significantly.

A liquid has a fixed volume, no fixed shape and a density similar to that of the solid.

Eventually the temperature of the liquid is increased to its boiling point. If the atoms are to escape into the gaseous phase, the interatomic forces in the liquid must be overcome. The energy input here is the latent heat of vaporisation. This goes to moving the molecules far enough apart for the interatomic forces, and hence the associated potential energy, to be negligible. (Remember that one of the assumptions of the kinetic theory is that, for an ideal gas, there are no forces between atoms or molecules.) Because these forces are negligible,

A gas does not have a fixed shape or volume.

The latent heat of vaporisation is much greater than the latent heat of fusion. The energy required to completely separate the atoms/molecules is greater than that required to break the rigid bonds in the solid. Furthermore, energy must be provided to push back the atmosphere as the liquid turns to vapour. The volume of vapour is much greater than the volume of the liquid.

Note that boiling occurs only at a particular temperature for a given atmospheric pressure. We have associated boiling with a definite input of energy, the latent heat. However, molecules are lost from the surface of a liquid at any temperature by the process of evaporation. The molecules in a liquid do not all move with the same speed, as we know from our work on the kinetic theory (see Topic 10). A molecule with a high enough speed may escape from the attractive forces of the molecules in the surface region of the liquid and leave the liquid entirely. This process can take place at any temperature. However, the higher the liquid temperature, the greater the number of very fast-moving molecules, and the greater the rate of loss of molecules from the surface. This fits with the everyday observation that the evaporation rate increases with an increase in the temperature of the surroundings.

The loss of the fastest molecules means that the average speed of those remaining falls, with a corresponding fall in average kinetic energy and hence in temperature. This means that evaporation tends to cool the remaining liquid. This, too, is an everyday observation.

So, although both evaporation and boiling represent a change of state from liquid to vapour, we distinguish between evaporation and boiling by the fact that evaporation can take place at any temperature, whereas boiling takes place at a fixed temperature for a given pressure of the surroundings. Furthermore, evaporation takes place at the surface of the liquid, whereas boiling occurs in the body of the liquid.

In describing this simple model, the words 'atoms' and 'molecules' are often used interchangeably. More precisely, we ought to consider whether we are describing an element, for which the word 'atom' is usually appropriate (though some elements, such as oxygen O_2, exist as molecules), or a simple molecular compound, for which 'molecule' should be used. For example, in talking about the boiling of water, we are concerned with the forces between water molecules and the loss of water molecules from the surface.

The structure of solids, liquids and gases

Let us now try to get an idea of the spacing of the atoms or molecules in our simple model of matter. We can estimate this quantity for a solid by using known values of its density and the mass containing one mole (the molar mass).

You will find that an answer of the order of 10^{-10} m will come up each time in calculations such as the example on page 219. We can take this as being the order of magnitude of the spacing of atoms in a solid. We can do similar calculations for liquids and gases, although of course the assumption about the cubic arrangement of atoms now does not make sense. In general, the average spacing of atoms in liquids is up to about twice that in solids, or still of the order of 10^{-10} m. For gases, the average spacing is of the order of 10^{-9} m.

www.iranino.co

IRANINOVIN

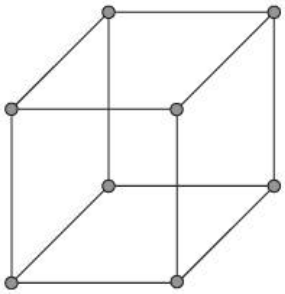


Figure 12.1 Unit cell of a simple cubic lattice – this cell is repeated regularly throughout the crystal

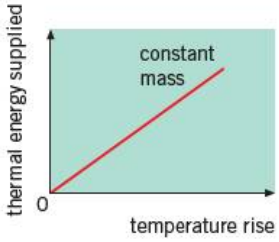


Figure 12.2

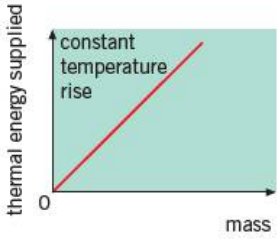
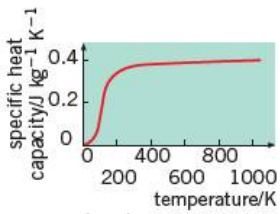


Figure 12.3

Table 12.1 Values of specific heat capacity for different materials

material	specific heat capacity/J kg^{-1} K^{-1}
ethanol	2500
glycerol	2420
ice	2100
mercury	140
water	4200
aluminium	913
copper	390
glass	640



www.iranino.co

Figure 12.4

Example

Assuming that the atoms in iron are arranged in a simple cubic lattice (Figure 12.1), estimate the spacing between atoms in iron. The density of iron is 7900 kg m^{-3} and its molar mass is 56×10^{-3} kg.

One mole contains $N_A = 6.0 \times 10^{23}$ atoms. From the density value, we know that 1.0 m^3 has mass 7900 kg, which is $7900/56 \times 10^{-3} = 1.4 \times 10^5$ mol. This contains $1.4 \times 10^5 \times 6.0 \times 10^{23} = 8.5 \times 10^{28}$ atoms. If these atoms are arranged regularly in a cubic lattice, along one side of the 1.0 m^3 cube there will be $\sqrt[3]{8.5 \times 10^{28}} = 4.4 \times 10^9$ atoms. But the side of the cube has length 1.0 m, so the interatomic spacing is $1/4.4 \times 10^9 = 2.3 \times 10^{-10}$ m. Because we have made an assumption about the simple cubic arrangement of atoms in iron (which is not, in fact, true), we should not give this estimate to more than one significant figure: interatomic spacing = $\mathbf{2 \times 10^{-10}\,m}$. You can carry out this calculation for other metals.

Now it's your turn

1 Estimate the interatomic spacing in aluminium, for which the density is 2700 kg m^{-3} and the molar mass is 27 g.

Thermal (heat) energy

In this section, we will be looking at the effect of heating on temperature.

Specific heat capacity

When a solid, a liquid or a gas is heated, its temperature rises. Plotting a graph of thermal (heat) energy supplied against temperature rise (Figure 12.2), it is seen that the temperature rise $\Delta\theta$ is proportional to the heat energy ΔQ supplied, for a particular mass of a particular substance.

$$\Delta Q \propto \Delta\theta$$

Similarly, the heat energy required to produce a particular temperature rise is proportional to the mass m of the substance being heated (Figure 12.3).

$$\Delta Q \propto m$$

Combining these two relations gives

$$\Delta Q \propto m\Delta\theta$$

or

$$\Delta Q = mc\Delta\theta$$

where c is the constant of proportionality known as the **specific heat capacity** of the substance. In this case, *specific* means *per unit mass*.

The numerical value of the specific heat capacity of a substance is the quantity of heat energy required to raise the temperature of unit mass of the substance by one degree.

The SI unit of specific heat capacity is J kg^{-1} K^{-1}. The unit of specific heat capacity is *not* the joule and this is why, in the definition of specific heat capacity, it is important to make reference to the *numerical value*. Specific heat capacity is different for different substances. Some values are given in Table 12.1.

It should be noted that, for relatively small changes in temperature, specific heat capacity is approximately constant. However, over a wide range of temperature, the value for a substance may vary considerably (Figure 12.4). Unless stated otherwise, specific heat capacity is assumed to be constant.

Determination of specific heat capacity

The specific heat capacity of a metal can be determined using a cylindrical block of the metal as shown in Figure 12.5.

IRANINOVIN

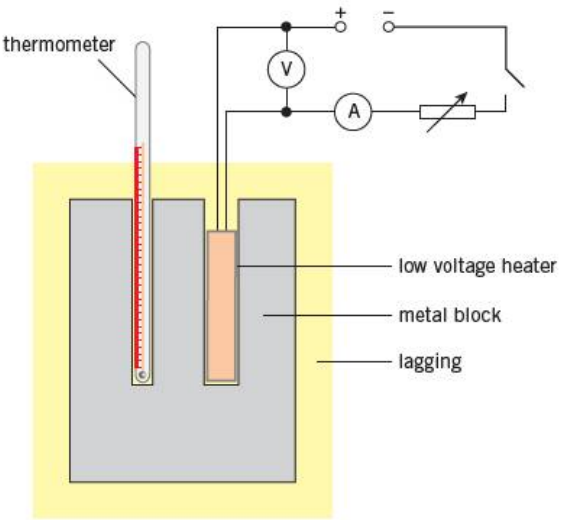


Figure 12.5 Apparatus to determine the specific heat capacity of a metal

The mass M of the metal block is measured using a balance and then placed in thermal lagging.

A low-voltage electrical heater and a thermometer are inserted into holes in the block. The initial temperature T_i of the block is measured. The current is switched on and a stopclock is started. The current is kept constant and values of the current I and the potential difference V are noted. After a suitable time, the current is switched off. The time t is recorded and the highest temperature T_f reached on the thermometer is noted.

Assuming no energy losses to the surroundings,

$$M \times c \times (T_f - T_i) = I \times V \times t$$

where c is specific heat capacity and t is the time in seconds.

The specific heat capacity of a liquid may be found using a similar method. A quantity of liquid is placed in a polystyrene cup and the mass is then measured on a balance. The cup acts as an insulator and, because it has negligible mass, the thermal energy gained by the cup is negligible.

A more reliable method for a liquid is based on continuous flow. The apparatus is illustrated in Figure 12.6.

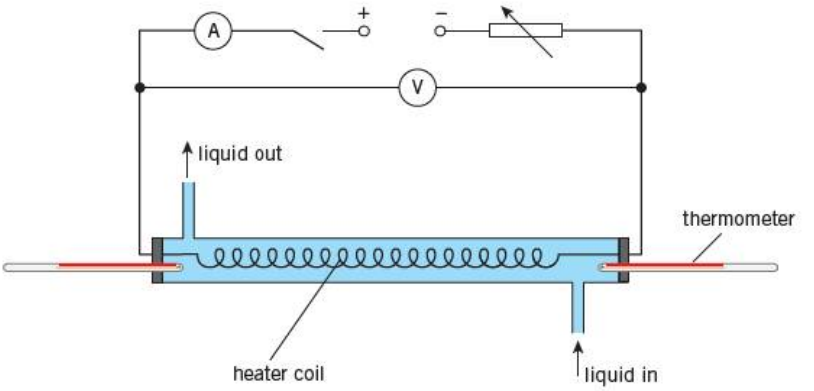


Figure 12.6 Determination of specific heat capacity of a liquid

The liquid flows over the heater at a constant rate. When the input temperature T_i and the output temperature T_o are constant, these are noted together with the voltmeter reading V_1, the ammeter reading I_1 and the mass of liquid M_1 collected in five minutes (300 seconds).

www.iranino.co

IRANINOVIN

The flow rate and the electrical power to the heater are then adjusted so that the input and output temperatures are unchanged. The new readings V_2, I_2 and M_2 are measured.

Initially,

$$\begin{matrix}\textit{thermal energy} \\ \textit{supplied}\end{matrix} = \begin{matrix}\textit{thermal energy} \\ \textit{gained by liquid}\end{matrix} + \begin{matrix}\textit{energy losses to} \\ \textit{surroundings}\end{matrix}$$

$$V_1 \times I_1 \times 300 = M_1 \times c \times (T_o - T_i) + h$$

where c is specific heat capacity and h is the thermal energy lost to the surroundings in 300 s.

For the second set of readings, the temperatures are unaltered and therefore h is unchanged.

So,

$$V_2 \times I_2 \times 300 = M_2 \times c \times (T_o - T_i) + h$$

Subtracting,

$$(V_2 \times I_2 \times 300) - (V_1 \times I_1 \times 300) = (M_2 - M_1) \times c \times (T_o - T_i)$$

Hence, thermal energy losses have been eliminated and c can be determined.

Example

Calculate the quantity of heat energy required to raise the temperature of a mass of 810 g of aluminium from 20 °C to 75 °C. The specific heat capacity of aluminium is 910 J kg^{-1} K^{-1}.

Heat energy required $= m \times c \times \Delta\theta$

$$= \frac{810}{1000} \times 910 \times (75 - 20)$$

$$= \mathbf{4.1 \times 10^4\ J}$$

Now it's your turn

2 Calculate the heat energy gained or lost for the following temperature changes. Use Table 12.1 to obtain values for specific heat capacity.
   (a) 45 g of copper heated from 10 °C to 90 °C
   (b) 1.3 g of ice at 0 °C cooled to −15 °C.

3 Calculate the specific heat capacity of water given that 0.20 MJ of energy are required to raise the temperature of a mass of 600 g of water by 80 K.

Thermal (heat) capacity

Specific heat capacity may be used for a single substance, but often objects are made of several different materials. In order to find the change in heat energy, the specific heat capacity and mass of every substance in the object would need to be known. This can be avoided using the **thermal (heat) capacity** of the object.

The numerical value of the thermal (heat) capacity of a body is the quantity of heat energy required to raise the temperature of the whole body by one degree.

For a body of thermal capacity C, the heat energy ΔQ supplied is related to the rise in temperature $\Delta\theta$ by the expression $\Delta Q = C\Delta\theta$.

Heat capacity has the SI unit J K^{-1}.

Example

A room has a thermal capacity of 1.2 MJ K^{-1}. Calculate the heat energy required to raise the temperature of the room from 18 °C to 20 °C.

$$\Delta Q = C \times \Delta\theta$$

$$= 1.2 \times 10^6 \times (20 - 18)$$

$$= \mathbf{2.4 \times 10^6\ J}$$

www.iranino.co

IRANINOVIN

Now it's your turn

4 A kettle contains 700 g of water. The specific heat capacity of water is 4200 J kg^{-1} K^{-1} and the kettle itself has heat capacity 540 J K^{-1}. The kettle and its contents are heated from 15 °C to the boiling point of water (100 °C).
   (a) Calculate the heat energy gained by:
      (i) the water,
      (ii) the kettle.
   (b) What fraction of the total energy was used in heating the water?

5 An aluminium saucepan has a copper base. Use the information below to determine the thermal capacity of the saucepan.

metal	mass/g	specific heat capacity/J kg^{-1} K^{-1}
aluminium	1160	900
copper	160	390

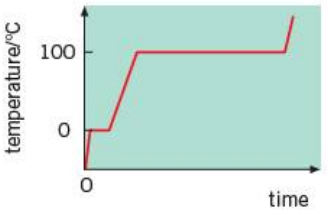


Figure 12.7

Specific latent heat

Figure 12.7 illustrates how the temperature of a mass of water varies with time when it is heated at a constant rate.

At times when the substance is changing phase (ice to water or water to steam), heat energy is being supplied without any change of temperature. Because the heat energy does not change the temperature of the substance, it is said to be *latent* (i.e. hidden). The latent heat required to melt (fuse) a solid is known as **latent heat of fusion**.

The numerical value of the **specific latent heat of fusion** *is the quantity of heat energy required to convert unit mass of solid to liquid without any change in temperature.*

The SI unit of specific latent heat of fusion is J kg^{-1}. For a substance with latent heat of fusion L_f, the quantity of heat energy ΔQ required to fuse (melt) a mass m of solid is given by

$$\Delta Q = mL_f$$

The latent heat required to vaporise a liquid without any change of temperature is referred to as **latent heat of vaporisation**.

The numerical value of the **specific latent heat of vaporisation** *is the quantity of heat energy required to convert unit mass of liquid to vapour without any change in temperature.*

The SI unit of specific latent heat of vaporisation is the same as that for fusion i.e. J kg^{-1}. For a substance with latent heat of vaporisation L_v, the quantity of heat energy ΔQ required to vaporise a mass m of liquid is given by

$$\Delta Q = mL_v$$

When a vapour condenses (vapour becomes liquid), the latent heat of vaporisation is released. Similarly, when a liquid solidifies (liquid becomes solid), the latent heat of fusion is released. Some values of specific latent heat of fusion and of vaporisation are given in Table 12.2.

Table 12.2 Values of specific latent heat

material	specific latent heat of fusion/kJ kg^{-1}	specific latent heat of vaporisation/kJ kg^{-1}
ice/water	330	2260
ethanol	108	840
copper	205	4840
sulfur	38.1	

www.iranino.co

IRANINOVIN

Example

Use the information given in Table 12.2 to determine the heat energy required to melt 50g of sulfur at its normal melting point.

Heat energy required = $m \times L_f$

$$= \frac{50}{1000} \times 38.1 \times 1000$$

$$= \mathbf{1900\,J}$$

Now it's your turn

6 Where appropriate, use the information given in Table 12.2.

(a) Calculate the heat energy required to:

(i) melt 50g of ice at 0°C,

(ii) evaporate 50g of water at 100°C.

(b) Using your answers to **a**, determine how many times more energy is required to evaporate a mass of water than to melt the same mass of ice.

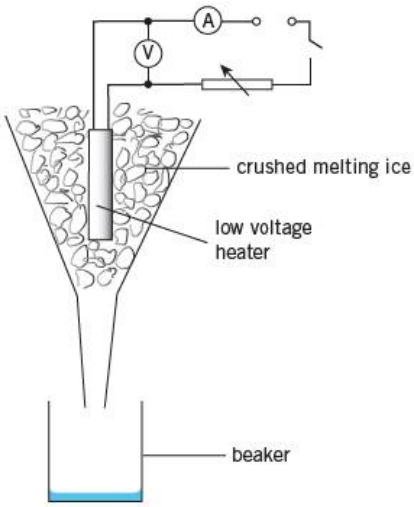


Figure 12.8 Apparatus to determine the latent heat of fusion of ice.

Determination of latent heat of fusion of ice

A large funnel is filled with crushed melting ice, as illustrated in Figure 12.8.

The apparatus is left until water drips out of the funnel at a constant rate. A weighed beaker is then placed under the funnel so that the mass m of ice melted in five minutes (300s) is determined. The heater is then switched on and the current is kept constant. Melting ice is added to the funnel to ensure that the heater is always surrounded by ice. When water drips out of the funnel at a constant rate, the mass M melted in 300s is found. The readings of the voltmeter V and the ammeter I are taken.

Ice will melt due to thermal energy gained from the surroundings. The mass of ice melted due to the heater is $(M - m)$. Thus

$$(M - m) \times L = V \times I \times 300.$$

The latent heat of fusion of ice L can be determined.

Determination of latent heat of vaporisation of water

A beaker of boiling water is placed on the pan of a balance, as illustrated in Figure 12.9.

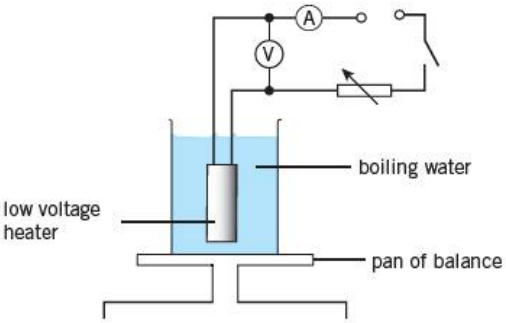


Figure 12.9 Apparatus to determine the latent heat of vaporisation of water

The heater is switched on and the current is kept constant. The readings of the voltmeter V_1 and the ammeter I_1 are taken. When the water is boiling at a steady rate, the change in the reading M_1 of the balance in 300s is determined. The current in the heater is now changed and the experiment is repeated. The new readings are V_2, I_2 and M_2.

Since the water is boiling at a steady rate,

thermal energy supplied by heater = *energy used to vaporise water* + *energy losses to surroundings*

$$V_1 \times I_1 \times 300 = M_1 \times L + h$$

www.iranino.co

IRANINOVIN

where L is the latent heat of vaporisation of the water and h is the energy lost to the surroundings in 300s.

The temperature of the apparatus does not change and so h does not change. For the second set of readings,

$$V_2 \times I_2 \times 300 = M_2 \times L + h$$

Subtracting

$$(V_2 I_2) \times 300 - (V_1 I_1) \times 300 = (M_2 - M_1) \times L$$

Hence, L can be calculated.

Exchanges of heat energy

When a hot object and a cold object come into contact, heat energy passes from the hot object to the cold one so that the two objects reach the same temperature. The law of conservation of energy applies in that the thermal energy (heat) gained by the cold object is equal to the thermal energy lost by the hot object. This does, of course, assume that no energy is lost to the surroundings. This simplification enables temperatures to be calculated.

Examples

1 A mass of 0.30kg of water at 95°C is mixed with 0.50kg of water at 20°C. Calculate the final temperature of the water, given that the specific heat capacity of water is 4200J kg^{-1} K^{-1}.

Hint: always start by writing out a word equation containing all the gains and losses of heat energy.

heat energy lost by hot water = heat energy gained by cold water

$$(m \times c \times \Delta\theta_1) = (M \times c \times \Delta\theta_2)$$

$$0.30 \times 4200 \times (95 - \theta) = 0.50 \times 4200 \times (\theta - 20)$$

where θ is the final temperature of the water.

$$1260 \times (95 - \theta) = 2100 \times (\theta - 20)$$

$$119700 - 1260\theta = 2100\theta - 42000$$

$$161700 = 3360\theta$$

$$\theta = \mathbf{48\,°C}$$

2 A mass of 12g of ice at 0°C is placed in a drink of mass 210g at 25°C. Calculate the final temperature of the drink, given that the specific latent heat of fusion of ice is 334kJ kg^{-1} and that the specific heat capacity of water and the drink is 4.2kJ kg^{-1} K^{-1}.

energy lost by drink = energy gained by melting ice + energy gained by ice water

$$(m \times c \times \Delta\theta_1) = (M \times L_f) + (M \times c \times \Delta\theta_2)$$

$$M \times c \times \Delta\theta_2 = \frac{12}{1000} \times 4.2 \times 1000 \times (\theta - 0)$$

$$= 50.4\theta$$

$$\frac{210}{1000} \times 4.2 \times 1000 \times (25 - \theta) = \frac{12}{1000} \times 334 \times 1000 + 50.4\theta$$

where θ is the final temperature of the drink. Simplifying,

$$22050 - 882\theta = 4008 + 50.4\theta$$

$$18\,042 = 932.4\theta$$

$$\theta = \mathbf{19\,°C}$$

Now it's your turn

Use the data in Table 12.1 and in Table 12.2 where appropriate.

7 A lump of copper of mass 120g is heated in a gas flame. It is then transferred to a mass of 450g of water, initially at 20°C. The final temperature of the copper and the water is 31°C. Calculate the temperature of the gas flame.

8 Steam at 100°C is passed into a mass of 350g of water, initially at 15°C. The steam condenses. Calculate the mass of steam required to raise the temperature of the water to 80°C.

www.iranino.co

IRANINOVIN

12.2 Internal energy

We have seen that the molecules of an ideal gas possess kinetic energy, and that this kinetic energy is proportional to the thermodynamic temperature of the gas (Topic 10). The sum of the kinetic energies of all the molecules, due to their random motion, is called the **internal energy** of the ideal gas. Not all molecules have the same kinetic energy, because they are moving with different speeds, but the sum of all the kinetic energies will be a constant if the gas is kept at a constant temperature.

For a real gas, the situation is a little more complicated. Because the molecules of a real gas exert forces on each other, at any instant there will be a certain potential energy associated with the positions that the molecules occupy in space. Because the molecules are moving randomly, the potential energy of a given molecule will also vary randomly. But at a given temperature the total potential energy of all the molecules will remain constant. If the temperature changes, the total potential energy will also change. Furthermore, the molecules of a real gas collide with each other, and will interchange kinetic energy during the collisions. For this real gas, the internal energy is given by the sum of the potential energies and the kinetic energies of all the molecules. It is important to realise that looking at a single molecule will give us very little information. Its kinetic energy will be changing all the time as it collides with other gas molecules, and its potential energy is also changing as its position relative to the other molecules in the gas changes. This single molecule has a kinetic energy which is part of the very wide range of kinetic energies of the molecules of the gas. We say that there is a *distribution* of molecular kinetic energies. The distribution is illustrated in Figure 12.10. Similarly, there is a distribution of molecular potential energies. But by adding up the kinetic and potential energies of all the molecules in the gas, the random nature of the kinetic and potential energies of the single molecule is removed.

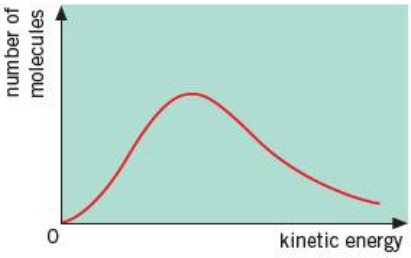


Figure 12.10 Distribution of molecular kinetic energies

The idea of internal energy can be extended to all states of matter. In a liquid, intermolecular forces are stronger as the molecules are closer together, so the potential energy contribution to internal energy becomes more significant. The kinetic energy contribution is still due to the random motion of the molecules in the liquid. We can think of a solid as being made up of atoms or molecules which oscillate (vibrate) about equilibrium positions. Here, the potential energy contribution is caused by the strong binding forces between atoms, and the kinetic energy contribution is due to the motion of the vibrating atoms.

Internal energy is determined by the state of the system and can be expressed as the sum of a random distribution of kinetic and potential energies associated with the molecules of the system.

The concept of internal energy is particularly useful as it helps us to distinguish between temperature and heat. Using an ideal gas as an example, temperature is a measure of the average kinetic energy of the molecules. It, therefore, does not depend on how many molecules are present in the gas. Internal energy (again for an ideal gas), however, is the *total* kinetic energy of the molecules, and clearly does depend on how many molecules there are. In general, heating refers to a *transfer* of energy from one substance to another, often as a result of a temperature difference. If two objects at different temperatures are placed in contact, there will be a flow of thermal energy from the object at the higher temperature to the one at the lower temperature. The direction of energy flow is determined by the difference in temperature, not by any difference in internal energies. If 10 g of a liquid at 30 °C is placed in contact with, or mixed into, 100 g of the same liquid at 20 °C, the direction of energy flow is from the liquid at 30 °C to the liquid at 20 °C, even though the liquid at 20 °C has a greater internal energy than the smaller mass of liquid at 30 °C.

www.iranino.co

IRANINOVIN

The first law of thermodynamics

We have already met the law of conservation of energy (Topic 6). There, it was stated in the following form: energy can neither be created nor destroyed, it can only be transformed from one form to another. In this section we shall see how this conservation law may be re-stated in relation to terms such as work, thermal energy and internal energy. This will lead to an understanding of the first law of thermodynamics.

Thermodynamics is the study of processes involving the transfer of thermal energy and the doing of work. In thermodynamics, it is necessary to define the system under consideration. For example, the system may be an ideal gas in a cylinder fitted with a piston, or an electric heating coil in a container of liquid.

In Topic 6 we established the scientific meaning of work. Work is done when energy is transferred by mechanical means. We have just seen that heating is a transfer of energy due to a difference in temperature. Work and heating both involve a transfer of energy, but by different means.

We also know that the internal energy of a system is the total energy, kinetic and potential, of the various parts of the system. For a system consisting of an ideal gas, the internal energy is simply the total kinetic energy of all the atoms or molecules of the gas. For such a system, we would expect the internal energy to increase if thermal energy (heat) were added to the gas, or if work were done on it. In both cases we are adding energy to the system. By the law of conservation of energy, this energy cannot just disappear; it must be transformed to another type of energy. It appears as an increase in the internal energy of the gas; that is, the total kinetic energy of the molecules is increased. But if the *total* kinetic energy of the molecules increases, their *average* kinetic energy is also increased. Because the average kinetic energy is a measure of temperature, the addition of energy to the ideal gas shows up as an increase in its temperature. We can express this transformation of energy as an equation.

The increase in internal energy of a system is equal to the sum of the thermal energy added to the system and the work done on it.

The increase in internal energy is given the symbol ΔU, thermal energy added is represented by q and work done on the system by w. The equation is then

$$\Delta U = q + w$$

Note the sign convention which has been adopted. A positive value of q means that heat has been added to the system. A positive value of w means that work is done on the system. A positive value for ΔU means an increase in internal energy.

If the system does work, then we show this by writing $-w$. If heat leaves the system, we show this by writing $-q$. Take care! There is an alternative sign convention that takes the work done *by* the system as a positive quantity. To avoid confusion, write down the sign convention you are using every time you quote the first law.

Let's see how the first law of thermodynamics applies to some simple processes. First, think about a change in the pressure and volume of a gas in a cylinder fitted with a piston. The cylinder and piston are insulated, so that no heat can enter or leave the gas. (The thermodynamic name for such a change is an **adiabatic** change. That is, no thermal energy is allowed to enter or leave the system.) If no thermal energy enters or leaves the gas, q in the first law equation is zero. Thus

$$\Delta U = 0 + w$$

If work is done on the gas by pushing the piston in, w is positive (remember the sign convention!) and ΔU will also be positive. That is, the internal energy increases and, because temperature is proportional to internal energy, the temperature of the gas rises. An adiabatic change can be achieved even if the cylinder and piston are not well insulated. Moving the piston rapidly, so that the heat has no time to flow in or out, is

www.iranino.co

IRANINOVIN

just as effective. You will have noticed that a cycle pump gets hot as a result of brisk pumping. This is because the gas in the pump is being compressed adiabatically. Work is being done on the gas, the pump strokes are too rapid for the heat to escape, and the internal energy, and hence the temperature, increases. Another example is the diesel engine, where air in the cylinder is compressed so rapidly that the temperature rises to a point that, when fuel is injected into the cylinder, it is above its ignition temperature.

Now think about an electric kettle containing water. Here the element provides thermal energy to the system. The quantity q in the first law equation is positive (the sign convention is 'thermal energy added, q positive'). No mechanical work is done on or by the water, so w in the first law equation is zero. Thus

$$\Delta U = q + 0$$

The fact that q is positive means that ΔU is also positive. Internal energy, and hence temperature, increases.

When a substance changes from solid to liquid, intermolecular bonds are broken, thus increasing the potential energy component of the internal energy. During the melting process, the temperature does not change, and therefore the kinetic energy of the molecules does not change. Most substances expand on melting, and thus external work is done. By the first law, thermal energy must be supplied to the system, and this thermal energy is the latent heat.

Volume changes associated with evaporation are much greater than those associated with melting. The external work done is much greater during vaporisation, and thus latent heat of vaporisation is much greater than latent heat of fusion.

Example

200J of thermal energy is added to a system, which does 150J of work. Find the change in internal energy of the system.

We use the first law of thermodynamics in the form $\Delta U = q + w$ with the sign convention that q is positive if heat is supplied to the system and w is positive if work is done on the system. Here $q = 200$J and $w = -150$J (the system is doing the work, hence the minus sign). Therefore $\Delta U = 200 - 150 =$ **50J**. This change in internal energy is an increase.

Now it's your turn

9 An **isothermal** change is one which takes place at constant temperature. Explain why, in any isothermal change, the change in internal energy is zero. In such a change, 200J of thermal energy is added to a system. How much work is done on or by the system?

www.iranino.co

IRANINOVIN

Summary

- Melting, boiling and evaporation are all examples of changes of phase (solid to liquid, and liquid to vapour).
- All these changes of phase require an input of energy (latent heat) to overcome the interatomic forces.
- Boiling takes place at a fixed temperature for a particular atmospheric pressure, but evaporation occurs at all temperatures. Boiling occurs in the body of the liquid; evaporation occurs at its surface.
- In a solid, the interatomic spacing is of the order of 10^{-10} m.
- Specific heat capacity is numerically equal to the heat energy required to raise the temperature of unit mass of substance by one degree. The SI unit of specific heat capacity is J kg^{-1} K^{-1}.
- The heat energy ΔQ required to raise the temperature of a mass m of substance of specific heat capacity c by an amount $\Delta\theta$ is given by the expression: $\Delta Q = mc\Delta\theta$
- The heat capacity of an object is numerically equal to the heat energy required to raise the temperature of the whole body by one degree. The SI unit of heat capacity is J K^{-1}.
- The heat energy ΔQ required to raise the temperature of an object having a heat capacity C by an amount $\Delta\theta$ is given by the expression: $\Delta Q = C\Delta\theta$
- Specific latent heat of fusion is numerically equal to the quantity of heat energy required to convert unit mass of solid to liquid without any change in temperature.
- Specific latent heat of vaporisation is numerically equal to the quantity of heat energy required to convert unit mass of liquid to vapour without any change in temperature.
- When a substance of mass m changes its state the quantity of heat energy required is given by $\Delta Q = mL$ where L is the appropriate specific latent heat.
- Specific latent heat has the SI unit of J kg^{-1}.
- The internal energy of a system is the sum of the random kinetic and potential energies of the various parts of the system. For an ideal gas, the internal energy is the total kinetic energy of random motion of the molecules. Internal energy is a measure of the temperature of the system.
- The first law of thermodynamics expresses the law of conservation of energy. The increase in internal energy ΔU of a system is equal to the sum of the heat q added to the system and the work w done on it:

  $\Delta U = q + w$

  (Sign convention: positive q, heat is added to the system; positive w, work is done on the system; positive ΔU, increase in internal energy.)

Examination style questions

In the following questions, use the data in Table 12.1 and in Table 12.2 where appropriate.

1 A piece of copper of mass 170 g is cooled in a freezer. It is then dropped into water at 0 °C, causing 4.0 g of water to freeze. Determine the temperature inside the freezer.

2 A liquid-in-glass thermometer consists of a mass of 62 g of glass and 3.5 g of mercury.

  a Calculate the thermal capacity of the thermometer.

  b The thermometer is used to measure the temperature of some glycerol of mass 90 g. Before the thermometer is inserted into the glycerol, the thermometer records 18 °C and the temperature of the glycerol is 42 °C. Calculate the temperature recorded on the thermometer when placed in the glycerol.

  c Using your answer to **b**, suggest why such thermometers cannot be used to measure reliably the temperature of a small mass of substance.

3 a A jet of steam at 100 °C is directed into a hole in a large block of ice. After the steam has been switched off, the condensed steam and the melted ice are both at 0 °C. The mass of water collected in the hole is 206 g. Calculate the mass of steam condensed.

  b Suggest why a scald with steam is much more serious than one involving boiling water.

4 A mass of 450 g of frozen vegetables is taken from a freezer at –20 °C. The vegetables are immediately placed in a saucepan containing 1100 g of boiling water. The saucepan has a thermal capacity of 900 J K^{-1}. The final temperature of the saucepan, water and vegetables is 83 °C.

  a Calculate the specific heat capacity of the vegetables.

  b The saucepan and its contents are then heated using a heater which provides 1200 J of thermal energy each second. Determine how long it takes to bring the water back to its boiling point.

www.iran-archive.com

IRANINOVIN

5 An ideal gas expands isothermally, doing 250 J of work. What is the change in internal energy? How much thermal energy is absorbed in the process?

6 50 J of heat energy is supplied to a fixed mass of gas in a cylinder. The gas expands, doing 20 J of work. Calculate the change in internal energy of the gas.

7 **a** Define *specific latent heat*. *[2]*

b The heater in an electric kettle has a power of 2.40 kW.

When the water in the kettle is boiling at a steady rate, the mass of water evaporated in 2.0 minutes is 106 g.

The specific latent heat of vaporisation of water is 2260 J g^{-1}.

Calculate the rate of loss of thermal energy to the surroundings of the kettle during the boiling process. *[3]*

Cambridge International AS and A Level Physics, 9702/41 May/June 2012 Q 3

8 **a** **i** State the basic assumption of the kinetic theory of gases that leads to the conclusion that the potential energy between the atoms of an ideal gas is zero. *[1]*

ii State what is meant by the *internal energy* of a substance. *[2]*

iii Explain why an increase in internal energy of an ideal gas is directly related to a rise in temperature of the gas. *[2]*

b A fixed mass of an ideal gas undergoes a cycle PQRP of changes as shown in Fig. 12.11.

i State the change in internal energy of the gas during one complete cycle PQRP. *[1]*

ii Calculate the work done on the gas during the change from P to Q. *[2]*

iii Some energy changes during the cycle PQRP are shown in Table 12.3.

Table 12.3

change	work done on gas/J	heating supplied to gas/J	increase in internal energy/J
P → Q	.....	−600	.....
Q → R	0	+720	.....
R → P	.....	+480	.....

Complete a copy of Table 12.3 to show all of the energy changes. *[3]*

Cambridge International AS and A Level Physics, 9702/42 Oct/Nov 2010 Q 2

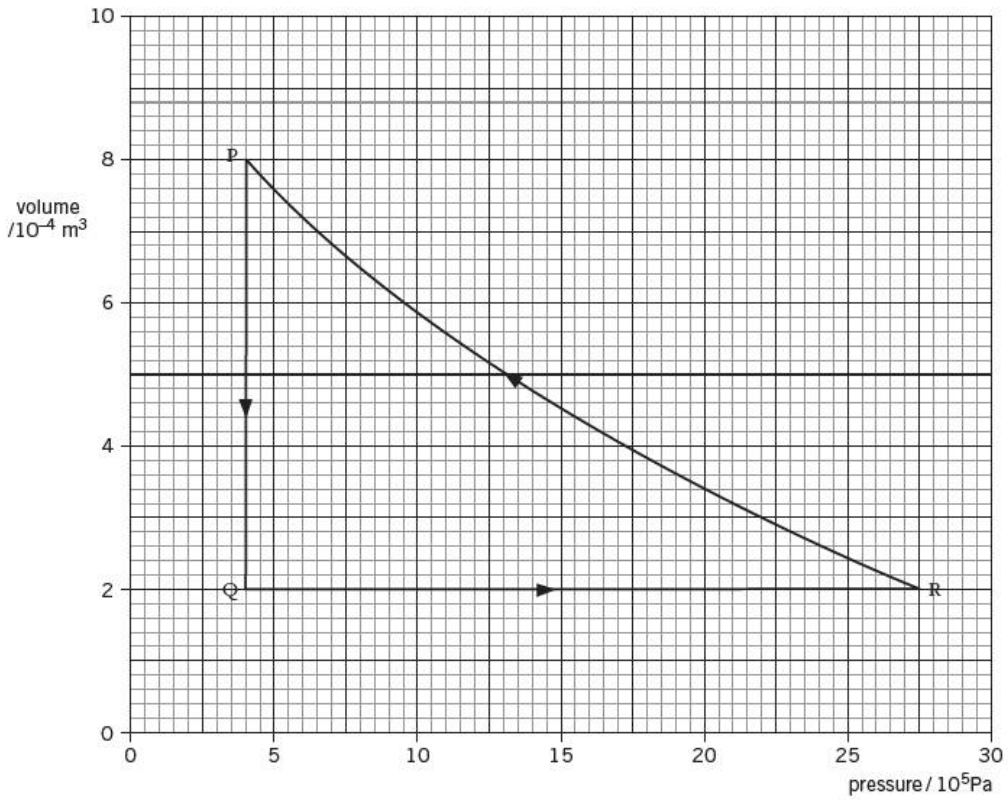


Fig. 12.11

www.iranino.co

IRANINOVIN

13 Oscillations

By the end of this topic, you will be able to:

13.1 (a) describe simple examples of free oscillations
(b) investigate the motion of an oscillator using experimental and graphical methods
(c) understand and use the terms amplitude, period, frequency, angular frequency and phase difference, and express the period in terms of both frequency and angular frequency
(d) recognise and use the equation $a = -\omega^2 x$ as the defining equation of simple harmonic motion
(e) recall and use $x = x_0 \sin \omega t$ as a solution to the equation $a = -\omega^2 x$
(f) recognise and use the equations $v = v_0 \cos \omega t$ and $v = \pm\omega\sqrt{(x_0^2 - x^2)}$
(g) describe, with graphical illustrations, the changes in displacement, velocity and acceleration during simple harmonic motion

13.2 (a) describe the interchange between kinetic and potential energy during simple harmonic motion

13.3 (a) describe practical examples of damped oscillations with particular reference to the effects of the degree of damping and the importance of critical damping
(b) describe practical examples of forced oscillations and resonance
(c) describe graphically how the amplitude of a forced oscillation changes with frequency near the natural frequency of the system, and understand qualitatively the factors that determine the frequency response and the sharpness of resonance
(d) appreciate that there are some circumstances in which resonance is useful and other circumstances in which resonance should be avoided

Starting points

- An object that moves to-and-fro continuously is said to be oscillating or vibrating.
- Oscillations occur in many different systems from the very small (e.g. atoms) to the very large (e.g. buildings).

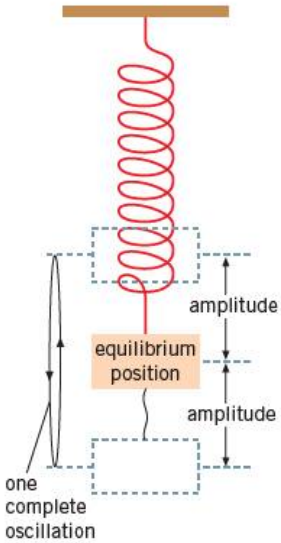


Figure 13.1 Oscillation of a mass on a spring

www.iranino.co

13.1 Oscillations

Some movements involve repetitive to-and-fro motion, such as a pendulum, the beating of a heart, the motion of a child on a swing, and the vibrations of a guitar string. Another example would be a mass bouncing up and down on a spring, as illustrated in Figure 13.1. One complete movement from the starting or rest position, up, then down and finally back up to the rest position, is known as an **oscillation**.

The time taken for one complete oscillation or vibration is referred to as the **period** T *of the oscillation.*

The oscillations repeat themselves.

The number of oscillations or vibrations per unit time is the **frequency** f.

Frequency may be measured in hertz (Hz), where one hertz is one oscillation per second ($1\,\text{Hz} = 1\,\text{s}^{-1}$). However, frequency may also be measured in min^{-1}, hour^{-1}, etc. For example, it would be appropriate to measure the frequency of the tides in h^{-1}.

Since period T is the time for one oscillation then

frequency $f = 1/T$

IRANINOVIN

As the mass oscillates, it moves from its rest or equilibrium position.

The distance from the equilibrium position is known as the **displacement**.

This is a vector quantity, since the displacement may be on either side of the equilibrium position.

The **amplitude** *(a scalar quantity) is the maximum displacement.*

Some oscillations maintain a constant period even when the amplitude of the oscillation changes. Galileo discovered this fact for a pendulum. He timed the swings of an oil lamp in Pisa Cathedral, using his pulse as a measure of time. Oscillators that have a constant time period are called isochronous, and may be made use of in timing devices. For example, in quartz watches the oscillations of a small quartz crystal provide constant time intervals. Galileo's experiment was not precise, and we now know that a pendulum swinging with a large amplitude is not isochronous.

The quantities period, frequency, displacement and amplitude should be familiar from our study of waves in Topic 14. It should not be a surprise to meet them again, as the idea of oscillations is vital to the understanding of waves.

Displacement–time graphs

It is possible to plot displacement–time graphs (see AS Level Topic 14) for oscillators. One experimental method is illustrated in Figure 13.2. A mass on a spring oscillates above a position sensor that is connected to a computer through a datalogging interface, causing a trace to appear on the monitor.

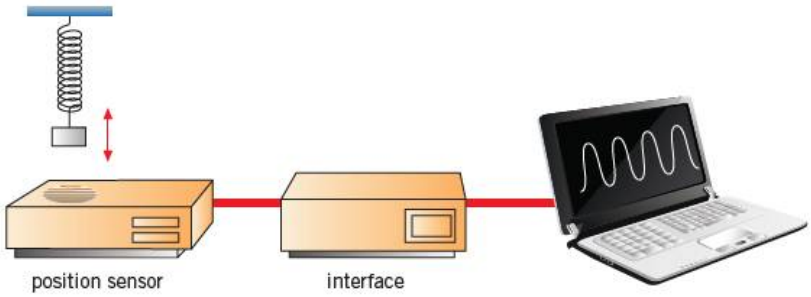


Figure 13.2 Apparatus for plotting displacement–time graphs for a mass on a spring

The graph describing the variation of displacement with time may have different shapes, depending on the oscillating system. For many oscillators the graph is approximately a sine (or cosine) curve. A sinusoidal displacement–time graph is a characteristic of an important type of oscillation called **simple harmonic motion** (s.h.m.). Oscillators which move in s.h.m. are called **harmonic** oscillators. We shall analyse simple harmonic motion in some detail, because it successfully describes many oscillating systems, both in real life and in theory. Fortunately, the mathematics of s.h.m. can be approached through a simple defining equation. The properties of the motion can be deduced from the relations between graphs of displacement against time, and velocity against time, which we met in Topic 3.

Simple harmonic motion (s.h.m.)

Simple harmonic motion is defined as the motion of a particle about a fixed point such that its acceleration a is proportional to its displacement x from the fixed point, and is directed towards the point.

Mathematically, we write this definition as

$$a = -\omega^2 x$$

www.iranino.co

IRANINOVIN

where ω^2 is a constant. We take the constant as a squared quantity, because this will ensure that the constant is always positive (the square of a positive number, or of a negative number, will always be positive). Why worry about keeping the constant positive? This is because the minus sign in the equation must be preserved. It has a special significance, because it tells us that the acceleration a is always in the opposite direction to the displacement x. Remember that both acceleration and displacement are vector quantities, so the minus sign is shorthand for the idea that the acceleration is always directed towards the fixed point from which the displacement is measured. This is illustrated in Figure 13.3.

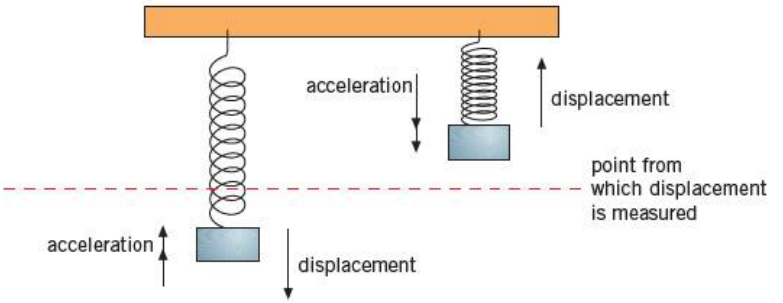


Figure 13.3 Directions of displacement and acceleration are always opposite.

The defining equation is represented in a graph of a against x as a straight line, of negative gradient, through the origin, as shown in Figure 13.4. The gradient is negative because of the minus sign in the equation. Note that both positive and negative values for the displacement should be considered.

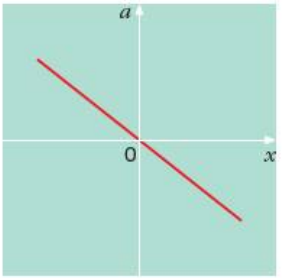


Figure 13.4 Graph of the defining equation for simple harmonic motion

The square root of the constant ω^2 (that is, ω) is known as the **angular frequency** of the oscillation. This angular frequency ω is related to the frequency f of the oscillation by the expression

$$\omega = 2\pi f$$

By Newton's second law, the force acting on a body is proportional to the acceleration of the body. The defining equation for simple harmonic motion can thus be related to the force acting on the particle. If the acceleration of the particle is proportional to its displacement from a fixed point, the resultant force acting on the particle is also proportional to the displacement. We can bring in the idea of the direction of the acceleration by specifying that the force is always acting towards the fixed point, or by calling it a **restoring** force.

Solution of equation for simple harmonic motion

In order to find the displacement–time relation for a particle moving in a simple harmonic motion, we need to solve the equation $a = -\omega^2 x$. To derive the solution requires mathematics which is beyond the requirements of A/AS Physics. However, you need to know the form of the solution. This is

$$x = x_0 \sin \omega t$$
or
$$x = x_0 \cos \omega t$$

where x_0 is the amplitude of the oscillation. The solution $x = x_0 \sin \omega t$ is used when, at time $t = 0$, the particle is at its equilibrium position where $x = 0$. Conversely, if at time $t = 0$ the particle is at its maximum displacement, $x = x_0$, the solution is $x = x_0 \cos \omega t$. The variation with time t of the displacement x for the two solutions is shown in Figure 13.5.

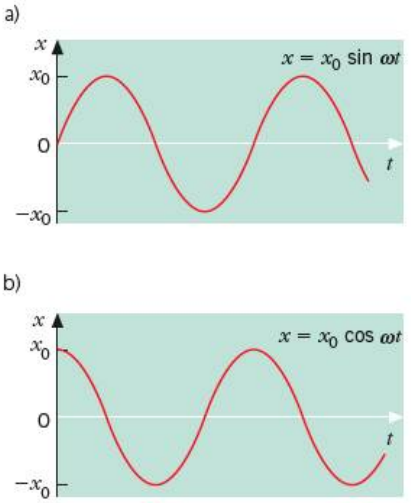


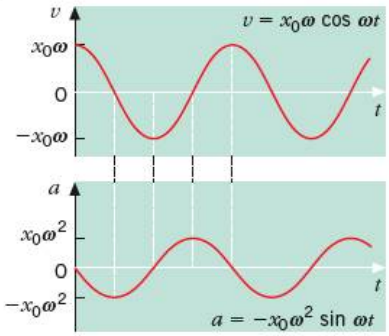
Figure 13.5 Displacement–time curves for the two solutions to the s.h.m. equation

In Topic 3 it was shown that the gradient of a displacement–time graph may be used to determine velocity. Referring to Figure 13.5, it can be seen that, at each time at which $x = x_0$, the gradient of the graph is zero (this applies to both solutions). Thus, the velocity is zero whenever the particle has its maximum displacement. If we think

www.iranino.co

IRANINOVIN

a)



b)

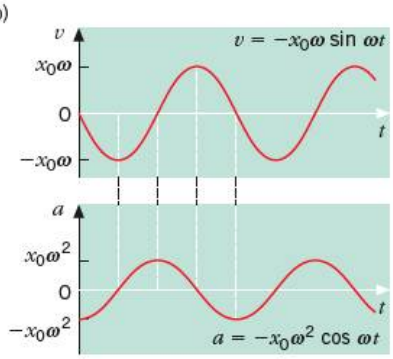


Figure 13.6 Velocity–time and acceleration–time graphs for the two solutions to the s.h.m. equation

about a mass vibrating up and down on a spring, this means that when the spring is fully stretched and the mass has its maximum displacement, the mass stops moving downwards and has zero velocity. Also from Figure 13.5, we can see that the gradient of the graph is at a maximum whenever $x = 0$. This means that when the spring is neither under- or over-stretched the speed of the mass is at a maximum. After passing this point, the spring forces the mass to slow down until it changes direction.

If a full analysis is carried out, it is found that the variation of velocity with time is cosinusoidal if the sinusoidal displacement solution is taken, and sinusoidal if the cosinusoidal displacement solution is taken. This is illustrated in Figure 13.6.

The velocity v of the particle is given by the expressions

$$v = x_0\omega \cos \omega t \text{ when } x = x_0 \sin \omega t$$

and

$$v = -x_0\omega \sin \omega t \text{ when } x = x_0 \cos \omega t$$

In each case, there is a phase difference between velocity and displacement. The velocity curve is $\pi/2$ rad ahead of the displacement curve. (If the phase angle is not considered, the variation with time of the velocity is the same in each case.) The maximum speed v_0 is given by

$$v_0 = x_0\omega$$

There is an alternative expression for the velocity.

$$v = \pm\omega\sqrt{(x_0^2 - x^2)}$$

which is derived on page 236.

For completeness, Figure 13.6 also shows the variation with time of the acceleration a of the particle. This could be derived from the velocity–time graph by taking the gradient. The equations for the acceleration are

$$a = -x_0\omega^2 \sin \omega t \text{ when } x = x_0 \sin \omega t$$

and

$$a = -x_0\omega^2 \cos \omega t \text{ when } x = x_0 \cos \omega t$$

Note that, for both solutions, these equations are consistent with the defining equation for simple harmonic motion, $a = -\omega^2 x$. You can easily prove this by eliminating $\sin \omega t$ from the first set of equations, and $\cos \omega t$ from the second.

Example

The displacement x at time t of a particle moving in simple harmonic motion is given by $x = 0.25 \cos 7.5t$, where x is in metres and t is in seconds.

(a) Use the equation to find the amplitude, frequency and period for the motion.

(b) Find the displacement when $t = 0.50$ s.

(a) Compare the equation with $x = x_0 \cos \omega t$. The amplitude $x_0 = \mathbf{0.25\,m}$. The angular frequency $\omega = 7.5$ rad s^{-1}. Remember that $\omega = 2\pi f$, so the frequency $f = \omega/2\pi = 7.5/2\pi = \mathbf{1.2\,Hz}$. The period $T = 1/f = 1/1.2 = \mathbf{0.84\,s}$.

(b) Substitute $t = 0.50$ s in the equation, remembering that the angle ωt is in radians and not degrees. $\omega t = 7.5 \times 0.50 = 3.75$ rad $= 215°$. So $x = 0.25 \cos 215° = \mathbf{-0.20\,m}$.

Now it's your turn

1 A mass oscillating on a spring has an amplitude of 0.10 m and a period of 2.0 s.

(a) Deduce the equation for the displacement x if timing starts at the instant when the mass has its maximum displacement.

(b) Calculate the time interval from $t = 0$ before the displacement is 0.08 m.

www.iranino.co

IRANINOVIN

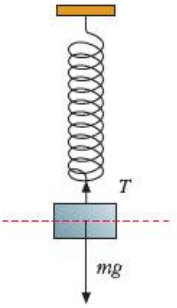


Figure 13.7 Mass on a helical spring

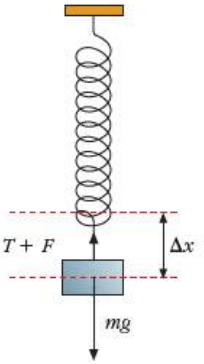


Figure 13.8 Mass on a helical spring: additional extension Δx

Examples of simple harmonic motion

By definition, an object whose acceleration is proportional to the displacement from the equilibrium position, and which is always directed towards the equilibrium position, is undergoing simple harmonic motion. We now look at two simple examples of oscillatory motion which approximate to this definition. They may easily be set up as demonstrations of s.h.m.

Mass on a helical spring

Figure 13.7 illustrates a mass m suspended from a spring. (This sort of spring is called a *helical* spring because it has the shape of a helix.)

The weight mg of the mass is balanced by the tension T in the spring. When the spring is extended by an amount Δx, as in Figure 13.8, there is an additional upward force in the spring given by

$$F = -k\Delta x$$

where k is a constant for a particular spring, known as the spring constant. The spring constant is a measure of the stiffness of the spring (see Topic 9). A stiff spring has a large value of k; a more flexible spring has a smaller value of k and, for the same force, would have a larger extension than one with a large spring constant. The spring constant k is given by the expression

$$k = \frac{F}{\Delta x}$$

and is measured in newtons per metre (N m^{-1}).

When the mass is released, the restoring force F pulls the mass towards the equilibrium position. (The minus sign in the expression for F shows the direction of this force.) The restoring force is proportional to the displacement. This means that the acceleration of the mass is proportional to the displacement from the equilibrium position and is directed towards the equilibrium position. This is the condition for simple harmonic motion.

The full theory shows that, for a mass m suspended from a light spring having spring constant k, the period T of the oscillations is given by

$$T = 2\pi\sqrt{\frac{m}{k}}$$

For oscillations to be simple harmonic, the spring must obey Hooke's law throughout – that is, the extensions must not exceed the limit of proportionality. Furthermore, for large amplitude oscillations, the spring may become slack. Ideally, the spring would have no mass, but if the suspended mass is more than about 20 times the mass of the spring, the error involved in assuming that the spring has no mass is less than 1%.

This example of simple harmonic motion is particularly useful in modelling the vibrations of molecules. A molecule containing two atoms oscillates as if the atoms were connected by a tiny spring. The spring constant of this spring depends upon the type of bonding between the atoms. The frequency of oscillation of the molecule can be measured experimentally using spectroscopy, and this gives direct information about the bonding. This model can be extended to solids, where atoms are often thought of as being connected to their neighbours by springs. Again, this leads to an experimental way of obtaining information about interatomic forces in the solid.

The simple pendulum

Figure 13.9 illustrates a simple pendulum. Ideally, a simple pendulum is a point mass m on a light, inelastic string. In real experiments we use a pendulum bob of finite size. When the bob is pulled aside through an angle θ and then released, there will be a restoring force acting in the direction of the equilibrium position.

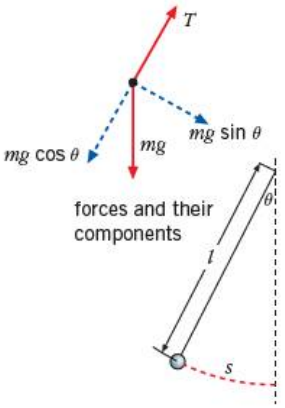


Figure 13.9 The simple pendulum

www.iranino.co

IRANINOVIN

Because the simple pendulum moves in the arc of a circle, the displacement will be an angular displacement θ, rather than the linear displacement x we have been using so far.

The two forces on the bob are its weight mg and the tension T in the string. The component of the weight along the direction of the string, $mg\cos\theta$, is equal to the tension in the string. The component of the weight at right angles to the direction of the string, $mg\sin\theta$, is the restoring force F. This makes the bob accelerate towards the equilibrium position.

The restoring force depends on $\sin\theta$. As θ increases, the restoring force is not proportional to the displacement (in this case θ), and so the motion is oscillatory but not simple harmonic. However, the situation is different if the angle θ is kept small (less than about 5°). For these small angles, θ is proportional to $\sin\theta$. In fact, if θ is measured in radians, then

θ in radians $\approx \sin\theta$

You can check this using your calculator. Some values of θ/°, θ/rad and $\sin\theta$ are given in Table 13.1.

Table 13.1

θ/°	θ/rad	$\sin\theta$
1.00	0.0175	0.0175
2.00	0.0349	0.0349
3.00	0.0524	0.0523
5.00	0.0873	0.0872
10.00	0.1745	0.1736

This means that, for small-amplitude oscillations (the angle of the string to the vertical should be less than about 5°), the pendulum bob oscillates with simple harmonic motion.

A full treatment of the theory shows that the period T is related to the length l of the pendulum (l is the distance between the centre of mass of the pendulum bob and the point of suspension – see Figure 13.9) by the expression

$$T = 2\pi\sqrt{\frac{l}{g}}$$

where g is the acceleration of free fall. An experiment in which the period of a simple pendulum is measured can be used to determine the acceleration of free fall. The experiment is repeated for different lengths of pendulum, and the gradient of a graph of T^2 against l is $4\pi^2/g$. This provides an alternative to dynamics experiments in which the time for a body to fall through measured distances is determined, and g is calculated from the equation of motion (see Topic 3).

Example

A helical spring is clamped at one end and hangs vertically. It extends by 10 cm when a mass of 50 g is hung from its free end. Calculate:

(a) the spring constant of the spring,

(b) the period of small amplitude oscillations of the mass.

(a) Using $k = F/\Delta x$, the spring constant $k = 50 \times 10^{-3} \times 9.8/10 \times 10^{-2} = \mathbf{4.9\,N\,m^{-1}}$

(b) Using $T = 2\pi\sqrt{(m/k)}$, the period $T = 2\pi\sqrt{(50 \times 10^{-3}/4.9)} = \mathbf{0.63\,s}$.

Now it's your turn

2 The acceleration of free fall at the Earth's surface is 9.8 m s^{-2}. Calculate the length of a simple pendulum which would have a period of 1.0 s.

3 The acceleration of free fall on the Moon's surface is 1.6 m s^{-2}. Calculate, for the pendulum in question 2, its period of oscillation on the Moon.

13.2 Energy changes in simple harmonic motion

Kinetic energy

On page 233, we saw that the velocity of a particle vibrating with simple harmonic motion varies with time and, consequently, with the displacement of the particle.

www.iranino.co

IRANINOVIN

For the case where displacement x is zero at time $t = 0$, displacement and velocity are given by

$$x = x_0 \sin \omega t$$

and

$$v = x_0\omega \cos \omega t$$

There is a trigonometrical relation between the sine and the cosine of an angle θ, which is $\sin^2 \theta + \cos^2 \theta = 1$. Applying this relation, we have

$$x^2/x_0^2 + v^2/x_0^2\omega^2 = 1$$

which leads to

$$v^2 = x_0^2\omega^2 - x^2\omega^2$$

and so

$$v = \pm\omega\sqrt{(x_0^2 - x^2)}$$

(If we had taken the displacement and velocity equations for the case when the displacement is a maximum at $t = 0$, exactly the same relation would have been obtained. Try it!)

The kinetic energy of the particle (of mass m) oscillating with simple harmonic motion is $\frac{1}{2}mv^2$. Thus, the kinetic energy E_k at displacement x is given by

$$E_k = \frac{1}{2}m\omega^2(x_0^2 - x^2)$$

The variation with displacement of the kinetic energy is shown in Figure 13.10.

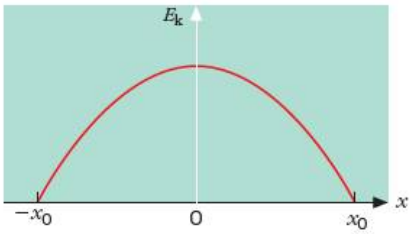


Figure 13.10 Variation of kinetic energy in s.h.m.

Potential energy

The defining equation for simple harmonic motion can be expressed in terms of the restoring force F_{res} acting on the particle. Since $F = ma$ and $a = -\omega^2 x$ then at displacement x, this force is

$$F_{res} = -m\omega^2 x$$

where m is the mass of the particle. To find the change in potential energy of the particle when the displacement increases by Δx, we need to find the work done against the restoring force.

The work done in moving the point of application of a force F by a distance Δx is $F\Delta x$. In the case of the particle undergoing simple harmonic motion, we know that the restoring force is directly proportional to displacement. To calculate the work done against the restoring force in giving the particle a displacement x, we take account of the fact that F_{res} depends on x by taking the average value of F_{res} during this displacement. The average value is just $\frac{1}{2}m\omega^2 x$, since the value of F_{res} is zero at $x = 0$ and increases linearly to $m\omega^2 x$ at displacement x. Thus, the potential energy E_p at displacement x is given by *average restoring force* × *displacement*, or

$$E_p = \frac{1}{2}m\omega^2 x^2$$

The variation with displacement of the potential energy is shown in Figure 13.11.

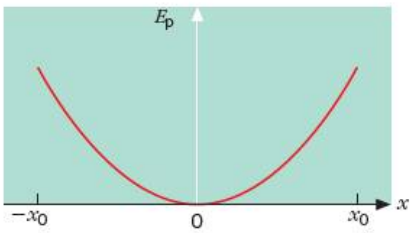


Figure 13.11 Variation of potential energy in s.h.m.

www.iranino.co

IRANINOVIN

Total energy

The total energy E_{tot} of the oscillating particle is given by

$$E_{tot} = E_k + E_p$$
$$= \frac{1}{2}m\omega^2 (x_0^2 - x^2) + \frac{1}{2}m\omega^2 x^2$$

$$E_{tot} = \frac{1}{2}m\omega^2 x_0^2$$

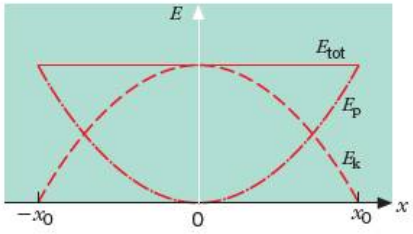


Figure 13.12 Energy variations in s.h.m.

This total energy is constant since m, ω and x_0 are all constant. We might have expected this result, as it merely expresses the law of conservation of energy.

The variations with displacement x of the total energy E_{tot}, the kinetic energy E_k and the potential energy E_p are shown in Figure 13.12.

Example

A particle of mass 60 g oscillates in simple harmonic motion with angular frequency 6.3 rad s^{-1} and amplitude 15 mm. Calculate

(a) the total energy,

(b) the kinetic and potential energies at half-amplitude (at displacement x = 7.5 mm).

(a) Using $E_{tot} = \frac{1}{2}m\omega^2 x_0^2$,

$$E_{tot} = \frac{1}{2} \times 60 \times 10^{-3} \times 6.3^2 \times (15 \times 10^{-3})^2$$
$$= \mathbf{2.7 \times 10^{-4}\ J}$$

(don't forget to convert g to kg and mm to m)

(b) Using $E_k = \frac{1}{2}m\omega^2 (x_0^2 - x^2)$,

$$E_k = \frac{1}{2} \times 60 \times 10^{-3} \times 6.3^2 \times [(15 \times 10^{-3})^2 - (7.5 \times 10^{-3})^2]$$
$$= \mathbf{2.0 \times 10^{-4}\ J}$$

Using $E_p = \frac{1}{2}m\omega^2 x^2$,

$$E_p = \frac{1}{2} \times 60 \times 10^{-3} \times 6.3^2 \times (7.5 \times 10^{-3})^2$$
$$= \mathbf{0.7 \times 10^{-4}\ J}$$

Note that $E_{tot} = E_k + E_p$, as expected.

Now it's your turn

4 A particle of mass 0.40 kg oscillates in simple harmonic motion with frequency 5.0 Hz and amplitude 12 cm. Calculate, for the particle at displacement 10 cm:

(a) the kinetic energy

(b) the potential energy

(c) the total energy.

13.3 Free and damped oscillations

A particle is said to be undergoing **free oscillations** *when the only external force acting on it is the restoring force.*

There are no forces to dissipate energy and so the oscillations have constant amplitude. Total energy remains constant. This is the situation we have been considering so far. Simple harmonic oscillations are free oscillations.

www.iranino.co

IRANINOVIN

In real situations, however, frictional and other resistive forces cause the oscillator's energy to be dissipated, and this energy is converted eventually into thermal energy. The oscillations are said to be **damped**.

The total energy of the oscillator decreases with time. The damping is said to be light when the amplitude of the oscillations decreases gradually with time. This is illustrated in Figure 13.13. The decrease in amplitude is, in fact, exponential with time. The period of the oscillation is slightly greater than that of the corresponding free oscillation.

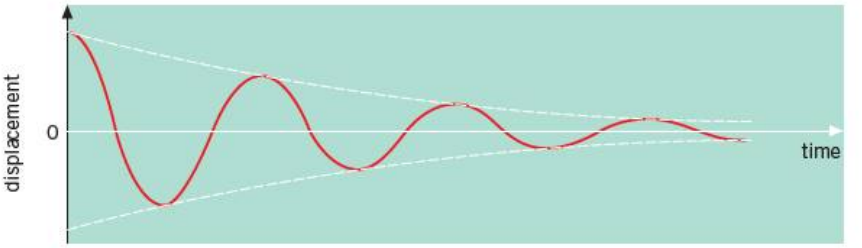


Figure 13.13 Lightly damped oscillations

Heavy damping causes the oscillations to die away more quickly. If the damping is increased further, then the system reaches **critical damping** point. Here the displacement decreases to zero in the shortest time, without any oscillation (Figure 13.14).

Any further increase in damping produces **overdamping**. The displacement decreases to zero in a longer time than for critical damping (Figure 13.14).

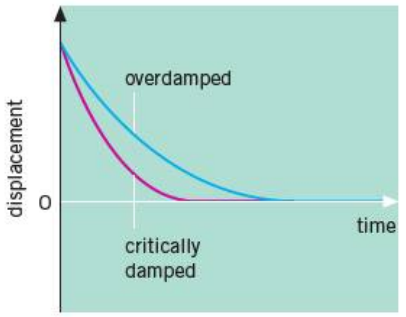


Figure 13.14 Critical damping and overdamping

Damping is often useful in an oscillating system. For example, vehicles have springs between the wheels and the frame to give a smoother and more comfortable ride (Figure 13.15). If there was no damping, a vehicle would move up and down for some time after hitting a bump in the road. Dampers (shock absorbers) are connected in parallel with the springs so that the suspension has critical damping and comes to rest in the shortest time possible. Dampers often work through hydraulic action. When the spring is compressed, a piston connected to the vehicle frame forces oil through a small hole in the piston, so that the energy of the oscillation is dissipated as thermal energy in the oil.



Figure 13.15 Vehicle suspension system showing springs and dampers

Many swing doors have a damping mechanism fitted to them. The purpose of the damper is so that the open door, when released, does not overshoot the closed position with the possibility of injuring someone approaching the door. Most door dampers operate in the overdamped mode.

Forced oscillations and resonance

When a vibrating body undergoes free (undamped) oscillations, it vibrates at its **natural frequency**. We met the idea of a natural frequency in Topic 15, when talking about stationary waves on strings. The natural frequency of such a system is the frequency of the first mode of vibration; that is, the fundamental frequency. A practical example is a guitar string, plucked at its centre, which oscillates at a particular frequency that depends on the speed of progressive waves on the string and the length of the string. The speed of progressive waves on the string depends on the mass per unit length of the string and the tension in the string.

Vibrating objects may have periodic forces acting on them. These periodic forces will make the object vibrate at the frequency of the applied force, rather than at the natural frequency of the system. The object is then said to be undergoing **forced vibrations**. Figure 13.16 illustrates apparatus which may be used to demonstrate the forced vibrations of a mass on a helical spring. The vibrator provides the forcing (driving) frequency and has a constant amplitude of vibration.

As the frequency of the vibrator is gradually increased from zero, the mass begins to oscillate. At first the amplitude of the oscillations is small, but it increases with increasing frequency. When the driving frequency equals the natural frequency of oscillation of the mass–spring system, the amplitude of the oscillations reaches a maximum. The frequency at which this occurs is called the **resonant frequency**, and **resonance** is said to occur.

www.iranino.co

IRANINOVIN

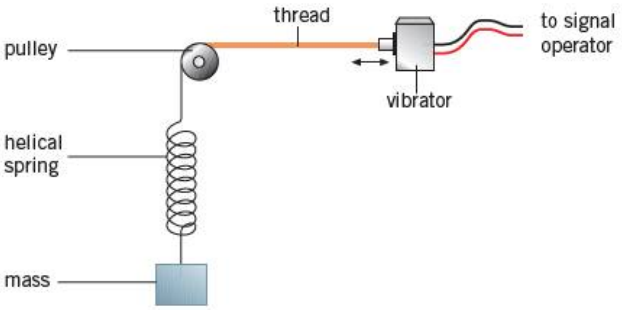


Figure 13.16 Demonstration of forced oscillations

Resonance occurs when the natural frequency of vibration of an object is equal to the driving frequency, giving a maximum amplitude of vibration.

If the driving frequency is increased further, the amplitude of oscillation of the mass decreases. The variation with driving frequency of the amplitude of vibration of the mass is illustrated in Figure 13.17. This graph is often called a **resonance curve**.

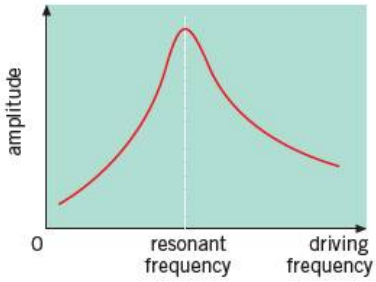


Figure 13.17 Resonance curve

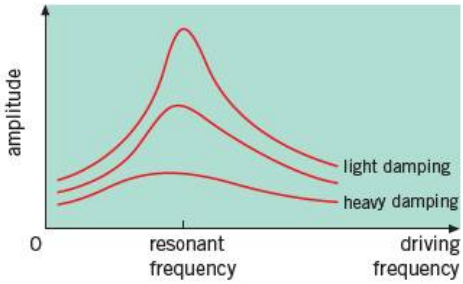


Figure 13.18 Effect of damping on the resonance curve

The effect of damping on the amplitude of forced oscillations can be investigated by attaching a light but stiff card to the mass in Figure 13.16. Movement of the card gives rise to air resistance and thus damping of the oscillations. The degree of damping may be varied by changing the area of the card. The effects of damping are illustrated in Figure 13.18. It can be seen that, as the degree of damping increases:

- the amplitude of oscillation at all frequencies is reduced
- the frequency at maximum amplitude shifts gradually towards lower frequencies
- the peak becomes flatter.

Barton's pendulums may also be used to demonstrate resonance and the effects of damping. The apparatus consists of a set of light pendulums, made (for example) from paper cones, and a more massive pendulum (the driver), all supported on a taut string. The arrangement is illustrated in Figure 13.19. The lighter pendulums have different lengths, but one has the same length as the driver. This has the same natural frequency as the driver and will, therefore, vibrate with the largest amplitude of all the pendulums (Figure 13.20).



Figure 13.20 Time-exposure photographs of Barton's pendulums with light damping, taken end-on. The longest arc, in the middle, is the driver.

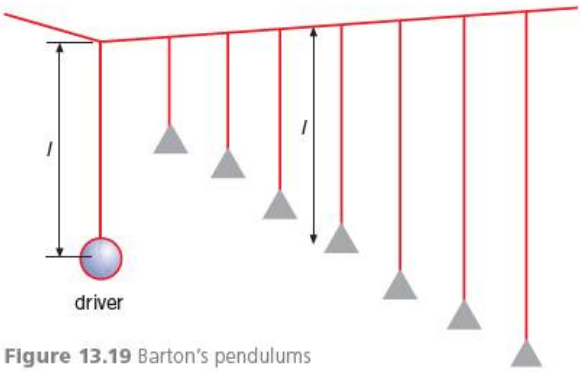


Figure 13.19 Barton's pendulums

www.iranino.co

IRANINOVIN



Figure 13.21 Pushing a child on a swing makes the swing go higher.

Adding weights to the paper cones reduces the effect of damping. With less damping, the amplitude of the resonant pendulum is much larger.

There are many examples of resonance in everyday life. One of the simplest is that of pushing a child on a swing (Figure 13.21). We push at the same frequency as the natural frequency of oscillation of the swing and child, so that the amplitude of the motion increases.

The operation of the engine of a vehicle causes a periodic force on the parts of the vehicle, which can cause them to resonate. For example, at particular frequencies of rotation of the engine, the mirrors may resonate. To prevent excessive vibration, the mountings of the mirrors provide damping.

Musical instruments rely on resonance to amplify the sound produced. The sound from a tuning fork is louder when it is held over a tube of just the right length, so that the column of air resonates. We met this phenomenon in Topic 15, in connection with the resonance tube method of measuring the speed of sound in air. Stringed instruments have a hollow wooden box with a hole under the strings which acts in a similar way. To amplify all notes from all of the strings, the sounding-box has to be a complex shape so that it resonates at many different frequencies.

A spectacular example of resonance that is often quoted is the failure of the first suspension bridge over the Tacoma Narrows in Washington State, USA. Wind caused the bridge to oscillate. It was used for months even though the roadway was oscillating with transverse vibrations. Approaching vehicles would appear, and then disappear, as the bridge deck vibrated up and down. One day strong winds set up twisting vibrations (Figure 13.22) and the amplitude of vibration increased due to resonance, until eventually the bridge collapsed. The driver of a car that was on the bridge managed to walk to safety before the collapse, although his dog could not be persuaded to leave the car.



Figure 13.22 The Tacoma Narrows bridge disaster

www.iranino.co

IRANINOVIN

Summary

Simple harmonic motion

- The period of an oscillation is the time taken to complete one oscillation.
- Frequency is the number of oscillations per unit time.
- Frequency f is related to period T by the expression $f = 1/T$
- The displacement of a particle is its distance from the equilibrium position.
- Amplitude is the maximum displacement.
- Simple harmonic motion (s.h.m.) is defined as the motion of a particle about a fixed point such that its acceleration a is proportional to its displacement x from the fixed point, and is directed towards the fixed point, $a \propto -x$ or $a = -\omega^2 x$
- The constant ω in the defining equation for simple harmonic motion is known as the angular frequency.
- For a particle oscillating in s.h.m. with frequency f, then $\omega = 2\pi f$
- Simple harmonic motion is described in terms of displacement x, amplitude x_0, frequency f, angular frequency ω by the following relations.

displacement:	$x = x_0 \sin \omega t$	or	$x = x_0 \cos \omega t$
velocity:	$v = x_0\omega \cos \omega t$	or	$v = -x_0\omega \sin \omega t$
acceleration:	$a = -x_0\omega^2 \sin \omega t$	or	$a = -x_0\omega^2 \cos \omega t$

- Remember that $\omega = 2\pi f$, and equations may appear in either form.

Energy of oscillations

- The kinetic energy E_k of a particle of mass m oscillating in simple harmonic motion with angular frequency ω and amplitude x_0 is $E_k = \frac{1}{2}m\omega^2(x_0^2 - x^2)$ where x is the displacement.
- The potential energy E_p of a particle of mass m oscillating in simple harmonic motion with angular frequency ω is $E_p = \frac{1}{2}m\omega^2x^2$ where x is the displacement.
- The total energy E_{tot} of a particle of mass m oscillating in simple harmonic motion with angular frequency ω and amplitude x_0 is
  $E_{tot} = \frac{1}{2}m\omega^2x_0^2$
- For a particle oscillating in simple harmonic motion
  $E_{tot} = E_k + E_p$
  and this expresses the law of conservation of energy.

Free and damped oscillations

- Free oscillations are oscillations where there are no resistive forces acting on the oscillating system.
- Damping is produced by resistive forces which dissipate the energy of the vibrating system.
- Light damping causes the amplitude of vibration of the oscillation to decrease gradually. Critical damping causes the displacement to be reduced to zero in the shortest time possible, without any oscillation of the object. Overdamping also causes an exponential reduction in displacement, but over a greater time than for critical damping.
- The natural frequency of vibration of an object is the frequency at which the object will vibrate when allowed to do so freely.
- Forced oscillations occur when a periodic driving force is applied to a system which is capable of vibration.
- Resonance occurs when the driving frequency on the system is equal to its natural frequency of vibration. The amplitude of vibration is a maximum at the resonant frequency.

www.iranino.co

IRANINOVIN

Examination style questions

1 A particle is oscillating in simple harmonic motion with period 4.5 ms and amplitude 3.0 cm.

At time $t = 0$, the particle is at the equilibrium position. Calculate, for this particle:

- **a** the frequency,
- **b** the angular frequency,
- **c** the maximum speed,
- **d** the magnitude of the maximum acceleration,
- **e** the speed at time $t = 1.0$ ms.

2 A particle is oscillating in simple harmonic motion with frequency 50 Hz and amplitude 15 mm. Calculate the speed when the displacement from the equilibrium position is 12 mm.

3 Geologists use the fact that the period of oscillation of a simple pendulum depends on the acceleration of free fall to map variations of g. A geologist determines the frequency of oscillation of a test pendulum of effective length 515.6 mm to be 0.6948 Hz. Calculate the acceleration of free fall at this locality.

4 A spring stretches by 85 mm when a mass of 50 g is hung from it. The spring is then stretched a further distance of 15 mm from the equilibrium position, and the mass is released at time $t = 0$. Calculate:

- **a** the spring constant,
- **b** the amplitude of the oscillations,
- **c** the period,
- **d** the displacement at time $t = 0.20$ s. ($g = 9.8$ m s^{-2})

5 One particle oscillating in simple harmonic motion has ten times the total energy of another particle, but the frequencies and masses are the same. Calculate the ratio of the amplitudes of the two motions.

6 **a** Calculate the displacement in terms of x_0, expressed as a fraction of the amplitude x_0, of a particle moving in simple harmonic motion with a speed equal to half the maximum value.

b Calculate the displacement at which the energy of the particle has equal amounts of kinetic and of potential energy.

7 The apparatus of Fig. 13.16 is used to demonstrate forced vibrations and resonance. A 50 g mass is suspended from the spring, which has a spring constant of 7.9 N m^{-1}.

- **a** Calculate the resonant frequency f_0 of the system.
- **b** A student suggests that resonance should also be observed at a frequency of $2f_0$. Discuss this suggestion.

8 **a** Define simple harmonic motion.

b Fig. 13.23a illustrates a U-tube of uniform cross-sectional area A containing liquid of density ρ. The total length of the liquid column is L. When the liquid is displaced by an amount Δx from its equilibrium position (see Fig. 13.23b), it oscillates with simple harmonic motion.

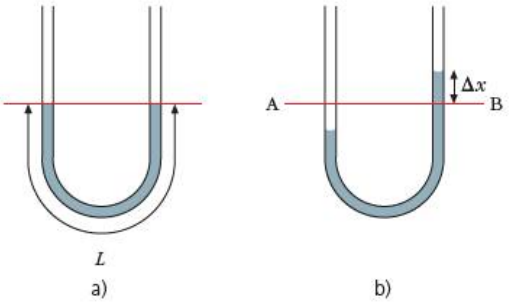


Fig. 13.23

The weight of liquid above AB in Fig. 13.23b provides the restoring force.

- **i** Write down an expression for the restoring force.
- **ii** Write down an expression for the acceleration of the liquid column caused by this force.
- **iii** Explain how this fulfils the condition for simple harmonic motion.
- **iv** Write down an expression for the frequency of the oscillations.

9 A 'baby bouncer' consists of a harness attached to a rubber cord. A baby of mass 6.5 kg is placed gently in the harness and the cord extends by 0.40 m. The baby is then pulled down another 0.10 m and, when released, begins to move with simple harmonic motion. Calculate

- **a** the period of the motion,
- **b** the maximum force on the baby.

(acceleration of free fall $g = 9.8$ m s^{-2})

10 A simple model of a hydrogen molecule assumes that it consists of two oscillating hydrogen atoms connected by a spring of spring constant 1.1×10^3 N m^{-1}.

- **a** The mass of a hydrogen atom is 1.7×10^{-27} kg. Calculate the frequency of oscillation of the hydrogen molecule.
- **b** Explain why light of wavelength 2.3 μm would be strongly absorbed by this model of the hydrogen molecule. (speed of light $c = 3.0 \times 10^8$ m s^{-1})

11 A ball is held between two fixed points A and B by means of two stretched springs, as shown in Fig. 13.24.

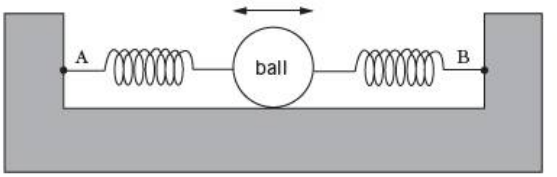


Fig. 13.24

The ball is free to oscillate along the straight line AB. The springs remain stretched and the motion of the ball is simple harmonic.

The variation with time t of the displacement x of the ball from its equilibrium position is shown in Fig. 13.25.

www.iranino.co

IRANINOVIN

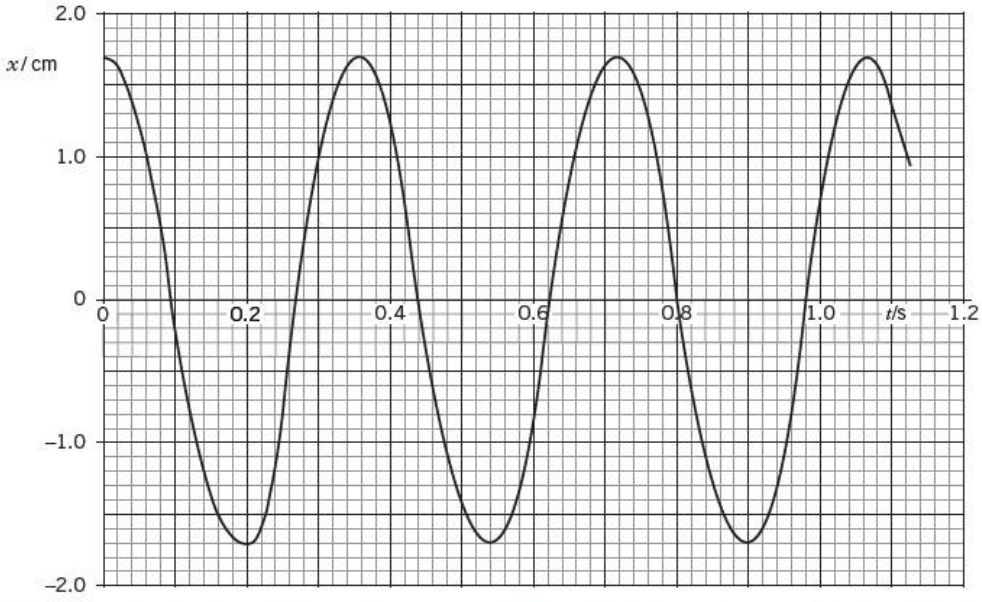


Fig. 13.25

a **i** Use Fig. 13.25 to determine, for the oscillations of the ball,

- the amplitude *[1]*
- the frequency. *[2]*

ii Show that the maximum acceleration of the ball is $5.2\,m\,s^{-2}$. *[2]*

b Use your answers in **a** to plot, on a copy of Fig. 13.26, the variation with displacement x of the acceleration a of the ball. *[2]*

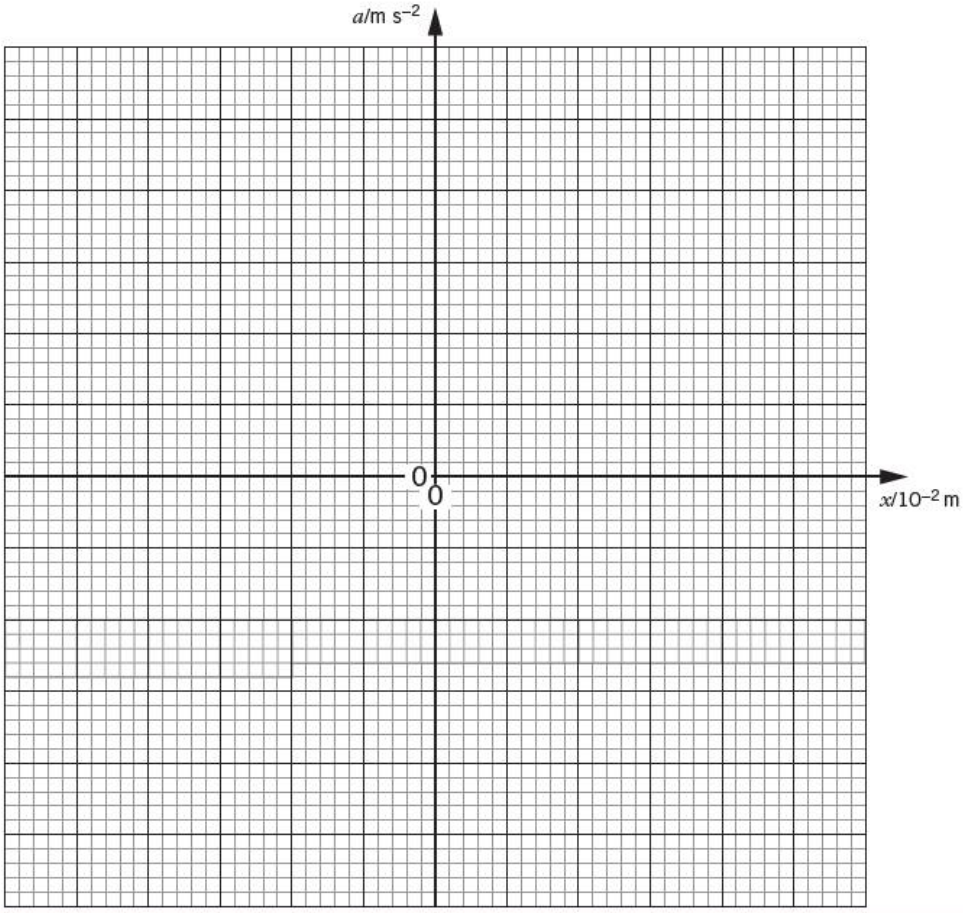


Fig. 13.26

www.iranino.com

IRANINOVIN

c Calculate the displacement of the ball at which its kinetic energy is equal to one half of the maximum kinetic energy. [3]

Cambridge International AS and A Level Physics, 9702/43 May/June 2013 Q 3

12 A long strip of springy steel is clamped at one end so that the strip is vertical. A mass of 65 g is attached to the free end of the strip, as shown in Fig. 13.27.

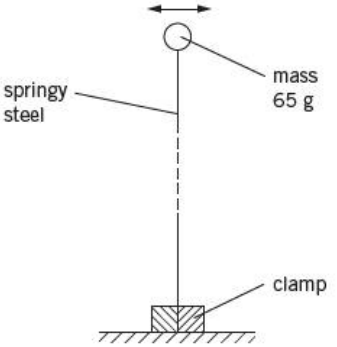


Fig. 13.27

The mass is pulled to one side and then released. The variation with time t of the horizontal displacement of the mass is shown in Fig. 13.28.

The mass undergoes damped simple harmonic motion.

a i Explain what is meant by *damping*. [2]

ii Suggest, with a reason, whether the damping is light, critical or heavy. [2]

b i Use Fig.13.28 to determine the frequency of vibration of the mass. [1]

ii Hence show that the initial energy stored in the steel strip before the mass is released is approximately 3.2 mJ. [2]

c After eight complete oscillations of the mass, the amplitude of vibration is reduced from 1.5 cm to 1.1 cm. State and explain whether, after a further eight complete oscillations, the amplitude will be 0.7 cm. [2]

Cambridge International AS and A Level Physics, 9702/42 May/June 2010 Q 2

13 A small metal ball is suspended from a fixed point by means of a string, as shown in Fig. 13.29.

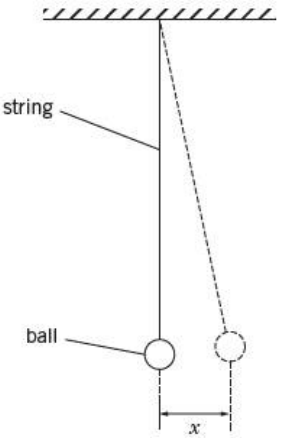


Fig. 13.29

The ball is pulled a small distance to one side and then released. The variation with time t of the horizontal displacement x of the ball is shown in Fig. 13.30.

The motion of the ball is simple harmonic.

a Use data from Fig. 13.30 to determine the horizontal acceleration of the ball for a displacement x of 2.0 cm. [3]

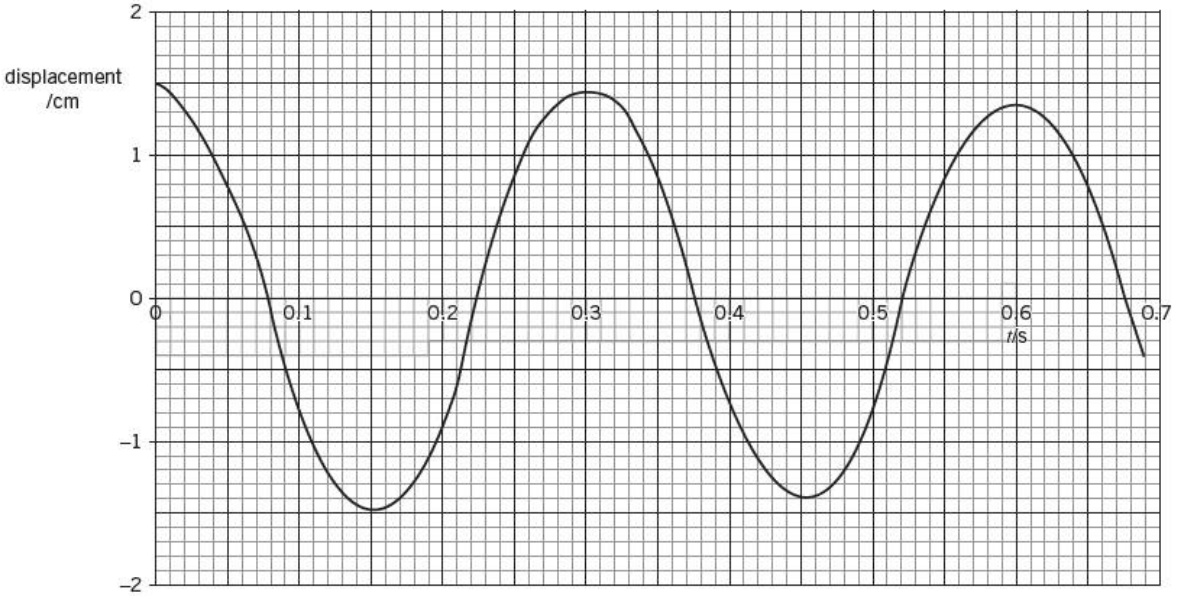


Fig. 13.28

www.iranino.co

IRANINOVIN

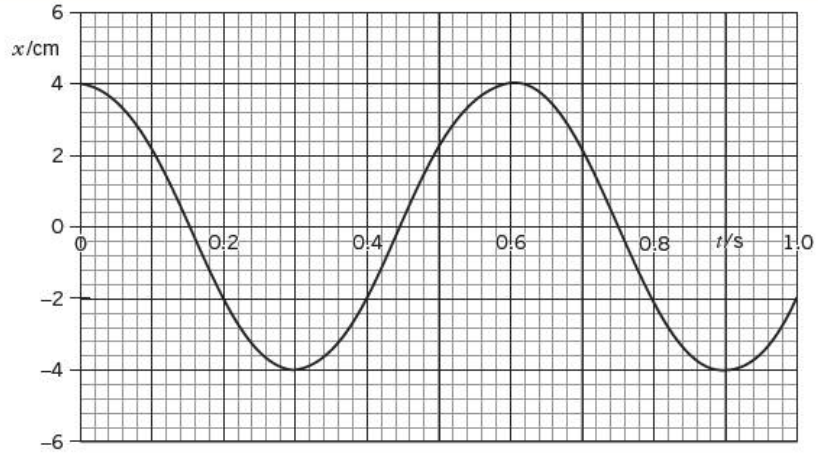


Fig. 13.30

b The maximum kinetic energy of the ball is E_K.

On a copy of the axes of Fig. 13.31, sketch a graph to show the variation with time t of the kinetic energy of the ball for the first 1.0 s of its motion. *[3]*

Cambridge International AS and A Level Physics, 9702/41 May/June 2012 Q 4

14 A student sets up the apparatus illustrated in Fig. 13.32 in order to investigate the oscillations of a metal cube suspended on a spring.

The amplitude of the vibrations produced by the oscillator is constant.

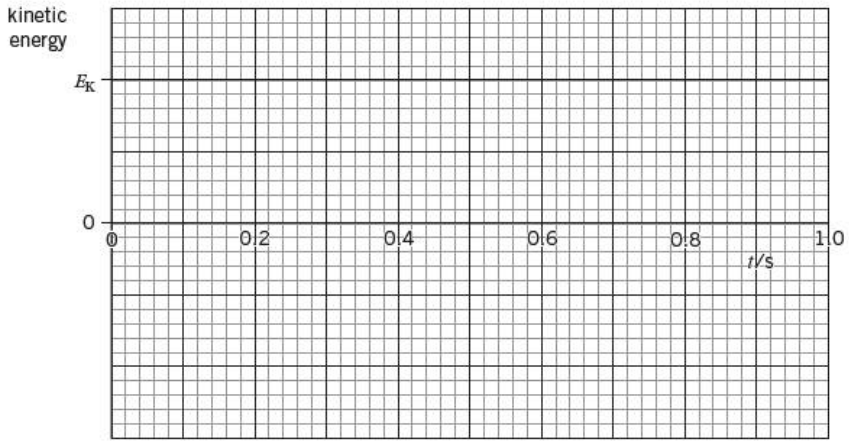


Fig. 13.31

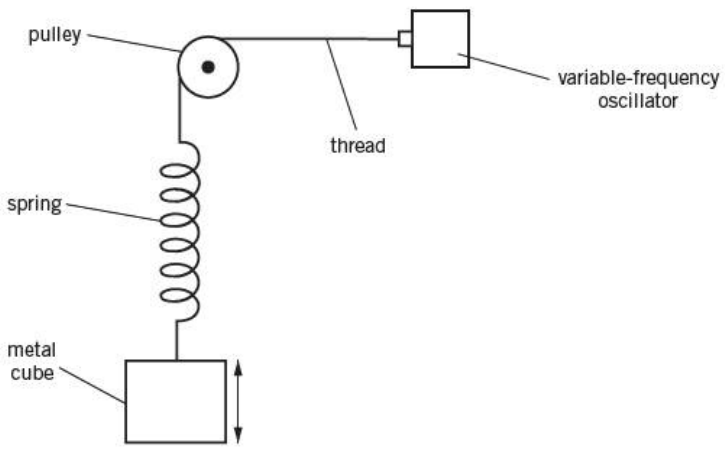


Fig. 13.32

www.iranino.co

IRANINOVIN

The variation with frequency of the amplitude of the oscillations of the metal cube is shown in Fig. 13.33.

a i State the phenomenon illustrated in Fig. 13.33. *[1]*

ii For the maximum amplitude of vibration, state the magnitudes of the amplitude and the frequency. *[1]*

b The oscillations of the metal cube of mass 150 g may be assumed to be simple harmonic.

Use your answers in **a ii** to determine, for the metal cube,

i its maximum acceleration, *[3]*

ii the maximum resultant force on the cube. *[2]*

c Some very light feathers are attached to the top surface of the cube so that the feathers extend outwards, beyond the vertical sides of the cube.

The investigation is now repeated.

On a copy of Fig. 13.33, draw a line to show the new variation with frequency of the amplitude of vibration for frequencies between 2 Hz and 10 Hz. *[2]*

Cambridge International AS and A Level Physics, 9702/42 Oct/Nov 2010 Q 3

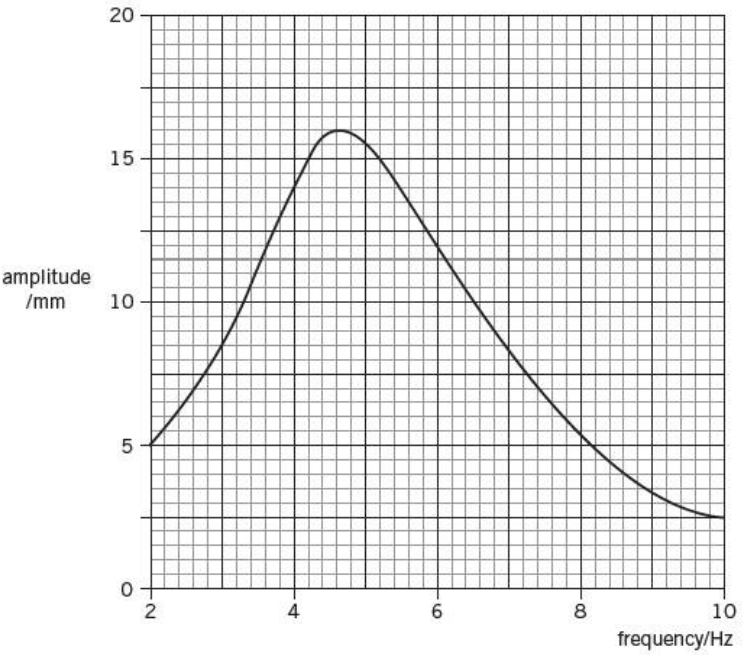


Fig. 13.33

www.iranino.co

IRANINOVIN

14 Ultrasound

By the end of this topic, you will be able to:

14.6 **(a)** explain the principles of the generation and detection of ultrasonic waves using piezo-electric transducers

(b) explain the main principles behind the use of ultrasound to obtain diagnostic information about internal structures

(c) understand the meaning of specific acoustic impedance and its importance to the intensity reflection coefficient at a boundary

(d) recall and solve problems by using the equation $I = I_0 e^{-\mu x}$ for the attenuation of ultrasound in matter

Starting points

- The wave nature of sound.
- An understanding of the terms wavelength, frequency, speed, amplitude and intensity.

14.6 The generation and use of ultrasound

Ultrasound waves may be generated using a **piezo-electric transducer**. A transducer is the name given to any device that converts energy from one form to another. In this case, electrical energy is converted into ultrasound energy by means of a piezo-electric crystal such as quartz.

The structure of quartz is made up of a large number of tetrahedral silicate units, as shown in Figure 14.1. These units build up to form a crystal of quartz that can be represented, in two dimensions, as shown in Figure 14.2.

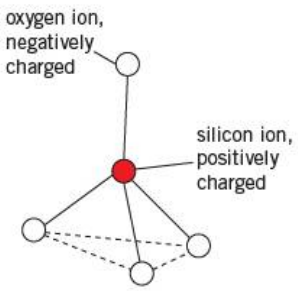


Figure 14.1 Tetrahedral silicate unit

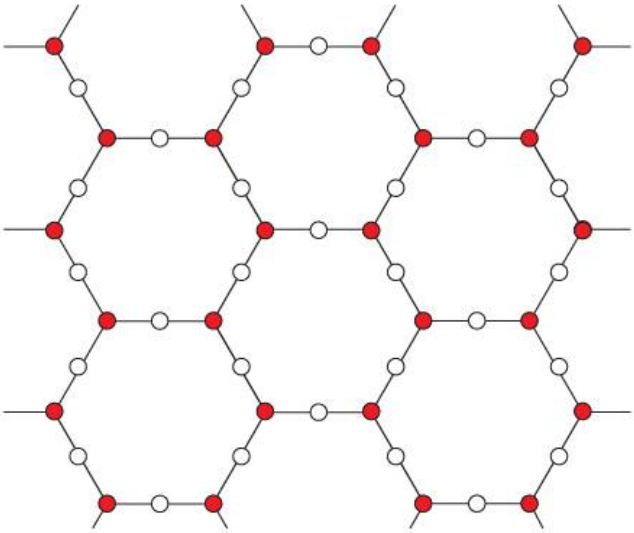


Figure 14.2 Two-dimensional representation of a quartz crystal

www.iranino.co

IRANINOVIN

When the crystal is unstressed, the centres of charge of the positive and the negative ions in any one unit coincide, as shown in Figure 14.3a. Electrodes may be formed on opposite sides of the crystal by depositing silver on its surfaces. When a potential difference is applied between the electrodes, an electric field is set up in the crystal. This field causes forces to act on the ions. The oxygen ions are negatively charged and the silicon ions have a positive charge. The ions are not held rigidly in position and, as a result, they will be displaced slightly when the electric field is applied across the crystal. The positive ions will be attracted towards the negative electrode and the negative ions will be attracted to the positive electrode. Dependent on the direction of the electric field, the crystal will become slightly thinner (Figure 14.3b) or slightly thicker (Figure 14.3c).

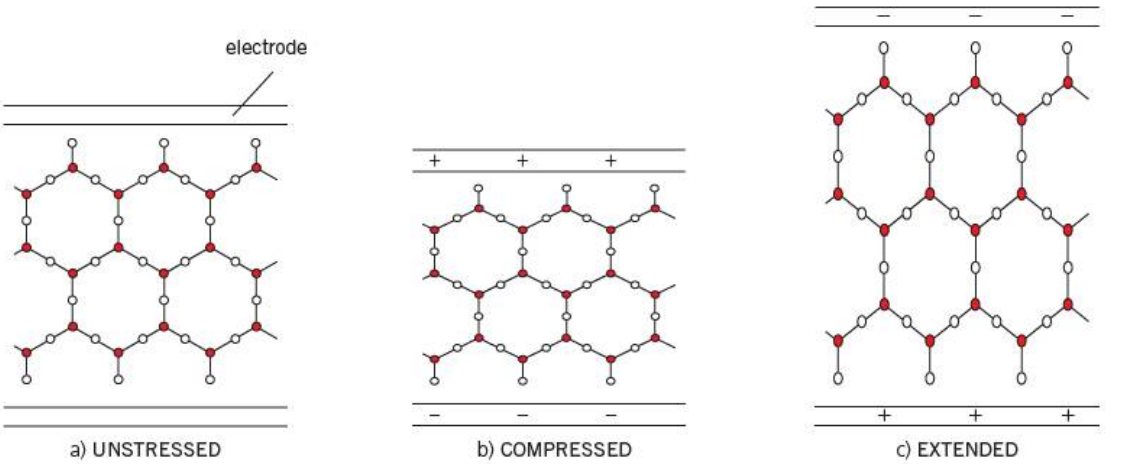


Figure 14.3 The effect of an electric field on a quartz crystal

An alternating voltage applied across the electrodes causes the crystal to vibrate with a frequency equal to that of the applied voltage. These oscillations are likely to have a small amplitude. However, if the frequency of the applied voltage is equal to the natural frequency of vibration of the crystal, resonance will occur (see Topic 13) and the amplitude of vibration will be a maximum. The dimensions of the crystal can be such that the oscillations are in the ultrasound range of frequencies (greater than about 20 kHz). These oscillations will give rise to ultrasound waves in any medium surrounding the crystal.

If a stress is applied to an uncharged quartz crystal, the forces involved will alter the positions of the positive and the negative ions, creating a potential difference across the crystal. Therefore, if an ultrasound wave is incident on the crystal, the pressure variations in the wave will give rise to voltage variations across the crystal. An ultrasound transducer may, therefore, also be used as a detector (or receiver).

A simplified diagram of a piezo-electric transducer/receiver is shown in Figure 14.4.

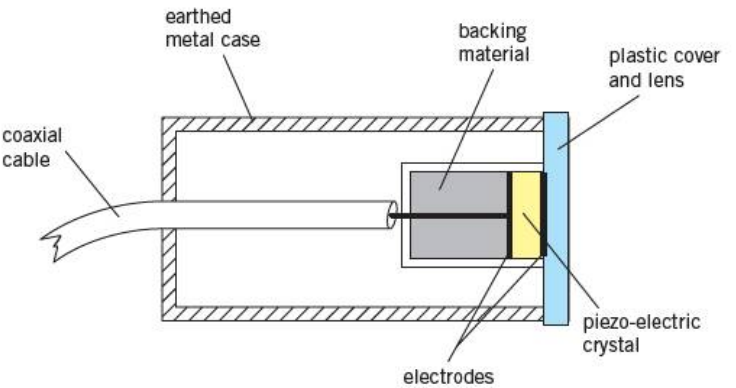


Figure 14.4 Piezo-electric transducer/receiver

www.iranino.co

IRANINOVIN

A transducer such as this is able to produce and detect ultrasound in the megahertz frequency range, which is typical of the frequency range used in medical diagnosis.

The reflection and absorption of ultrasound

Ultrasound is typical of many types of wave in that, when it is incident on a boundary between two media, some of the wave power is reflected and some is transmitted. This is illustrated in Figure 14.5.

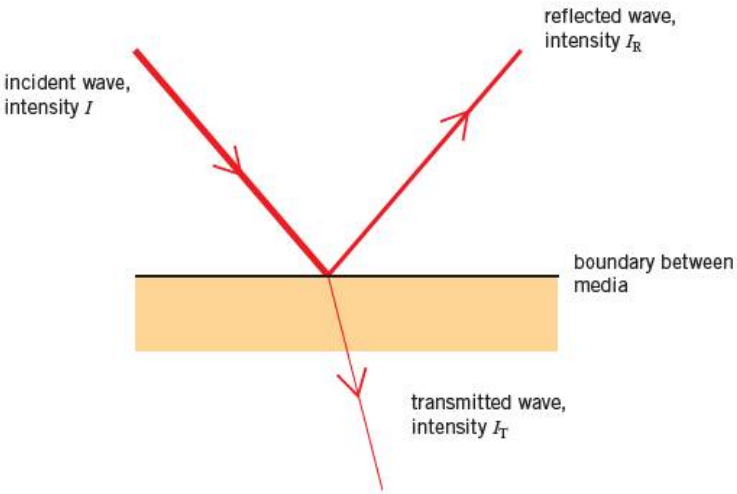


Figure 14.5 The reflection and transmission of a wave at a boundary

For a wave of incident intensity I, reflected intensity I_R and transmitted intensity I_T, by conservation of energy,

$$I = I_R + I_T$$

Although, for a beam of constant intensity, the sum of the reflected and transmitted intensities is constant, their relative magnitudes depend not only on the angle of incidence of the beam on the boundary but also on the media themselves. The relative magnitudes of I_R and I_T are quantified by reference to the **specific acoustic impedance** Z of each of the media. This is defined as the product of the density ρ of the medium and the speed c of the wave in the medium. That is,

$$Z = \rho c$$

For a wave incident normally on a boundary between two media having specific acoustic impedances of Z_1 and Z_2, the ratio of the reflected intensity I_R to the incident intensity I is given by

$$\frac{I_R}{I} = \frac{(Z_2 - Z_1)^2}{(Z_2 + Z_1)^2}$$

The ratio I_R/I is known as the **intensity reflection coefficient** for the boundary and is given the symbol α. As the above equation shows, α depends on the difference between the specific acoustic impedances of the two media. Some typical values of specific acoustic impedance are given in Table 14.1, together with the approximate speed of ultrasound in the medium.

www.iranino.co

IRANINOVIN

Table 14.1 Values of speed of ultrasound and specific acoustic impedance for some media

medium	speed/m s^{-1}	specific acoustic impedance/kg m^{-2} s^{-1}
air	330	430
blood	1600	1.6×10^6
bone	4100	5.6×10^6–7.8×10^6
fat	1500	1.4×10^6
muscle	1600	1.7×10^6
soft tissue	1600	1.6×10^6
water	1500	1.5×10^6

Example

Using data from Table 14.1, calculate the intensity reflection coefficient for a parallel beam of ultrasound incident normally on the boundary between

(a) air and soft tissue,

(b) muscle and bone that has a specific acoustic impedance of 6.5×10^6 kg m^{-2} s^{-1}.

(a) $\alpha = (Z_2 - Z_1)^2/(Z_2 + Z_1)^2$

$= (1.6 \times 10^6 - 430)^2/(1.6 \times 10^6 + 430)^2$

$=$ **0.999**

(b) $\alpha = (6.5 \times 10^6 - 1.7 \times 10^6)^2/(6.5 \times 10^6 + 1.7 \times 10^6)^2$

$=$ **0.34**

Now it's your turn

1 Using data from Table 14.1,

(a) suggest why, although the speed of ultrasound in blood and muscle is approximately the same, the specific acoustic impedance is different.

(b) calculate the intensity reflection coefficient for a parallel beam of ultrasound incident normally on the boundary between fat and muscle.

It can be seen that the intensity reflection coefficient for a boundary between air and soft tissue is approximately equal to unity. This means that, when ultrasound is incident on the body, very little is transmitted into the body. In order that ultrasound may be transmitted into the body and also that the ultrasound may return to the transducer, it is important that there is no air between the transducer and the skin (soft tissue). This is achieved by means of a water-based jelly. This jelly has a specific acoustic impedance of approximately 1.5×10^6 kg m^{-2} s^{-1}.

Once the ultrasound wave is within the medium, the intensity of the wave will be reduced by absorption of energy as it passes through the medium. The medium is heated. In fact, the heating effect produced by ultrasound of appropriate frequencies is used in physiotherapy to assist recovery from sprains and similar injuries.

For a parallel beam, this absorption is approximately exponential (as with X-rays – see Topic 25) and is shown in Figure 14.6.

For such a beam of ultrasound incident normally on a medium of thickness x, the transmitted intensity I is related to the incident intensity I_0 by the expression

$$I = I_0 e^{-kx} \text{ or } I = I_0 \exp(-kx)$$

where k is a constant for the medium known as the linear absorption (attenuation) coefficient. The coefficient k depends not only on the medium itself but also on the frequency of the ultrasound. Some values of the linear absorption (attenuation) coefficient are shown in Table 14.2.

www.iranino.co

IRANINOVIN

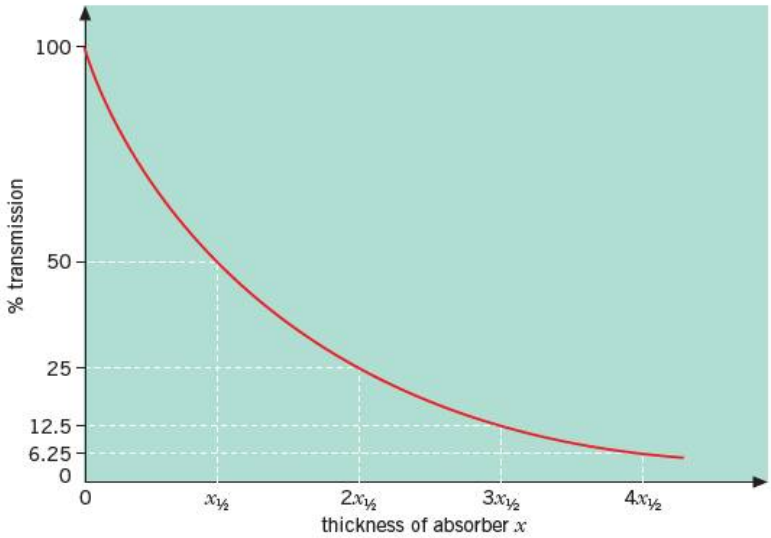


Figure 14.6 The percentage transmission of ultrasound in a medium

Note that the expression for the change in the transmitted intensity applies only to a *parallel* beam. If the beam is divergent, then the intensity would decrease, without any absorption by the medium.

Table 14.2 Some values of linear absorption (attenuation) coefficient for ultrasound

medium	linear absorption (attenuation) coefficient/cm^{-1}
air	1.2
bone	0.13
muscle	0.23
water	0.0002

Example

A parallel beam of ultrasound is incident on the surface of a muscle and passes through a thickness of 3.5 cm of the muscle. It is then reflected at the surface of a bone and returns through the muscle to its surface. Using data from Tables 14.1 and 14.2, calculate the fraction of the incident intensity that arrives back at the surface of the muscle.

The beam passes through a total thickness of 7.0 cm of muscle. For the attenuation in the muscle,

$$I = I_0 \exp(-0.23 \times 7.0) = 0.20 I_0$$

Fraction reflected at muscle–bone interface $= (6.5 \times 10^6 - 1.7 \times 10^6)^2/(6.5 \times 10^6 + 1.7 \times 10^6)^2$

$$= 0.34$$

Fraction received back at surface $= 0.20 \times 0.34 =$ **0.068**

Now it's your turn

2 A parallel beam of ultrasound passes through a thickness of 4.0 cm of muscle. It is then incident normally on a bone having a specific acoustic impedance of 6.4×10^6 kg m^{-2} s^{-1}. The bone is 1.5 cm thick. Using data from Table 14.2, calculate the fraction of the incident intensity that is transmitted through the muscle and bone.

Obtaining diagnostic information using ultrasound

The ultrasound transducer is placed on the skin, with the water-based jelly excluding any air between the transducer and the skin (Figure 14.7).

Short pulses of ultrasound are transmitted into the body where they are partly reflected and partly transmitted at the boundaries between media in the body such as fat–muscle and muscle–bone. The reflected pulses return to the transducer where they are detected and converted into voltage pulses. These voltage pulses can be amplified and processed by electronic circuits such that the output of the circuits may be displayed on a screen as in, for example, a cathode-ray oscilloscope.

Pulses of ultrasound are necessary so that the reflected ultrasound pulses can be detected in the time intervals between the transmitted pulses. The time between the transmission of a pulse and its receipt back at the transducer gives information as to the distance of the boundary from the transducer. The intensity of the reflected pulse



Figure 14.7 Ultrasound diagnosis

www.iranino.co

IRANINOVIN

gives information as to the nature of the boundary. Two techniques are in common use for the display of an ultrasound scan.

In an **A-scan**, a short pulse of ultrasound is transmitted into the body through the coupling medium (the water-based jelly). At each boundary between media, some of the energy of the pulse is reflected and some is transmitted. The transducer detects the reflected pulses as it now acts as the receiver. The signal is amplified and displayed on a cathode-ray oscilloscope (c.r.o.). Reflected pulses (echoes) received at the transducer from deeper in the body tend to have lower intensity than those reflected from boundaries near the skin. This is caused not only by absorption of wave energy in the various media but also, on the return of the reflected pulse to the transducer, some of the energy of the pulse will again be reflected at intervening boundaries. To allow for this, echoes received later at the transducer are amplified more than those received earlier. A vertical line is observed on the screen of the c.r.o. corresponding to the detection of each reflected pulse. The time-base of the c.r.o. is calibrated so that, knowing the speed of the ultrasound wave in each medium, the distance between boundaries can be determined. An example of an A-scan is illustrated in Figure 14.8.



Figure 14.8 An A-scan

A **B-scan** consists of a series of A-scans, all taken from different angles so that, on the screen of the c.r.o., a two-dimensional image is formed. Such an image is shown in Figure 14.9.

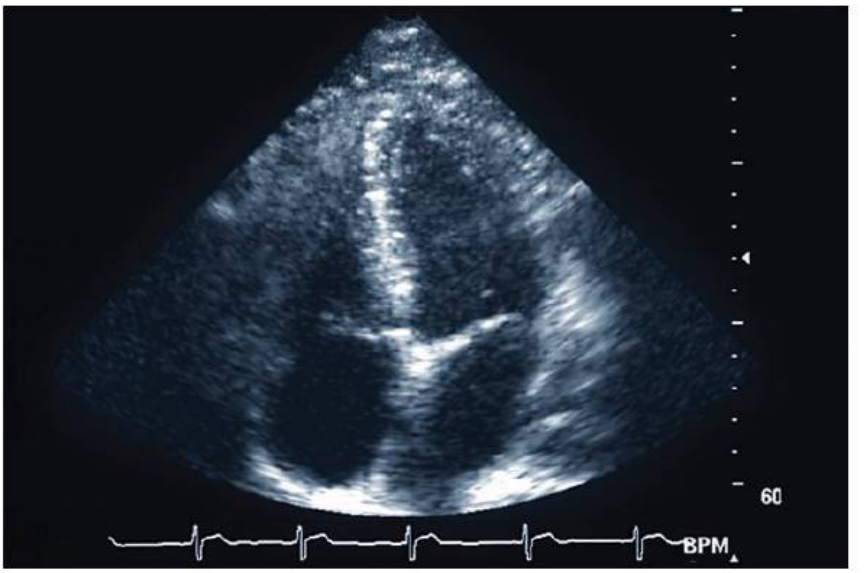


Figure 14.9 An image of a healthy heart produced from a B-scan

www.iranino.co

IRANINOVIN

The ultrasound probe for a B-scan does not consist of a single crystal. Rather, it has an array of small crystals, each one at a slightly different angle to its neighbours. The separate signals received from each of the crystals in the probe are processed. Each reflected pulse is shown on the screen of the c.r.o. as a bright spot in the direction of orientation of the particular crystal that gave rise to the signal. The pattern of spots builds up to form a two-dimensional image representing the positions of the boundaries within the body. The image may be either viewed immediately or photographed or stored in a computer memory.

The main advantage of ultrasound scanning compared to X-ray diagnosis is that the health risk to both the patient and to the operator is very much less. Also, ultrasound equipment is much more portable and is relatively simple to use.

Higher frequency ultrasound enables greater resolution to be obtained since the wavelength will be shorter and there will be less diffraction around small features. That is, more detail can be seen. Furthermore, as modern techniques allow for the detection of very low intensity reflected pulses, boundaries between tissues where there is little change in acoustic impedance can be detected.

Summary

- Ultrasound may be generated and detected by piezo-electric crystals.
- Ultrasound images are formed as a result of the detection and processing of ultrasound pulses that have been reflected from tissue boundaries.
- The acoustic impedance Z of a medium is ρc.
- The intensity reflection coefficient is $(Z_2 - Z_1)^2/(Z_2 + Z_1)^2$
- Two-dimensional scans may be obtained using a generator/detector consisting of many separate crystals all at different angles of orientation.

Examination style questions

1 Explain why, when obtaining an ultrasound scan,

a the ultrasound is pulsed and is not continuous,

b the reflected signal received from deeper in the body is amplified more than that received from near the skin.

2 The specific acoustic impedance of fat, muscle and bone are 1.4×10^6 kg m^{-2} s^{-1}, 1.6×10^6 kg m^{-2} s^{-1} and 6.5×10^6 kg m^{-2} s^{-1} respectively. The linear absorption coefficients in fat and in muscle are 0.24 cm^{-1} and 0.23 cm^{-1} respectively.

A parallel beam of ultrasound of intensity I is incident on the layer of fat. Discuss quantitatively, in terms of I, the reflection and the transmission of the beam of ultrasound as it passes through the layer of fat of thickness 4.0 mm, into the muscle of thickness 43.5 mm and finally into the bone.

3 a i State what is meant by the *acoustic impedance* of a medium. *[1]*

ii Data for some media are given in Table 14.3.

Table 14.3

medium	speed of ultrasound/m s^{-1}	acoustic impedance/ kg m^{-2} s^{-1}
air	330	4.3×10^2
gel	1500	1.5×10^6
soft tissue	1600	1.6×10^6
bone	4100	7.0×10^6

Use data from Table 14.3 to calculate a value for the density of bone. *[1]*

b A parallel beam of ultrasound has intensity I. It is incident at right angles to a boundary between two media, as shown in Fig. 14.10.

www.iranino.co

IRANINOVIN

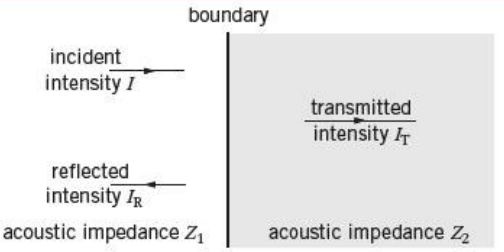


Fig. 14.10

The media have acoustic impedances of Z_1 and Z_2. The transmitted intensity of the ultrasound beam is I_T and the reflected intensity is I_R.

i State the relation between I, I_T and I_R. [1]

ii The reflection coefficient α is given by the expression

$$\alpha = \frac{(Z_2 - Z_1)^2}{(Z_2 + Z_1)^2}$$

Use data from Table 14.3 to determine the reflection coefficient α for a boundary between

- gel and soft tissue, [2]
- air and soft tissue. [1]

c By reference to your answers in **b ii**, explain the use of a gel on the surface of skin during ultrasound diagnosis. [3]

Cambridge International AS and A Level Physics, 9702/42 Oct/Nov 2010 Q 10

4 a By reference to ultrasound waves, state what is meant by *acoustic impedance*. [2]

b An ultrasound wave is incident on the boundary between two media. The acoustic impedances of the two media are Z_1 and Z_2, as illustrated in Fig. 14.11.

Explain the importance of the difference between Z_1 and Z_2 for the transmission of ultrasound across the boundary. [3]

c Ultrasound frequencies as high as 10 MHz are used in medical diagnosis.

State and explain one advantage of the use of high-frequency ultrasound compared with lower-frequency ultrasound. [2]

Cambridge International AS and A Level Physics, 9702/42 May/June 2013 Q 10

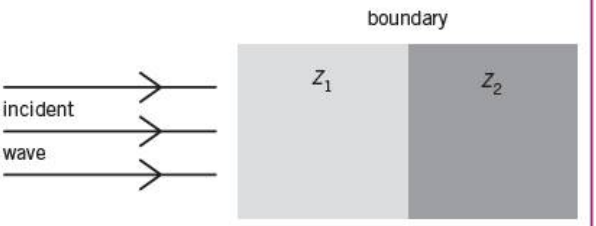


Fig. 14.11

www.iranino.co

IRANINOVIN

16 Communication

By the end of this topic, you will be able to:

16.1 **(a)** appreciate that information may be carried by a number of different channels, including wire-pairs, coaxial cables, radio and microwave links, and optic fibres

16.2 **(a)** understand the term modulation and be able to distinguish between amplitude modulation (AM) and frequency modulation (FM)

(b) recall that a carrier wave, amplitude modulated by a single audio frequency, is equivalent to the carrier wave frequency together with two sideband frequencies

(c) understand the term bandwidth

(d) recall the frequencies and wavelengths used in different channels of communication

(e) demonstrate an awareness of the relative advantages of AM and FM transmissions

16.3 **(a)** recall the advantages of the transmission of data in digital form compared with the transmission of data in analogue form

(b) understand that the digital transmission of speech or music involves analogue-to-digital conversion (ADC) on transmission and digital-to-analogue conversion (DAC) on reception

(c) understand the effect of the sampling rate and the number of bits in each sample on the reproduction of an input signal

16.4 **(a)** discuss the relative advantages and disadvantages of channels of communication in terms of available bandwidth, noise, cross-linking, security, signal attenuation, repeaters and regeneration

(b) recall the relative merits of both geostationary and polar-orbiting satellites for communicating information

16.5 **(a)** understand and use signal attenuation expressed in dB and dB per unit length

(b) recall and use the expression number of dB = $10 \lg \left(\frac{P_1}{P_2}\right)$, for the ratio of two powers

Starting points

- Electromagnetic waves are transverse waves that all have the same speed in a vacuum.
- The electromagnetic spectrum can be divided into different frequency or wavelength regions, each having different properties.
- Communication is the means by which information is transferred from one place to another.

16.1 Communication channels

For communication to take place, information is carried from one place to another. The information is carried in the form of a signal which may be, for example, changes in the frequency or amplitude of a wave or changes in the intensity of light. The signal may be transferred in various ways using different channels of communication including wire-pairs, coaxial cables, radio and microwave links, and optic fibres.

Wire-pairs

Wire-pairs provide a very simple link between a transmitter of information and the receiver. In the early days of electrical communication using Morse code, the transmitter and receiver were connected directly to one another by means of two copper wires (or one copper wire and an 'earth return'). In modern communication systems, wire-pairs are used mainly for short-distance communication at relatively low

www.iranino.co

IRANINOVIN

frequencies, for example linking telephones to the nearest exchange. In the simplest of applications, they are used for linking a door bell in a house to the switch outside the door.

Coaxial cables

The construction of a coaxial cable is shown in Figure 16.1.

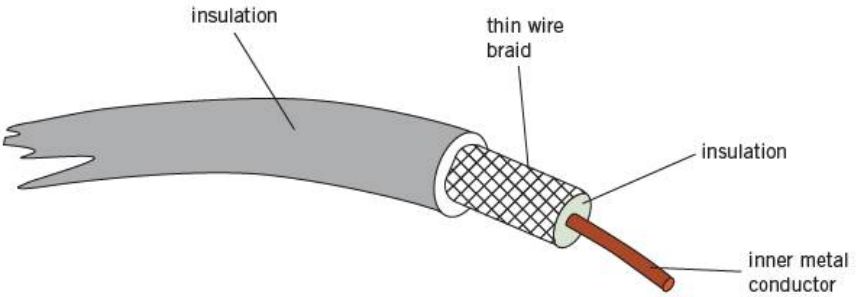


Figure 16.1 Coaxial cable

It consists, basically, of two wire conductors. Any electrical signal is transmitted along a central inner conductor that is covered by an insulator. The second conductor is in the form of thin wire braid that completely surrounds the insulator. This wire braid acts as the 'return' for the signal and is earthed. The braid is covered by a protective layer of insulation.

The earthed outer braiding shields the inner conductor from external interference, for example, other stray electrical signals. Consequently, the signals in coaxial cables suffer far less from distortion than wire-pairs and provide better security. Coaxial cables are used, for example, to connect an aerial to a television receiver.

Radio waves

Alternating current in a wire acts as an aerial. Energy is radiated from the aerial in the form of electromagnetic waves. These waves travel outwards from the aerial with the speed of light. Electromagnetic waves in the frequency range 30 kHz to 3 GHz are generally referred to as radio waves.

The first radio waves to be used for communication were of very low frequency and very long wavelength. The radio waves were switched on and off so that communication was by Morse code. Later, the use of higher frequencies and further developments (including FM broadcasts and the use of different carrier frequencies) enabled higher-quality communication and also more radio stations to operate in the same area.

The choice of aerial for broadcasting determines whether the radio waves are emitted in all directions (for broadcasting to an area) or in one direction only (for point-to-point communication). Similarly, for the receiving of radio signals, the choice of aerial is determined by whether the signal from one direction, or all directions, is to be received. Aerials with dish reflectors enable the radio waves to be transmitted as a parallel beam.

The intensity of the radio waves will always be reduced (attenuated) as the distance from the transmitter increases. The degree of attenuation depends on the frequency of the waves.

Microwaves

Microwaves are electromagnetic waves that can be considered alongside other radio waves. They are in the frequency range 3 GHz to 30 GHz and are generally used for point-to-point communication since, for use on Earth, the range of the transmissions is limited to line-of-sight.

Reflecting parabolic dishes are used so that the transmission is in the form of a parallel beam and then as much wave power as possible can be focused on to the receiving aerial, as shown in Figure 16.2.

www.iranino.co

IRANINOVIN

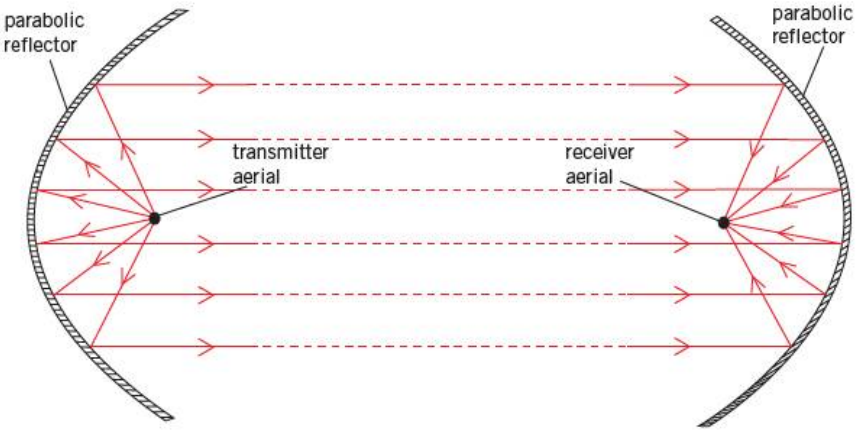


Figure 16.2 The use of parabolic reflectors with microwaves



Figure 16.3 Optic fibres making up an optic cable

It should be remembered that the reflecting parabolic dish is not the aerial. The aerial is found at the focus of the reflecting dish.

Optic fibres

An optic fibre consists of a fine strand of very pure glass, surrounded by a protective covering. The glass fibre itself is thinner than a hair. Pulses of light or infra-red radiation travel along the fibre as a result of total internal reflection. These pulses carry digital information along the fibre.

The radiation pulses are provided by lasers and have very high frequencies, of the order of 10^8 MHz. In theory, a single pulse need only last for 10^{-14} s. However, lasers cannot be controlled at such high frequencies and the duration of a single pulse, or bit, is governed by the frequency at which the laser can be switched on and off. Technology is always improving and, at present, the frequency is about 800 MHz.

Fibre optic transmission is being installed in many different types of communication.

16.2 Modulation

For any system of communication, there must be a transmitter and a receiver. A simple system could be one person speaking to another. One person is the transmitter, the other is the receiver, and the communication system is sound waves.

An alternative system, if the two people are in different rooms, is a microphone connected to a loudspeaker by means of a wire-pair. The microphone converts the sound wave into a varying electrical signal that is transmitted along the wires to the loudspeaker where it is converted back to sound waves.

Communication of speech could also be achieved using radio waves. The signal from the microphone would be amplified and applied to an aerial. The radio waves produced would then be picked up by a receiving aerial. After amplification, the received signals would be passed to a loudspeaker. This simple system is shown in Figure 16.4.

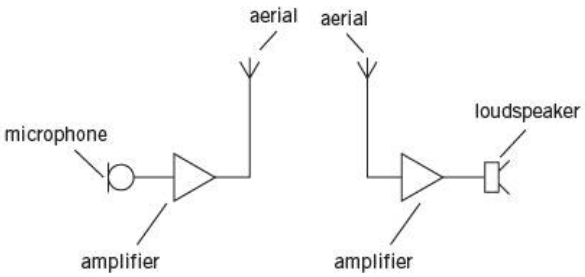


Figure 16.4 Simple radio communication system

www.iranino.co

IRANINOVIN

This simple system has two serious disadvantages. First, only one system could operate in an area because the receiving aerial would not be selective – it would pick up all signals. Second, the aerial required for the transmission of the low frequencies of sound waves (about 20 Hz to 15 kHz) would be very long and the electrical power required for transmission over long distances would be very large.

These problems are solved by the process known as **modulation**. A high-frequency wave, known as the **carrier wave**, is transmitted. This high-frequency carrier wave has either its amplitude varied or its frequency varied so as to carry information. This is illustrated in Figure 16.5.

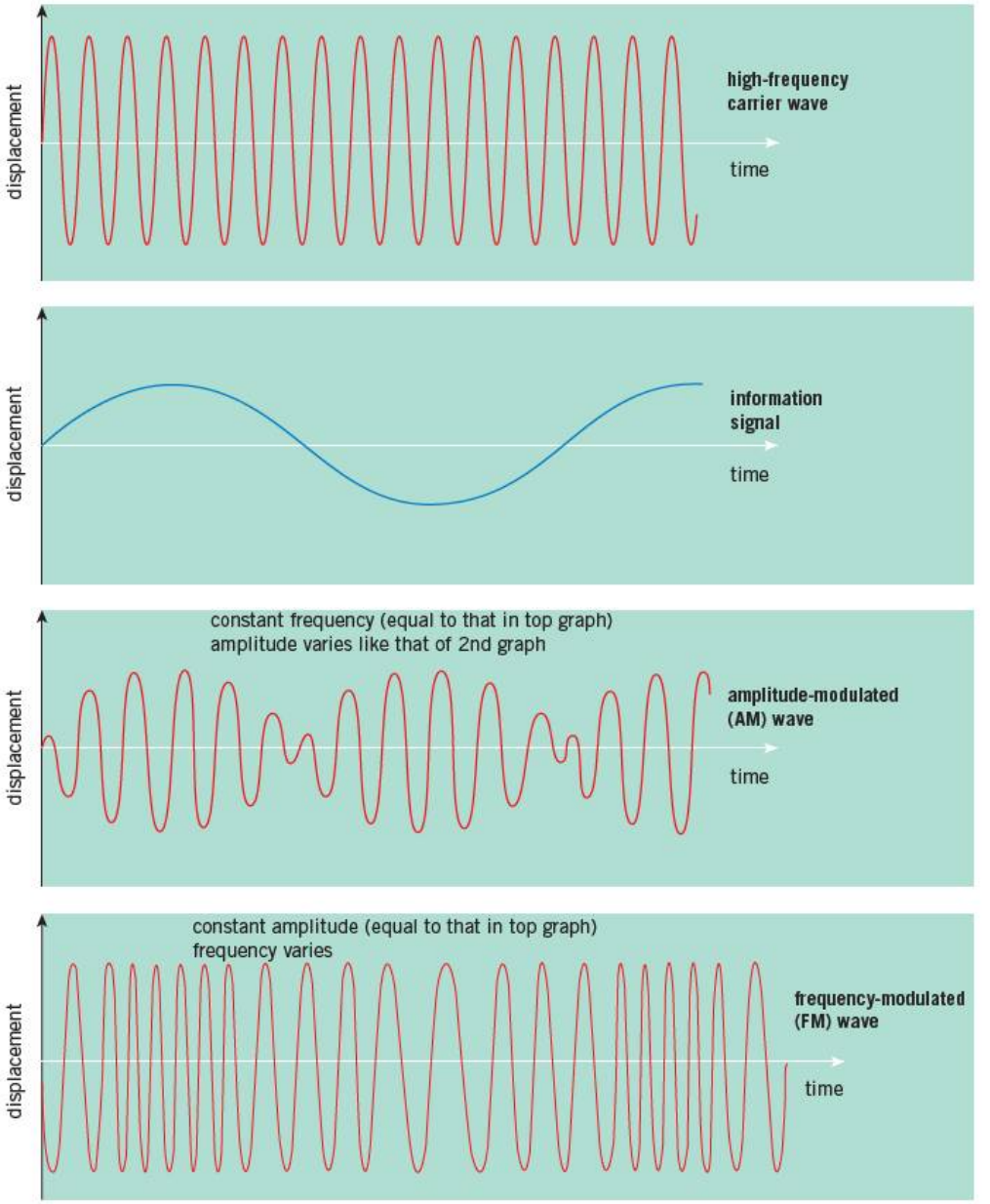


Figure 16.5 Amplitude modulation and frequency modulation of a carrier wave

In **amplitude modulation (AM)**, *the carrier wave has constant frequency. The amplitude of the carrier wave is made to vary. These variations are in synchrony with the displacement of the information signal.*

www.iranino.co

IRANINOVIN

The rate at which the amplitude of the carrier wave varies is related directly to the frequency of the information signal.

Note: Amplitude modulation is not the same as superposition. Superposition involves the addition of displacements whereas amplitude modulation is achieved through the multiplication of displacements.

In **frequency modulation (FM)**, *the amplitude of the carrier wave remains constant. The frequency of the carrier wave is made to vary in synchrony with the displacement of the information signal.*

The change in the frequency of the carrier wave is a measure of the displacement of the information signal. The rate at which the carrier-wave frequency varies is a measure of the frequency of the information signal.

The use of a carrier wave allows different radio stations in the same area to transmit at the same time. Each radio station has a different carrier wave frequency. The receiver is adjusted, or tuned, to the frequency of whichever transmitter is desired. The receiver accepts the signal transmitted on that particular carrier wave and rejects other carrier-wave frequencies.

Example

A sinusoidal carrier wave has a frequency of 800 kHz and an amplitude of 5.0 V. The frequency deviation of the carrier wave is 30 kHz V^{-1}. That is, for every 1.0 V change in displacement of the signal, the frequency of the carrier wave changes by 30 kHz. The carrier wave is frequency-modulated by a sinusoidal signal of frequency 10 kHz and amplitude 2.0 V. Describe the modulated carrier wave.

Amplitude of information signal = 2.0 V giving variation of (2 × 30) = 60 kHz.

The carrier wave has a constant amplitude of 5.0 V. Its frequency changes from 740 kHz to 860 kHz and back to 740 kHz. This change of frequency occurs 10 000 times per second.

Now it's your turn

1 A sinusoidal carrier wave has a frequency of 750 kHz and an unmodulated amplitude of 4.0 V. The carrier wave is to be amplitude-modulated by a sinusoidal signal of frequency 3 kHz and amplitude 0.5 V. Describe the modulated carrier wave.

2 The variation with time of the displacement of an amplitude-modulated radio wave is shown in Figure 16.6.

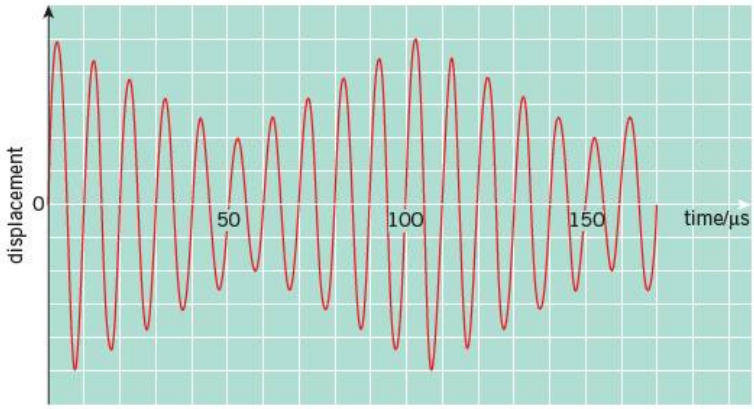


Figure 16.6

For this wave determine the frequency of:

(a) the carrier wave,

(b) the information signal.

www.iranino.co

IRANINOVIN

Figure 16.6 shows a carrier wave, frequency f_C, that has been amplitude-modulated by a signal having only one frequency f_S. If this waveform is analysed, it is found to be made up of three frequencies $(f_C + f_S)$, f_C and $(f_C - f_S)$. The frequency spectrum of this modulated wave – that is, a graph showing the variation with frequency of the amplitudes of each component – is shown in Figure 16.7.

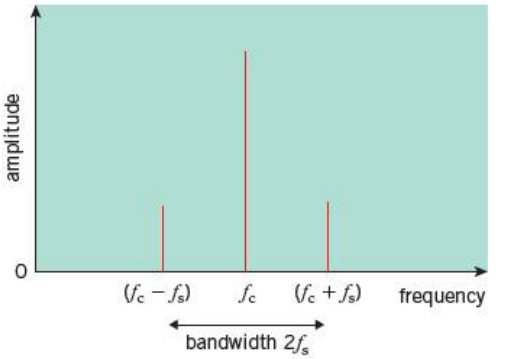


Figure 16.7 The spectrum of a carrier wave amplitude-modulated by a single frequency

The central frequency f_C is the frequency of the carrier wave and it has the largest amplitude. The other two frequencies, $(f_C + f_S)$ and $(f_C - f_S)$, are known as **sidebands** or **sideband frequencies**. The frequency f_S is the frequency of the information signal.

The **bandwidth** *is the range of frequencies occupied by the amplitude-modulated waveform.*

This bandwidth is equal to $2f_S$.

For the broadcast of music, the information signal will contain a wide range of frequencies, possibly from about 20 Hz to 15 kHz. A typical frequency spectrum of such an amplitude-modulated wave is shown in Figure 16.8.



Figure 16.8 Typical spectrum for amplitude-modulated wave transmitting music

In practice, the very high frequencies in music may not be broadcast, so as to reduce the bandwidth of the transmitted signal. Some quality of reproduction will be lost but, for normal broadcasting, this loss would be minimal.

The frequency spectrum of a frequency-modulated (FM) waveform differs from that for an amplitude-moderated (AM) waveform. The FM waveform has additional side frequencies that are multiples of the information signal frequency, resulting in a greater bandwidth for the same range of broadcast frequencies.

Radio-wave frequencies

Electromagnetic waves in the frequency range 30 kHz to 3 GHz are generally referred to as radio waves.

The intensity of the radio waves will always be reduced (attenuated) as the distance from the transmitter increases. The degree of attenuation depends on the frequency of the waves. Some data on frequencies and ranges are given in Table 16.1.

www.iranino.co

IRANINOVIN

Table 16.1 Radio waves

name of radio wave	frequency range	distance
space wave	greater than 30 MHz	line-of-sight between transmitter and receiver, plus satellite communication
sky wave	3 MHz → 30 MHz	worldwide, as a result of multiple reflections from the ionosphere and the ground or sea
surface wave	below 3 MHz	up to 1000 km

For simple radio communication, AM broadcasts on the medium-wave (MW) and long-wave (LW) wavebands are relatively cheap and technically less complex. They are transmitted as surface waves and provide coverage over large areas.

The wavelength of the radio waves determines the length of the aerial. For mobile phones, the aerial must be, for the sake of convenience, short and hence the wavelength must also be relatively short. High-frequency space waves are used. Remember that, as the frequency of the carrier wave increases, the bandwidth also increases.

The frequency bands used for radio communication are summarised in Table 16.2.

Table 16.2 Radio-frequency bands

communication type	frequency range	wavelength in air	frequency band
LW radio	30 kHz → 300 kHz	10 km → 1 km	low frequencies LF
MW radio	300 kHz → 3 MHz	1 km → 100 m	medium frequencies MF
short-wave (SW) radio	3 MHz → 30 MHz	100 m → 10 m	high frequencies HF
FM radio	30 MHz → 300 MHz	10 m → 1 m	very high frequencies VHF
TV broadcasting and mobile phones	300 MHz → 3 GHz	1 m → 10 cm	ultra-high frequencies UHF
microwave links	3 GHz → 30 GHz	10 cm → 1 cm	super-high frequencies SHF
satellite links	30 GHz → 300 GHz	1 cm → 1 mm	extra-high frequencies EHF

Relative advantages of AM and FM transmissions

AM radio transmissions on the long-wave (LW), medium-wave (MW) and short-wave (SW) wavebands are broadcast over very large distances so that one transmitter can serve a large area. FM transmissions have a range of only about 30 km and this range is by line-of-sight. Consequently, many transmitters are required to broadcast over a large area. FM can be used to serve small local areas but it is simpler and cheaper to broadcast using AM.

The bandwidth of AM broadcasts on the LW and MW wavebands is 9 kHz. This means that the highest frequency that can be broadcast is 4.5 kHz. This is quite adequate for speech. However, the lack of quality for music is easily noticed. On the very-high-frequency (VHF) waveband, the bandwidth of an FM radio signal is about 200 kHz, giving a maximum frequency that can be broadcast of about 15 kHz. FM broadcasting does, therefore, offer higher quality.

The long-wave (LW) waveband has a range of frequencies from about 30 kHz to 300 kHz. If the bandwidth on each AM broadcast is 9 kHz, then, theoretically, $(300 - 30)/9 = 30$ transmitters could broadcast in the same area without causing interference. In practice, the number would be less than this, so that 'space' in the frequency spectrum is left between transmitters. For FM broadcasting, $(300 - 30)/200 = 1$ transmitter could broadcast in the LW waveband. The number of transmitters that can share the same waveband is much larger for AM than for FM. For this reason, FM is broadcast only at frequencies in excess of 1 MHz.

Electrical equipment that produces sparks also produces electromagnetic waves. (For this reason, the ignition systems of petrol engines must be 'suppressed'.) These electromagnetic waves are detected as interference by the aerial in a radio. If the radio is tuned into an AM signal, the interference will add to the displacement of the AM signal and will appear

www.iranino.co

IRANINOVIN

as noise in the output of the radio. An FM signal is based on changes of frequency, not displacement. Interference is not picked up by the aerial of the FM radio because the interference does not alter the frequency of the signal. The quality of the reception of FM is generally better than that of AM, since there will be less noise or interference.

AM transmitters and receivers are much simpler, electronically, than those for FM. It is, therefore, cheaper to broadcast and receive AM.

In summary, FM transmissions are more expensive than AM and the area covered by one FM transmitter is much smaller. This difference in coverage is an advantage where local radio is concerned but a disadvantage as regards national radio. The bandwidth necessary for FM is greater. The quality of the received FM signal is much better because of the increased frequency spectrum, and it also suffers less noise.

Example

A particular transmitter is broadcasting an AM signal of frequency 200 kHz. The transmitter is broadcasting a programme of music with a maximum frequency of 4.5 kHz. Determine, for this AM signal:

(a) the wavelength,
(b) the bandwidth.

(a) speed of electromagnetic waves is 3×10^8 m s^{-1}
wavelength = *speed*/*frequency* = $(3 \times 10^8)/(200 \times 10^3)$ = **1500 m**
(b) bandwidth = 2×4.5 = **9.0 kHz**

Now it's your turn

3 Discuss the relative advantages of AM and FM radio transmissions. Wherever appropriate, give numerical values to illustrate your answer.

16.3 Analogue and digital signals

An information signal that has the same variations with time as the information itself is known as an **analogue** signal. For example, the signal voltage produced by a microphone is analogous to the sound wave incident on the microphone. The voltage output from the microphone is an analogue signal.

Much of the information that we wish to transmit and communicate is analogue in nature (speech, music, television pictures, etc.). When any signal is transmitted over a long distance, it will pick up noise. Noise is not just unwanted sound, but any unwanted random signal that adds to the signal that is being transmitted. Also, the power of the signal becomes less; that is, the signal is attenuated. For long-distance transmission, the signal has to be amplified at regular intervals. The problem is that, on amplification of an analogue signal, the noise is also amplified. The signal becomes distorted or 'noisy'. This effect is shown in Figure 16.9.

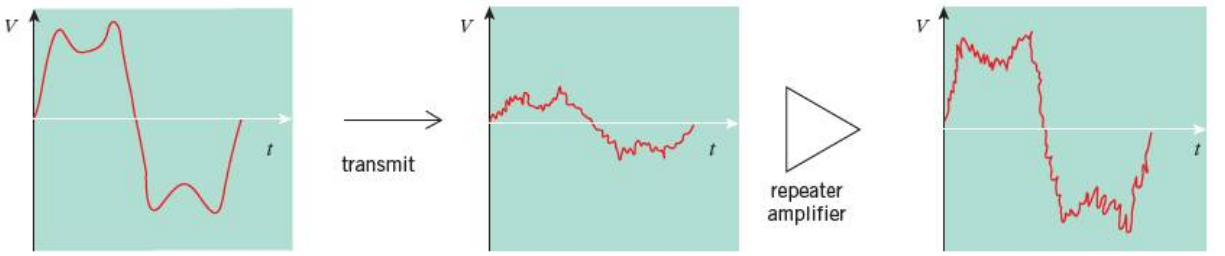


Figure 16.9 Amplification of a noisy analogue signal

A **digital** signal consists of a series of 'highs' and 'lows' with no values between the 'highs' and the 'lows'. The data in the signal is transmitted as a particular sequence of 'highs' and 'lows' or, effectively, a sequence of 1s and 0s. The signal will still suffer from noise and attenuation. However, on amplification, the noisy 1s and 0s can be re-shaped or regenerated to return the signal to its original form. Such amplifiers are

www.iranino.co

IRANINOVIN

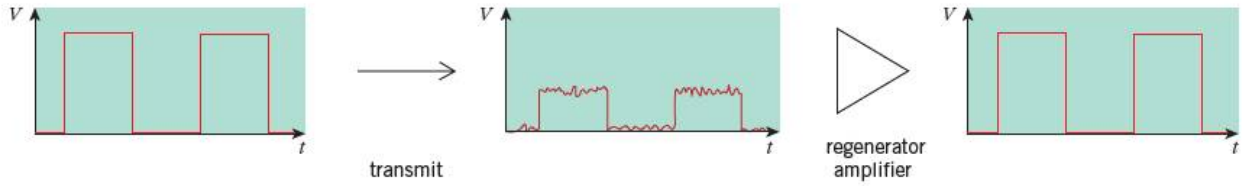


Figure 16.10 Amplification of a noisy digital signal

known as regenerator amplifiers. These amplifiers 'filter out' any noise, and in so doing, restore the signal. This is illustrated in Figure 16.10.

Unlike an analogue signal, a digital signal can be transmitted over a long distance with regular regenerations without the signal becoming degraded.

Modern digital electronic circuits are, in general, more reliable and cheaper to produce than analogue circuits. In the future, the vast majority of communications systems are likely to be digital.

An additional advantage of digital systems is that extra information, or data, can be added to the transmissions. These extra data are a code for the receiving system so that the transmitted signal may be checked and corrected before the signal is finally reproduced.

Table 16.3

decimal number	binary number
0	0000
1	0001
2	0010
3	0011
4	0100
5	0101
6	0110
7	0111
8	1000
12	1100
13	1101
15	1111

Binary numbers

A binary number is a number that has the base 2, whereas a decimal number has the base 10. Table 16.3 shows some decimal numbers and their equivalents in binary notation. A binary number consists of a number of digits, or **bits**. All the binary numbers shown are four-bit numbers. Binary numbers may also be called digital numbers.

Larger numbers would require digital numbers with more bits. When reading a digital number, the bit on the left-hand side of the number is called the **most significant bit**, or MSB. This bit has the highest value. The bit on the right-hand side has the least value and is known as the **least significant bit**, or LSB.

When the LSB is 1, this corresponds to decimal number 1. When the second bit shows 1, this corresponds to decimal number 2. When successive bits show 1 they correspond to decimal numbers 4, 8, 16, 32, 64, etc. Thus, the binary number 1101 corresponds to decimal number 8 + 4 + 0 + 1, or 13. Conversely, decimal number 11 which equals 8 + 0 + 2 + 1 corresponds to binary number 1011.

Example

Convert the following decimal numbers into five-bit digital numbers:

(a) 9,

(b) 23.

(a) **01001**

(b) **10111**

Now it's your turn

4 Convert the following five-bit numbers into decimal numbers:

(a) 11001,

(b) 01010.

The transmission of a signal

In an analogue signal, such as speech or music, the generated voltage signal varies continuously. For digital transmission, this analogue signal must be converted into a digital signal. This is achieved using an **analogue-to-digital converter (ADC)**.

In an analogue-to-digital converter, the analogue voltage is sampled at regular intervals of time, at what is known as the **sampling frequency** *or* **sampling rate**. *The value of the sample voltage measured at each sampling time is converted into a digital (binary) number that represents the voltage value.*

www.iranino.co

IRANINOVIN

For example, assuming that a four-bit number is being used, then the number representing a signal that is sampled as 5.0V would be 0101. Note that when sampling, the number representing the sample would be the whole number *below* the actual value of the sampled voltage. If the signal were to be sampled as 11.4V, then the 4-bit number would be 1011. A sampled signal of 11.8V would also be 1011.

Figure 16.11a shows a sinusoidal analogue signal that is to be sampled at a sampling rate of 10kHz. A four-bit system is used for the digital numbers generated.

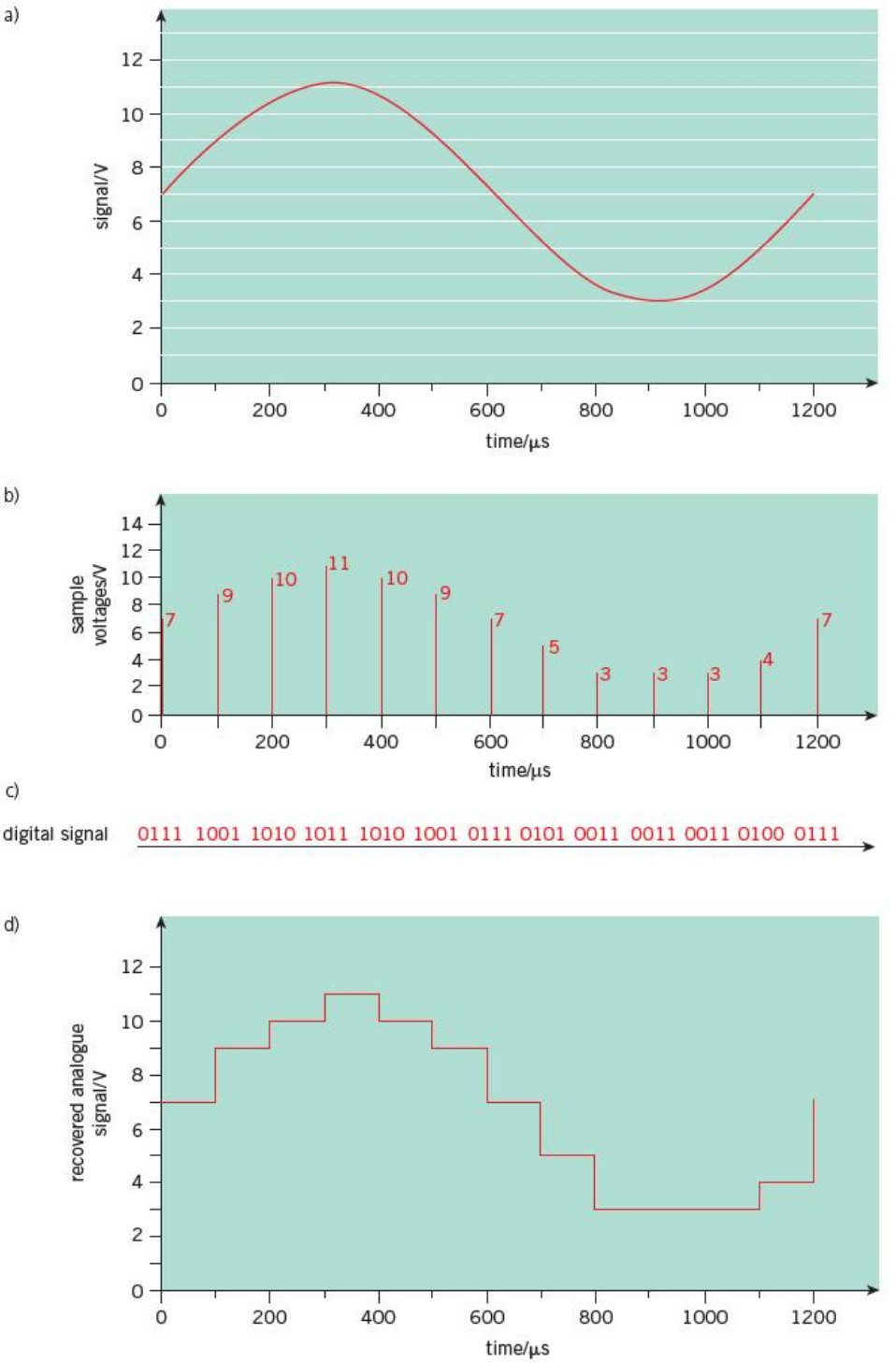


Figure 16.11 Analogue-to-digital and digital-to-analogue conversion

www.iranino.co

IRANINOVIN

The sample voltages are shown in Figure 16.11b. These sample voltages are converted into a digital signal, shown in Figure 16.11c, by the analogue-to-digital converter (ADC). After this digital signal had been transmitted, it is converted back into an analogue signal using a **digital-to-analogue converter (DAC)**.

In a digital-to-analogue converter, a digital signal is converted into an analogue signal.

Figure 16.11d shows the analogue signal that has been recovered. The recovered signal is very 'grainy', consisting of large 'steps'. The size of these steps, and hence the faithfulness of the reproduction of the initial analogue signal, can be improved by using more voltage levels and sampling at a higher frequency.

The number of bits in each digital number limits the number of voltage levels. In this example, there are four bits and $2^4 = 16$ levels. In practice, eight or more bits would be used for sampling. An eight-bit number would give $2^8 = 256$ levels.

The choice of sampling frequency also determines the amount of information that can be transmitted. About 100 years ago, Nyquist showed that, in order to recover an analogue signal of frequency f, then the signal must be sampled at a frequency greater than $2f$. The greater the sampling frequency, the more faithful is the reproduction of the original signal. For example, for good quality reproduction of music, the higher audible frequencies must be present; that is, frequencies up to about 20 kHz. For compact discs (CDs) the sampling frequency is therefore 44.1 kHz. This quality of reproduction is not required for speech, and it would prove costly. In a telephone system, the sampling frequency is 8 kHz and the highest frequency to be transmitted is limited to 3.4 kHz.

16.4 Relative merits of channels of communication

Wire-pairs

High-frequency electrical signals lose their energy over short distances in wire-pairs – the signals have a high attenuation. This high attenuation is partly due to the heating caused by the electrical resistance of the wires and partly due to the emission of radiation (radio waves), since the wires act as aerials. A signal going any distance in a wire-pair must be amplified at regular intervals along the length of the wires.

A further problem is that, since the wires acts as aerials, they pick up any electromagnetic waves and these unwanted signals, or noise, will cause deterioration of the signal. Wire-pairs close to one another pick up each other's signals. This effect is called **cross-talk** or **cross-linking** and means that wire-pairs give rise to poor security since the signals can be 'tapped' easily.

The bandwidth of a wire-pair is only about 500 kHz and, as a result, wire-pairs are limited as to the amount of information that they can carry.

To summarise, wire-pairs:

- are used mainly for short-distance communication
- cause high attenuation of a signal
- easily pick up noise
- suffer from cross-talk and are of low security
- have limited bandwidth.

Coaxial cables

Attenuation is reduced when compared to wire pairs. Consequently, repeater amplifiers can be further apart on coaxial cables.

The bandwidth of a coaxial cable is about 50 MHz. Much more information can be carried along a coaxial cable than along a wire-pair. However, coaxial cable is more costly. Coaxial cable is used to connect, for example, an aerial and a television.

www.iranino.co

IRANINOVIN

To summarise, comparing a wire-pair and a coaxial cable, the coaxial cable:

- is more costly
- causes less attenuation
- is less noisy and is more secure
- has a larger bandwidth.

Radio and microwave links

Microwaves are electromagnetic waves that can be considered alongside other radio waves, as shown in Table 16.2. They are in the frequency range 3 GHz to 30 GHz and are generally used for point-to-point communication since, for use on Earth, the range of the transmissions is limited to line-of-sight.

Reflecting parabolic dishes are used so that the transmission is in the form of a parallel beam and so that as much wave power as possible can be focused onto the receiving aerial, as shown in Figure 16.2.

The bandwidth of a microwave link is of the order of 1 GHz. This large bandwidth means that the microwave beam has a large capacity for transmitting information.

To summarise the information about radio waves and microwaves:

- these are electromagnetic waves with a wide range of frequencies.
- their mode of use and range depend on frequency.
- security is achieved only by encoding the information.

Optic fibres

Pulses of light or infra-red radiation travel along the fibre as a result of total internal reflection. These pulses carry digital information along the fibre.

The radiation pulses are provided by lasers and have very high frequencies, of the order of 10^8 MHz. In theory, a single pulse need only last for 10^{-14} s. However, lasers cannot be controlled at such high frequencies and the duration of a single pulse, or bit, is governed by the frequency at which the laser can be switched on and off. Technology is always improving and, at present, the frequency is about 800 MHz. Such high frequencies mean a large bandwidth, so many different phone calls can share the same optic fibre.

The advantages of optic fibres are summarised below:

- large bandwidth, giving rise to large transmission capacity
- much lower cost than metal wires
- diameter and weight of cable is much less than metal cable, hence easier handling and storage
- much less signal attenuation, so far fewer regenerator amplifiers are required, reducing the cost of installation
- do not pick up electromagnetic interference, so very high security and negligible cross-talk
- can be laid alongside existing routes such as electric railway lines and power lines.

Communication satellites

Long-distance communication using radio waves is possible on the MW waveband (as surface waves) and the SW waveband (as sky waves). However, for modern communication systems, there are three major disadvantages.

1. Long-distance communication using sky waves is unreliable in that it depends on reflection from layers of ions in the upper atmosphere. These layers of ions vary in height and density, giving rise to variable quality of signal. Surface waves are also unreliable because there is poor reception in hilly areas.
2. The wavebands available on MW and SW are already crowded.
3. The bandwidths that are available are narrow and, consequently, unable to carry large amounts of information.

www.iranino.co

IRANINOVIN

Satellite communication enables more wavebands to be made available and at much higher frequencies, thus giving rise to a much greater data-carrying capacity. The basic principle of satellite communication is shown in Figure 16.12.

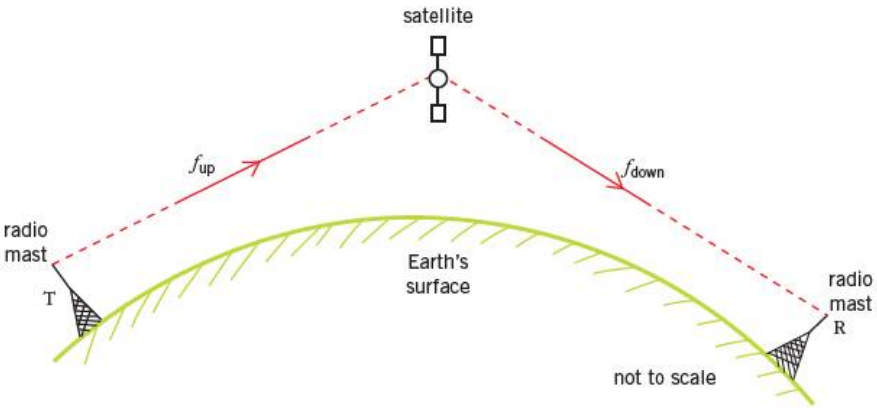


Figure 16.12 Satellite communication

A carrier wave of frequency f_{up} is sent from a transmitter T on Earth to a satellite. The satellite receives the greatly attenuated signal. The signal is amplified and the carrier frequency is changed to a lower value, f_{down}. The carrier wave is then directed back to a receiver R on Earth. The carrier wave frequencies f_{up} and f_{down} are different so that the very low power signal received from Earth is not swamped by (can be distinguished from) the high power signal that is transmitted back to Earth.

Values of the 'uplink' frequency f_{up} and the 'downlink' frequency f_{down} might be 6 GHz and 4 GHz respectively (the 6/4 GHz band). Other bands are 14/11 GHz and 30/20 GHz.

The satellite will probably be a **geostationary satellite** *in a geostationary orbit. In this case, the satellite orbits the Earth with a period of 24 hours at a height of 3.6 × 10⁴ km above the Earth's surface. The geostationary satellite orbits in the same direction as the rotation of the Earth (from west to east) and the orbit is above the Equator.*

From the viewpoint of a person on Earth, the satellite remains above the same point on the Earth's surface. The transmitting aerial and the receiving aerial on Earth both have large parabolic reflectors. One advantage of geostationary satellites is that these aerials can be fixed in position since the satellite does not have to be tracked.

A geostationary satellite can have a permanent link with a transmitting ground station. Communication is maintained with any point on the Earth's surface that can receive the signal from the satellite. A number of satellites with overlapping areas, such that any area can receive a signal from one of the satellites, allows for long-distance communication. This removes the need for long-distance submarine cables. International television broadcasting is possible, allowing 'live' events in one country to be viewed in another.

Geostationary satellites are in equatorial orbits. This means that communication in polar regions may not be possible because a satellite will not be in line-of-sight. A further problem results because of the height of the orbit. Between transmission and receipt of the signal, the wave must travel at least twice the distance between the satellite and Earth; that is, 7.2×10^4 km. For the wave travelling at 3.0×10^8 m s^{-1}, the time taken is 0.24 s. This delay would be increased where several satellites are involved, and would not be acceptable during a telephone conversation. To avoid these problems, geostationary satellites may be used in conjunction with optic fibres.

Polar satellites *are satellites that have low orbits and pass over the poles.*

www.iranino.co

IRANINOVIN

The orbital period of polar satellites is about 100 minutes. Since the Earth rotates below such satellites, then in any period of 24 hours the satellite will pass over every region of the Earth's surface.

Continuous communication with a single polar satellite is not possible. However, information may be transmitted to the satellite while it is overhead. The data can be stored in the satellite and then transmitted back to Earth when the satellite is over the appropriate area. Continuous communication is possible using a number of polar satellites in orbits that are inclined to one another so that at least one satellite is always above the transmitter and receiver. In this case, the aerials must track the satellites in their orbits. The advantage of using such satellites is that their orbital height is only of the order of 10^5 m (a few hundred km) and thus delays in telephone conversations are not noticed.

Since polar satellites pass over the whole of the Earth in any 24-hour period, they are used for remote sensing. Such uses include military espionage, geological prospecting and weather forecasting. The Global Positioning System (GPS, or 'sat nav') uses the signals from a number of satellites that are not in geostationary orbits.

16.5 Signal attenuation

When a signal passes along a wire or a fibre, it loses power. This loss of power is referred to as **attenuation**, and the amount of attenuation increases as the distance that the signal travels increases. In the case of an electrical signal in a metal wire, signal power is lost as heating of the wire. In optic fibres, light power is lost as a result of absorption in impurities in the glass and also scattering due to imperfections. A beam of electromagnetic waves travelling through air also loses power as a result of absorption and scattering.

In order that a signal may be detected, the power of the signal must be a minimum number of times greater than the noise power. This ratio, the **signal-to-noise ratio**, may be, for example, 50 to 100. Signals are amplified and the output of the amplifier is a certain number of times greater than the input. The **amplifier gain** could be 100000, or 10^5. When a microwave signal is sent from Earth to a satellite, the signal power may be reduced by a factor of 10^{19}! It can be seen that, in comparing power levels, the ratio of two numbers is involved and that this ratio can be very large or very small. In order to condense the scale of such variations and to make the numbers more manageable, the power levels are compared on a logarithmic scale. The result of this comparison gives the ratio in a unit known as the **bel**, which has the symbol B.

$$\text{number of bels} = \lg\left(\frac{P_2}{P_1}\right)$$

where P_2 and P_1 are the two powers that are being compared. Since the bel is a large unit, the ratio is usually expressed as the **decibel**, where 1 bel = 10 decibels or 10 dB.

$$\text{number of decibels (dB)} = 10\lg\left(\frac{P_2}{P_1}\right)$$

It should be noted that if P_2 is greater than P_1, the dB number is positive and there has been amplification. Where P_2 is less than P_1, there is attenuation and the number of dB will be negative.

Example

A signal having a power of 2.4 μW is amplified in a two-stage amplifier. The first stage has a gain of 18 dB and the second stage provides a further amplification of 25 dB. Calculate:

(a) the total gain of the two-stage amplifier,

(b) the power of the output signal from the amplifier.

(a) For the first stage, the input is V_I and the output is V, giving a gain of V/V_I.
For the second stage, the input is V and the output is V_O, giving a gain of V_O/V.

www.iranino.co

IRANINOVIN

Total gain = $V_O/V_I = V_O/V \times V/V_I$

= gain of first stage × gain of second stage

Gain, when expressed in dB, is a logarithm. Therefore, when two gains are multiplied together, then the gains in dB must be added (see Maths Note).

lg (total gain) = lg (gain of first stage) + lg (gain of second stage)

= 18 + 25 = **43 dB**

(b) Gain in dB = 10 lg (P_2/P_1)

$43 = 10 \lg (P_2/(2.4 \times 10^{-6}))$

$4.3 = \lg (P_2/(2.4 \times 10^{-6}))$

Taking antilogs,

$10^{4.3} = P_2/(2.4 \times 10^{-6})$

$P_2 = \mathbf{0.048\ W}$

Now it's your turn

5 A signal is amplified by two amplifiers, A and B, connected in series. The input power to amplifier A is 17 μW. Amplifier A has a gain of 20 dB. The output power from amplifier B is 0.26 W. Determine:

(a) the overall gain, in dB, of the two amplifiers,

(b) the gain of amplifier B.

Maths Note

The logarithm of a number, to the base 10, is the power to which the number ten must be raised in order to give that number.

For example,

$100 = 10^2$ and so the logarithm to the base ten of 100 (lg 100) is 2.00

and

$2.00 = 10^{0.301}$ and so the logarithm to the base ten of 2.00 (lg 2.00) is 0.301

also

$10^{1.699} = 50.0$ and so lg 50.0 = 1.699

Remember that when two numbers are multiplied together, if the numbers are expressed as numbers to the base ten, then the powers of the base ten are added. For division, the powers are subtracted.

For example, $75 \times 3.00 = 10^{1.875} \times 10^{0.477}$

Adding the powers, $75 \times 3.00 = 10^{2.352}$

$10^{2.352} = 225$

(We already know that $75 \times 3 = 225$!)

Also, $75 \div 3.00 = 10^{1.875} \div 10^{0.477}$

Subtracting the powers, $75 \div 3.00 = 10^{1.398}$

$10^{1.398} = 25$

(We already know that $75 \div 3 = 25$!)

The gain of an amplifier is usually expressed in dB.

Now, gain (in dB) = 10 lg(P_{out}/P_{in}) and so the gain, in dB, is a logarithm to the base ten.

When two amplifiers are connected in series, the combined gain is found *either* by multiplying together the two actual gains *or* by adding together the gains, when expressed in dB.

In a transmission line, the amount of attenuation is dependent on the length of the line. As a result, **attenuation per unit length** is normally quoted.

$$\text{attenuation per unit length} = \left(\frac{1}{L}\right) 10 \lg \left(\frac{P_2}{P_1}\right)$$

where L is the length of the transmission line. Attenuation per unit length is measured in dB km^{-1}.

An optic fibre may have an attenuation per unit length of between 1 and 3 dB km^{-1}, whereas for a coaxial cable, it varies with frequency, increasing from about 3 dB km^{-1} at 10 MHz to about 40 dB km^{-1} at 3 GHz.

Example

The signal input to an optic fibre is 7.0 mW. The average noise power in the fibre is 5.5×10^{-19} W and the signal-to-noise ratio must not fall below 24 dB. The fibre has an attenuation of 1.8 dB km^{-1}. Calculate:

(a) the minimum effective signal power in the optic fibre,

(b) the maximum uninterrupted length of the optic fibre through which the signal can be transmitted.

www.iranino.co

IRANINOVIN

(a) number of decibels (dB) = 10 lg (P_2/P_1)

$$24 = 10 \lg (P_{min}/5.5 \times 10^{-19})$$

where P_{min} is the minimum effective signal power.

$$10^{2.4} = P_{min}/5.5 \times 10^{-19}$$

minimum effective signal power P_{min} = **1.38×10^{-16} W**

(b) total attenuation of signal to reach minimum = 10 lg (P_{input}/P_{min})

$$= 10 \lg (7.0 \times 10^{-3})/(1.38 \times 10^{-16})$$

$$= 137 \text{ dB}$$

maximum uninterrupted length = total attenuation/attenuation per unit length

$$= 137/1.8 = \textbf{76 km}$$

Now it's your turn

6 A laser is used to provide the input to an optic fibre of length 65 m. The power input to the fibre is 6.20 mW and the output at the other end of the fibre is 4.9×10^{-12} W.

(a) Calculate the power loss, in dB km^{-1}, in the optic fibre.

(b) The average noise power in the fibre is 2.4×10^{-18} W and the minimum acceptable signal-to-noise ratio is 22 dB. Calculate the maximum length of uninterrupted fibre through which the signal could be transmitted.

Summary

- Information can be carried by different channels including wire-pairs, coaxial cables, radio waves and microwaves, and optic fibres.
- Modulation is the process whereby either the amplitude (in AM) or the frequency (in FM) of a carrier wave is varied so as to carry information.
- Bandwidth is the range of frequencies that is used in any particular broadcast.
- Bandwidth is higher for FM broadcasts, giving better sound quality.
- Information may be transmitted as an analogue signal or as a digital signal.
- A digital signal can be transmitted over a long distance without the signal becoming degraded if regenerator amplifiers are used.
- Digital communication involves analogue-to-digital conversion (ADC) on transmission and digital-to-analogue conversion (DAC) on reception.
- For good reproduction of an audio signal, the number of bits in a sample and the sampling rate need to be as high as possible.
- The ratio of two powers may be expressed logarithmically and is measured in decibels (dB):
  number of dB = 10 lg (P_2/P_1).
- The signal attenuation, the power gain on amplification, and the signal-to-noise ratio, are all given as a ratio of two powers.
- Satellites in geostationary and in polar orbits are used for communication.

Examination style questions

1 Both wire-pairs and coaxial cables are used in communication systems.

a Suggest two uses of wire-pairs in communication systems.

b Compare the use of wire-pairs and coaxial cables in communication systems.

2 The use of geostationary satellites and optic fibres for worldwide communication have, for many applications, replaced the use of sky waves.

a Outline the use of sky waves for worldwide communication.

b Suggest reasons why geostationary satellites and optic fibres have replaced sky waves.

c Explain why a combination of geostationary satellites and optic fibres is preferable to geostationary satellites alone for communication over very long distances.

3 a Outline the principles of communication using geostationary satellites.

b Suggest one disadvantage of using several geostationary satellites for communication between opposite sides of the Earth.

c Discuss the advantages and disadvantages of using polar, rather than geostationary, satellites for radio communication.

www.iranino.co

IRANINOVIN

4 During the development of commercial radio, AM broadcasting preceded FM broadcasting. Explain why, in broadcasting:

a AM was developed before FM,

b there has been a change-over from AM to FM.

5 Fig. 16.13 shows the variation with time of an analogue signal.

The analogue signal is to be sampled at a rate of 1.0 kHz and converted into four-bit digital numbers. After digital transmission, the analogue signal is to be recovered using a DAC.

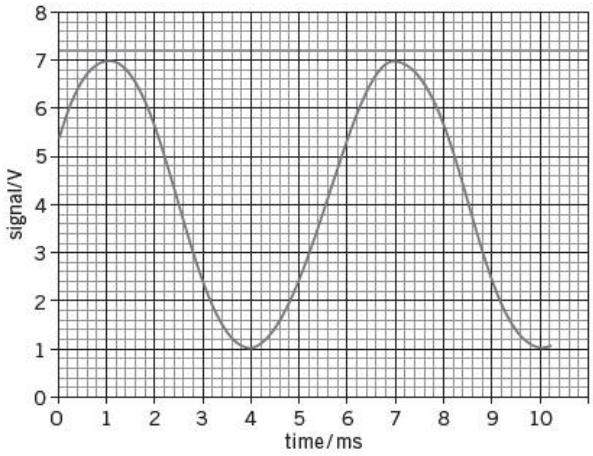


Fig. 16.13

Draw diagrams (similar to those in Fig. 16.11) to show

a the sample voltages in analogue form,

b the sample voltages in digital form,

c the recovered analogue signal.

6 State **three** advantages of a coaxial cable compared with a wire-pair for the transmission of an electrical signal. *[3]*

Cambridge International AS and A Level Physics, 9702/06 May/June 2003 Q 14

7 A radio signal may be transmitted between a transmitter and a receiving aerial by means of sky waves, ground (surface) waves or space waves. Copy and complete Table 16.4 by giving a typical wavelength and the maximum transmission range for each type of wave.

Table 16.4

type	wavelength / m	range
sky wave	.....	.....
ground (surface) wave	.....	.....
space wave	.....	.....

[5]

Cambridge International AS and A Level Physics, 9702/06 May/June 2003 Q 15

8 A radio station emits an amplitude-modulated wave for the transmission of music.

a i State what is meant by an *amplitude-modulated* (AM) wave. *[2]*

ii Give two reasons why the transmitted wave is modulated, rather than transmitting the information signal directly as a radio wave. *[2]*

b The variation with frequency f of the amplitude A of the transmitted wave is shown in Fig. 16.14.

For this transmission, determine

i the wavelength of the carrier wave, *[2]*

ii the bandwidth, *[1]*

iii the maximum frequency, in Hz, of the transmitted audio signal. *[1]*

Cambridge International AS and A Level Physics, 9702/43 May/June 2013 Q 11

9 The digital transmission of speech may be represented by the block diagram of Fig. 16.15.

a State the purpose of the parallel-to-serial converter. *[2]*

b Part of the signal from the microphone is shown in Fig. 16.16.

The ADC (analogue-to-digital converter) samples the analogue signal at a frequency of 5.0 kHz.

Each sample from the ADC is a four-bit digital number where the smallest bit represents 1.0 mV.

The first sample is taken at time zero.

Use Fig. 16.16 to determine the four-bit digital number produced by the ADC at times

i 0.4 ms, *[1]*

ii 0.8 ms. *[1]*

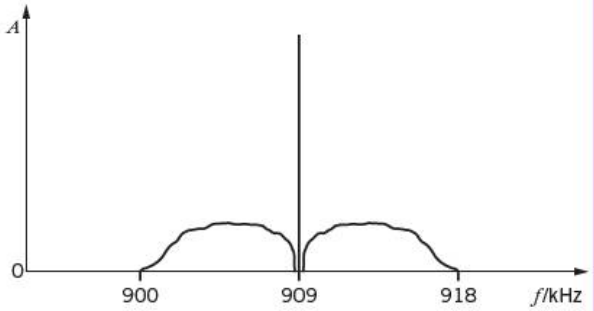


Fig. 16.14

www.iranino.co

IRANINOVIN



Fig. 16.15

c The digital signal is transmitted and then converted to an analogue form by the DAC (digital-to-analogue converter).

Using data from Fig. 16.16, draw, on the axes of a copy of Fig. 16.17, the output level of the transmitted analogue signal for time zero to time 1.2 ms.

d State and explain the effect on the transmitted analogue waveform of increasing, for the ADC and the DAC, both the sampling frequency and the number of bits in each sample. *[3]*

Cambridge International AS and A Level Physics, 9702/42 May/June 2013 Q 12

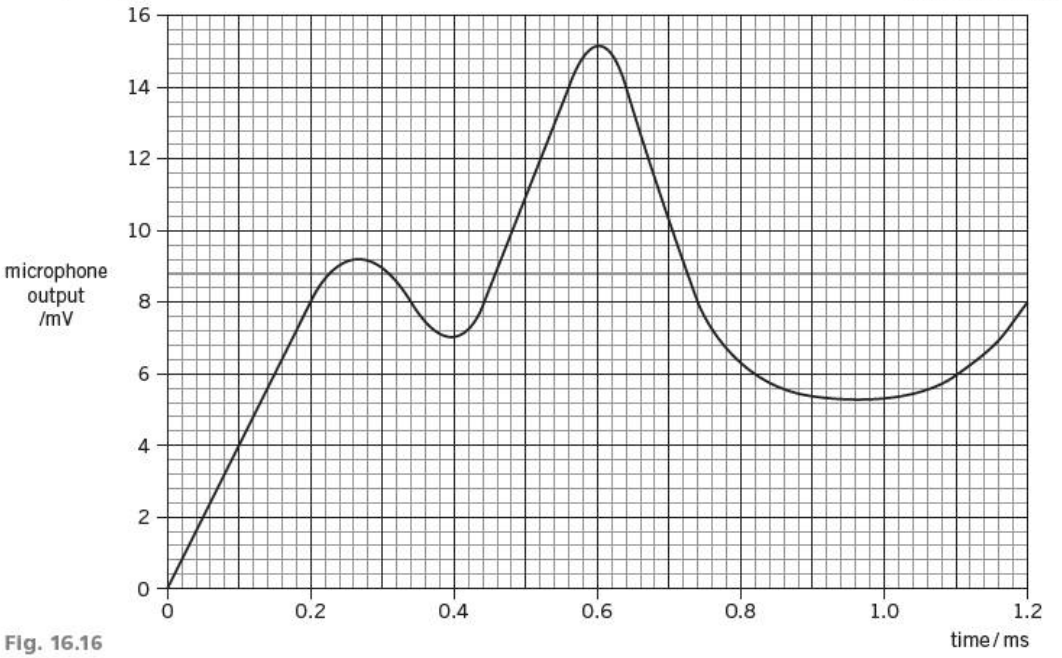


Fig. 16.16

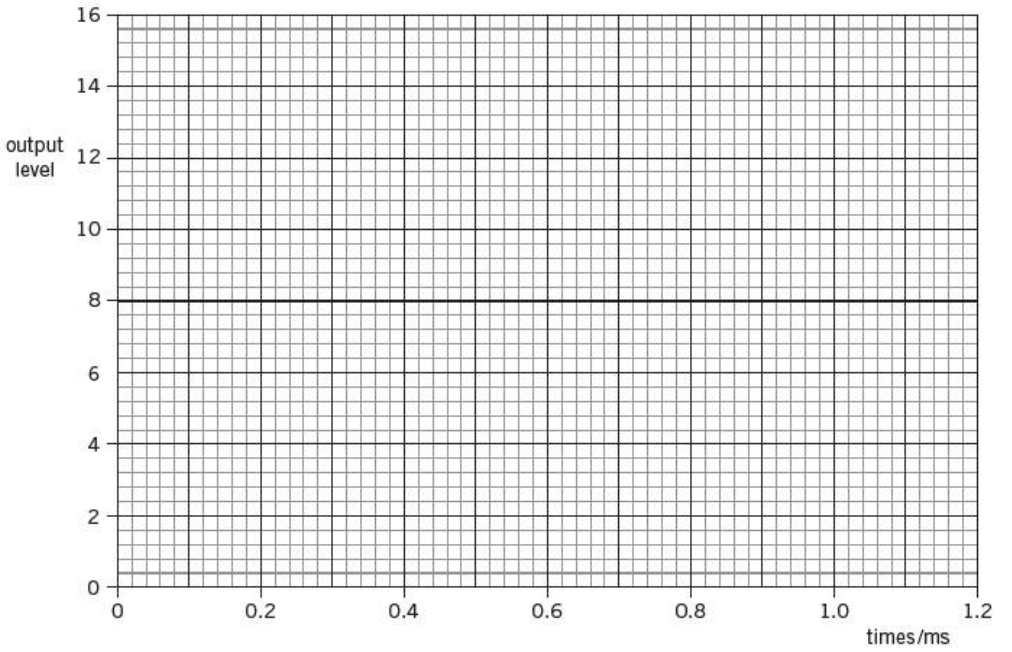


Fig. 16.17

www.iranino.co

IRANINOVIN

10 An optic fibre is used for the transmission of digital telephone signals. The power input to the optic fibre is 9.8 mW. The effective noise level in the receiver circuit is 0.36 μW, as illustrated in Fig. 16.18.

The signal-to-noise ratio at the receiver must not fall below 28 dB.

For this transmission without any repeater amplifiers, the maximum length of the optic fibre is 85 km.

a Calculate the minimum input signal power to the receiver. *[2]*

b Use your answer in **a** to calculate the attenuation in the fibre. *[2]*

c Determine the attenuation per unit length of the fibre. *[1]*

Cambridge International AS and A Level Physics, 9702/43 May/June 2013 Q 12

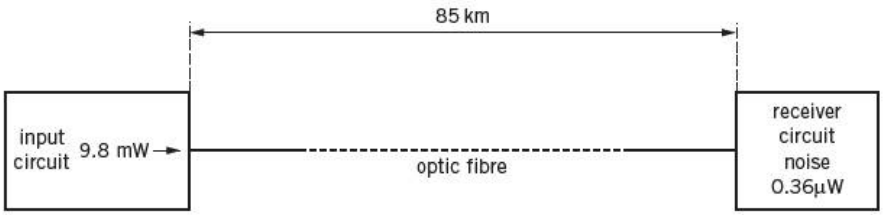


Fig. 16.18

www.iranino.co

IRANINOVIN

17 Electric fields

By the end of this topic, you will be able to:

17.3 (a) understand that, for any point outside a spherical conductor, the charge on the sphere may be considered to act as a point charge at its centre
(b) recall and use Coulomb's law in the form $F = Q_1Q_2/4\pi\varepsilon_0 r^2$ for the force between two point charges in free space or air

17.4 (a) recall and use $E = Q/4\pi\varepsilon_0 r^2$ for the field strength of a point charge in free space or air

17.5 (a) define the potential at a point as the work done per unit positive charge in bringing a small test charge from infinity to the point
(b) state that the field strength of the field at a point is equal to the negative of the potential gradient at that point
(c) use the equation $V = Q/4\pi\varepsilon_0 r$ for the potential in the field of a point charge
(d) recognise the analogy between certain qualitative and quantitative aspects of electric fields and gravitational fields

Starting points

- An electric field is a region of space where electric charge experiences a force.
- Electric fields may be represented by field lines.
- Be able to calculate the forces on charges in uniform electric fields.
- Be able to describe the effect of uniform electric fields on the motion of charged particles.

17.3 Point charges

In AS Level Topic 17, we saw that the electric field surrounding a charged sphere is radial, as shown here in Figure 17.1.

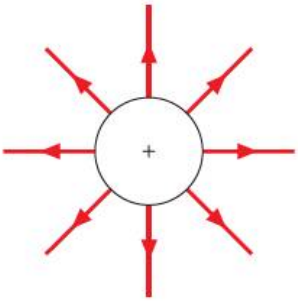


Figure 17.1 Electric field near an isolated charged sphere

The sphere is a conductor and so the charge on the sphere distributes itself evenly around the surface of the sphere. However, from any position outside the sphere, the electric field lines appear to radiate from the centre of the sphere. Consequently:

for any point outside a spherical conductor, the charge on the sphere may be considered to act as a point charge at the centre of the sphere.

Force between point charges

We are familiar with the 'law of charges', namely

- *Like charges repel.*
- *Unlike charges attract.*

However, this law is purely qualitative and does not give any indication as to the magnitude of the forces.

In the late eighteenth century, the French scientist Charles Coulomb investigated the magnitude of the force between charges, and how this varies with the charges involved and the distance between them. He discovered the following rule.

www.iranino.co

IRANINOVIN

The force between two point charges is proportional to the product of the charges and inversely proportional to the square of the distance between them. This is known as **Coulomb's law**.

Coulomb's experiments made use of small, charged insulated spheres. Strictly speaking, the law applies to point charges, but it can be used for charged spheres provided that their radii are small compared with their separation.

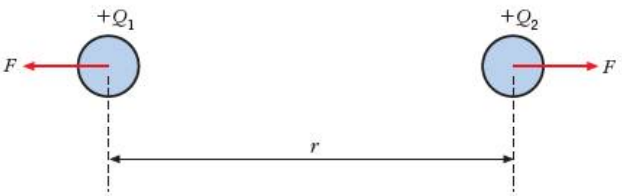


Figure 17.2 Force between charged spheres

For point charges Q_1 and Q_2 a distance r apart (Figure 17.2), Coulomb's law gives the force F as

$$F \propto Q_1Q_2/r^2$$

or

$$F = \frac{kQ_1Q_2}{r^2}$$

where k is a constant of proportionality, the value of which depends on the insulating medium around the charges and the system of units employed. In SI units, F is measured in newtons, Q in coulombs and r in metres. Then the constant k is given as

$$k = \frac{1}{4\pi\varepsilon_0}$$

and so

$$F = \frac{Q_1Q_2}{4\pi\varepsilon_0 r^2}$$

when the charges are in a vacuum. The quantity ε_0 is called the **permittivity of free space** (or the permittivity of a vacuum).

The value of the permittivity of air is very close to that of a vacuum ($1.0005\varepsilon_0$), so the equation can be used for the force between charges in a vacuum or in air.

The value of the permittivity of free space is given as

$$\varepsilon_0 = 8.85 \times 10^{-12}\ \text{C}^2\ \text{N}^{-1}\ \text{m}^{-2}$$

(We shall meet an alternative unit for ε_0 later, that is F m^{-1}.) This numerical value for ε_0 means that

$$k = \frac{1}{4\pi\varepsilon_0} = 8.99 \times 10^9\ \text{C}^{-2}\ \text{N}\ \text{m}^2 \approx 9 \times 10^9\ \text{C}^{-2}\ \text{N}\ \text{m}^2$$

Coulomb's law is often referred to as an **inverse square** law of force, because the variation of force with distance r between the charges is proportional to $1/r^2$. We have met another important inverse square law of force when we considered the gravitational force between two point masses (Topic 8).

Example

Calculate the force between two point charges, each of 1.0×10^{-9} C, which are 4.0 cm apart in a vacuum.

Using $F = kQ_1Q_2/r^2$, $F = 9.0 \times 10^9 \times (1.0 \times 10^{-9})^2/(4.0 \times 10^{-2})^2$

$= \mathbf{5.6 \times 10^{-6}\ N}$

www.iranino.co

IRANINOVIN

Now it's your turn

1 Calculate the force between two electrons which are 1.0×10^{-10} m apart in a vacuum. (charge e on electron = -1.6×10^{-19} C)

17.4 Electric field strength due to a point charge

The **electric field strength** at a point is defined as the force per unit charge acting on a small stationary positive charge placed at that point.

The electric field due to an isolated point charge is **radial** (see Figure 17.1). We have to mention that the point charge is isolated. If any other object, charged or otherwise, is near it, the field would be distorted.

From Coulomb's law, the force on a test charge q a distance r from the isolated point charge Q is given by

$$F = Qq/4\pi\varepsilon_0 r^2$$

The electric field E at the location of the test charge q is given by $E = F/q$. Thus, the electric field due to the isolated point charge is

$$E = \frac{Q}{4\pi\varepsilon_0 r^2}$$

or

$$E = \frac{kQ}{r^2} \text{ where } k = 1/4\pi\varepsilon_0$$

This is an inverse square law field of force because the force varies in proportion to $1/r^2$.

Example

In a simplified model of the hydrogen atom, the electron is at a distance of 5.3×10^{-11} m from the proton. The proton charge is $+1.6 \times 10^{-19}$ C. Calculate the electric field strength of the proton at this distance.

Assuming that the field is radial,

$E = kQ/r^2 = 9.0 \times 10^9 \times 1.6 \times 10^{-19}/(5.3 \times 10^{-11})^2$

$= \mathbf{5.1 \times 10^{11}\ N\ C^{-1}}$

Now it's your turn

2 A Van de Graaff generator has a dome of radius 15 cm. The dome carries a charge of 2.5×10^{-6} C. It can be assumed that this charge acts as if it were all concentrated at the centre of the spherical dome. Calculate the electric field strength at the dome's surface.

17.5 Electric potential energy and electric potential

We already know how useful the concept of energy is in dealing with mechanical problems. Energy is a conserved quantity and is thus a fundamental aspect of nature. We can also use the idea of energy in electricity.

We define electric potential energy in exactly the same way as we do for other types of potential energy. That is, the change in electric potential energy when a charge Q is moved between two points, A and B, in an electric field is the work done by the electric force in moving the charge from B back to A.

www.iranino.co

IRANINOVIN

We defined electric field as the force per unit positive charge. Similarly, electric potential is defined as the potential energy per unit positive charge. The symbol for potential is V, and its unit is the volt. If a point charge Q has potential energy E_{PA} at a point A, the electric potential at that point is

$$V_A = E_{PA}/Q$$

We already know that only *differences* in potential energy are measurable. We need to specify a convenient reference point to act as a zero of potential energy and potential. In dealing with gravitational energy we often take the floor of the laboratory or the Earth's surface as zero, and measure mgh from one of these. Similarly, in electrical problems, it is often convenient to take earth potential as zero, especially if part of the circuit is earthed. But the 'official' definition of the zero of electric potential is the potential of a point an infinite distance away. This means that

The **electrical potential** *at a point in an electric field is defined as the work done per unit positive charge in bringing a small test charge from infinity to the point.*

There is an important link between the electric field at a point and the electric potential at that point.

The electric field strength is equal to the negative of the potential gradient at that point.

Thus, if we have a graph showing how the potential changes in a field, the gradient of this graph at any point gives us the numerical value of the field strength at that point (Figure 17.3).

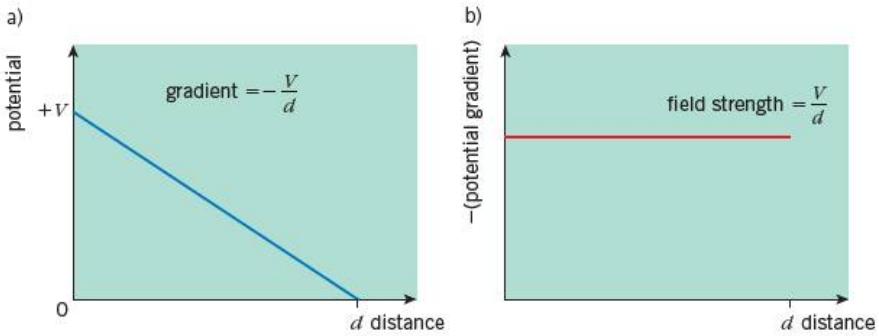


Figure 17.3 Graphs of the potential and the (negative of the) potential gradient for a uniform electric field.

This link leads to an expression for the potential at points in the field of a point charge Q. The potential V is given by

$$V = \frac{kQ}{r} = \frac{Q}{4\pi\varepsilon_0 r}$$

The potential very near a positive charge is large, and decreases towards zero the further away we move from the charge which is the source of the field. If the charge producing the field is negative, the potential is also negative and increases towards zero at large distances. Note that the variation of potential with distance is an inverse proportionality, and not the inverse square relationship that applies for the variation of field strength with distance.

Example

How much work must be done by an external force in moving a charge q of +2.0 μC from infinity to a point A, 0.40 m from a charge Q of +30 μC?

The work is simply the change in electric potential energy. The potential energy at infinity is zero, so

$W = qV = qkQ/r = 2.0 \times 10^{-6} \times 9.0 \times 10^{9} \times 30 \times 10^{-6}/0.40 =$ **1.4J**

(assuming $1/4\pi\varepsilon_0 = 9.0 \times 10^{9}$ mF^{-1})

www.iranino.co

IRANINOVIN

Now it's your turn

3 Two +30 μC charges are placed on a straight line 0.40 m apart. A +0.5 μC charge is to be moved a distance of 0.10 m along the line from a point midway between the charges. How much work must be done?

Electric fields and gravitational fields

Electric fields and gravitational fields, whilst being different in their natures, do have some features in common. These include

- Both can be represented by field lines.
- Gravitational fields are always attractive, whereas electric fields can be attractive or repulsive.
- For spherical bodies, the charge/mass on the sphere behaves as a point charge/mass at its centre.
- The field strength due to a point charge/mass follows an inverse square law with distance from the charge/mass.
- Gravitational potential is always negative, whereas electric potential can be positive or negative.
- The potential due to a point charge/mass varies inversely with distance from the charge/mass.

Summary

- Like charges repel; unlike charges attract each other.
- The force between two point charges is proportional to the product of the charges and inversely proportional to the square of the distance between them. This is Coulomb's law: $F = Q_1Q_2/4\pi\varepsilon_0 r^2$
- ε_0 is the permittivity of free space; its value is 8.85×10^{-12} C^2 N^{-1} m^{-2} (F m^{-1}).
- An electric field is a region of space where a stationary charge experiences a force.
- Electric field strength is the force per unit positive charge $E = F/Q$
- The electric field strength at a point in the field of an isolated point charge is given by $E = Q/4\pi\varepsilon_0 r^2$
- The electric potential at a point in an electric field is the work done per unit positive charge in bringing a small test charge from infinity to the point.
- The potential at a point in the field of an isolated point charge is given by $V = Q/4\pi\varepsilon_0 r$

Examination style questions

1 In a simplified model, a uranium nucleus is a sphere of radius 8.0×10^{-15} m. The nucleus contains 92 protons (and rather more neutrons). The charge on a proton is 1.6×10^{-19} C. It can be assumed that the charge of these protons acts as if it were all concentrated at the centre of the nucleus. The nucleus releases an α-particle containing two protons (and two neutrons) at the surface of the nucleus. Calculate

a the electric field strength at the surface of the nucleus before emission of the α-particle,

b the electric force on the α-particle at the surface of the nucleus,

c the electric potential at the surface of the nucleus before emission of the α-particle,

d the electric potential energy of the α-particle when it is at the surface of the nucleus.

2 Two point charges of +2.4 μC and −2.9 μC are placed at points A and B respectively in a vacuum. The distance AB is 0.15 m, as shown in Fig. 17.4.

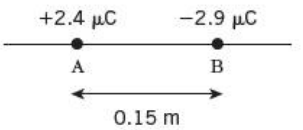


Fig. 17.4

It is required to find a point P at which the resultant electric field due to these two charges is zero.

a Explain why the point P must lie along the line AB, extended if necessary.

b Deduce the position of point P.

www.iran

IRANINOVIN

3 Two point charges A and B each have a charge of $+6.4 \times 10^{-19}$ C. They are separated in a vacuum by a distance of 12.0 μm, as shown in Fig. 17.5.

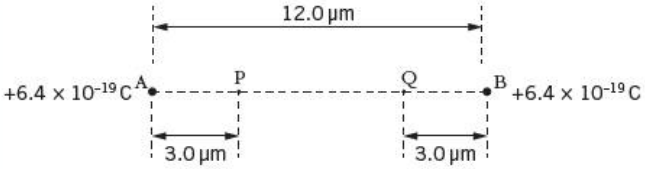


Fig. 17.5

Points P and Q are situated on the line AB. Point P is 3.0 μm from charge A and point Q is 3.0 μm from charge B.

a Calculate the force of repulsion between the charges A and B. *[3]*

b Explain why, without any calculation, when a small test charge is moved from point P to point Q, the net work done is zero. *[2]*

c Calculate the work done by an electron in moving from the midpoint of line AB to point P. *[4]*

Cambridge International AS and A Level Physics, 9702/42 May/June 2010 Q 4

4 In a simplified model of the hydrogen atom, the electron is separated from the proton by a distance r of 5.3×10^{-11} m. Use the following data to calculate the ratio of the electric force to the gravitational force at this distance. What would be the ratio if the charges were separated to a distance $2r$?

electron charge $e = -1.6 \times 10^{-19}$ C

electron mass $m_e = 9.1 \times 10^{-31}$ kg

proton mass $m_p = 1.7 \times 10^{-27}$ kg

permittivity of free space $\varepsilon_0 = 8.9 \times 10^{-12}$ F m^{-1}

Gravitational constant $G = 6.7 \times 10^{-11}$ N m^2 kg^{-2}

5 Two small charged metal spheres A and B are situated in a vacuum. The distance between the centres of the spheres is 12.0 cm, as shown in Fig. 17.6.

The charge on each sphere may be assumed to be a point charge at the centre of the sphere.

Point P is a movable point that lies on the line joining the centres of the spheres and is distance x from the centre of sphere A.

The variation with distance x of the electric field strength E at point P is shown in Fig. 17.7.

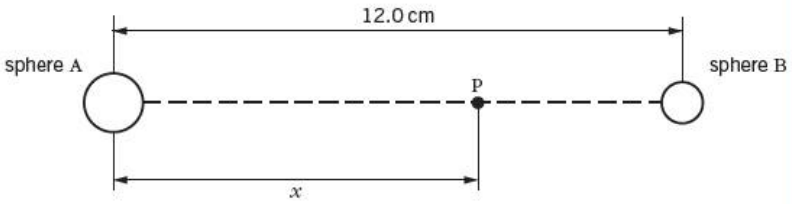


Fig. 17.6 (not to scale)

a State the evidence provided by Fig. 17.7 for the statements that

i the spheres are conductors, *[1]*

ii the charges on the spheres are either both positive or both negative. *[2]*

b i State the relation between electric field strength E and potential gradient at a point. *[1]*

ii Use Fig. 17.7 to state and explain the distance x at which the rate of change of potential with distance is

- maximum, *[2]*
- minimum. *[2]*

Cambridge International AS and A Level Physics, 9702/41 Oct/Nov 2011 Q 4

www.iranino.co

IRANINOVIN

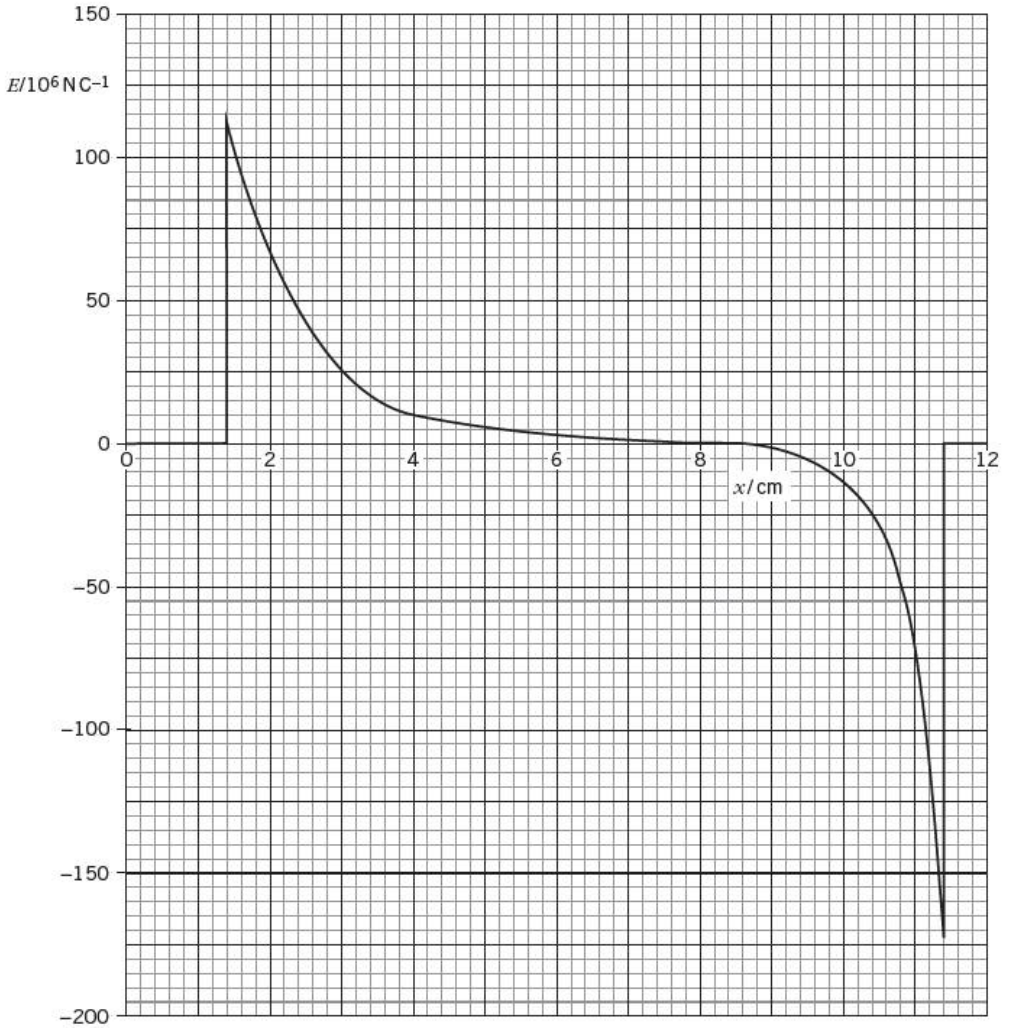


Fig. 17.7

www.iranino.co

IRANINOVIN

18 Capacitance

By the end of this topic, you will be able to:

18.1 **(a)** define capacitance and the farad, as applied to both isolated conductors and to parallel plate capacitors
(b) recall and use $C = Q/V$
(c) derive, using the formula $C = Q/V$, conservation of charge and the addition of potential differences, formulae for the combined capacitance for capacitors in series and in parallel
(d) solve problems using the capacitance formulae for capacitors in series and in parallel

18.2 **(a)** deduce, from the area under a potential–charge graph, the equation $W = \frac{1}{2}QV$ and hence $W = \frac{1}{2}CV^2$
(b) show an understanding of the function of capacitors in simple circuits

Starting points

- When charge is put onto an isolated conductor, the potential of the conductor rises.
- Capacitance relates the increase in charge to the rise in potential.
- The increase in potential of the conductor implies that energy is being stored.

18.1 Capacitors and capacitance

Capacitance

Consider an isolated spherical conductor connected to a high voltage supply (Figure 18.1).

It is found that, as the potential of the sphere is increased, the charge stored on the sphere also increases. The graph showing the variation of charge Q on the conductor with potential V is shown in Figure 18.2.

It can be seen that charge Q is related to potential V by

$$Q \propto V$$

Hence,

$$Q = CV$$

where C is a constant which depends on the size and shape of the conductor. C is known as the capacitance of the conductor.

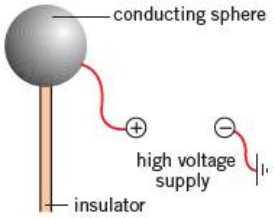


Figure 18.1 Charged spherical conductor

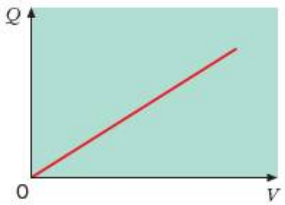


Figure 18.2 Relation between charge and potential

Capacitance *is the ratio of charge to potential for a conductor.*

$$C = \frac{Q}{V}$$

Another chance for confusion! The letter C is used as an abbreviation for the unit of charge, the coulomb. As an italic letter C it is used as the symbol for capacitance.

The unit of capacitance is the ***farad*** *(symbol F). One farad is one coulomb per volt.*

www.iranino.co

IRANINOVIN

The farad is an inconveniently large unit. In electronic circuits and laboratory experiments, the range of useful values of capacitance is from about 10^{-12} F (1 picofarad, or 1 pF) to 10^{-3} F (1 millifarad, or 1 mF). (See Topic 1 for a list of decimal multiples and submultiples for use with units.)

Note that capacitance cannot apply to an insulator. When charge is placed on the sphere in Figure 18.1, the charge distributes itself so that there is one value of potential for the whole sphere. To define capacitance, the whole sphere must have the same value of potential. With an insulator, the charge would not be able to move, so that there would be different potentials at different points on the insulator. Thus, capacitance does not apply to an insulator.

Circuit components which store charge and therefore have capacitance are called **capacitors**.

Example

Calculate the charge stored on an isolated conductor of capacitance 470 μF when it is at a potential of 20 V.

Using $C = Q/V$, we have $Q = CV = 470 \times 10^{-6} \times 20$

$= \mathbf{9.4 \times 10^{-3}\,C}$.

Now it's your turn

1 The charge on a certain isolated conductor is 2.5×10^{-2} C when its potential is 25 V. Calculate its capacitance.

Capacitors

The simplest capacitor in an electric circuit consists of two metal plates, with an air gap between them which acts as an insulator. This is called a parallel-plate capacitor. Figure 18.3a shows the circuit symbol for a capacitor. When the plates are connected to a battery, the battery transfers electrons from the plate connected to the positive terminal of the battery to the plate connected to the negative terminal. Thus the plates carry equal but opposite charges.

The **capacitance** of a capacitor is defined as the charge stored on one plate per unit potential difference between the plates.

Note that there are equal but opposite charges on the two plates. Thus, a capacitor does not store charge. We shall see later that a capacitor functions to store energy.

The capacitance of an air-filled capacitor can be increased by putting an insulating material, such as mica or waxed paper, between the plates. The material between the plates is called the **dielectric**. In a type of capacitor known as an **electrolytic** capacitor the dielectric is deposited by an electrochemical reaction. These capacitors must be connected with the correct polarity for their plates, or they will be damaged. The circuit symbol for an electrolytic capacitor is shown in Figure 18.3b. Electrolytic capacitors are available with capacitances up to about 1 mF.

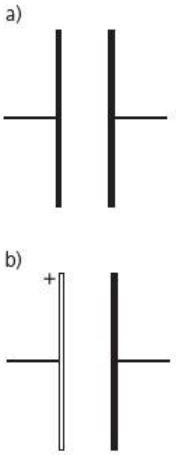


Figure 18.3 Circuit symbols for a) a capacitor and b) an electrolytic capacitor

Factors affecting capacitance

As stated previously, the material used as a dielectric affects the capacitance of a capacitor. The other factors determining the capacitance are the area of the plates and the distance between them. Experiment shows that

The capacitance is directly proportional to the area A of the plates, and inversely proportional to the distance d between them.

Putting these two factors together gives

$C \propto \frac{A}{d}$, where A is the area of one of the plates.

For a capacitor with air or a vacuum between the plates, the constant of proportionality is the permittivity of free space ε_0. Thus

$$C = \varepsilon_0 \frac{A}{d}$$

www.iranino.co

IRANINOVIN

Since C is measured in farads, A in square metres and d in metres, we can see that an alternative unit for ε_0 is farads per metre, F m^{-1} (see Topic 17), ε_0 has the value 8.85×10^{-12} F m^{-1}.

We introduce a quantity called the **relative permittivity ε_r** of a dielectric to account for the fact that the use of a dielectric increases the capacitance.

The relative permittivity is defined as the capacitance of a parallel-plate capacitor with the dielectric between the plates divided by the capacitance of the same capacitor with a vacuum between the plates.

Relative permittivity ε_r is a ratio and has no units. Some values of relative permittivity are given in Table 18.1.

Table 18.1 Relative permittivity of different dielectric materials

material	relative permittivity ε_r
air	1.0005
polyethylene (polythene)	2.3
sulfur	4
paraffin oil	4.7
mica	6
barium titanate	1200

Including the relative permittivity factor, the full expression for the capacitance of a parallel-plate capacitor is

$$C = \varepsilon_0 \varepsilon_r \frac{A}{d}$$

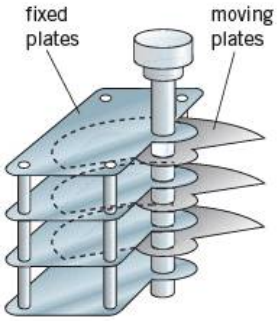


Figure 18.4 Variable capacitor

Variable capacitors (Figure 18.4) have one set of plates mounted on a spindle, so that the area of overlap can be changed.

Varying the capacitance in an electronic tuning circuit is one way of tuning in to different radio or TV stations.

Example

A parallel-plate, air-filled capacitor has square plates of side 30 cm that are a distance 1.0 mm apart. Calculate the capacitance of the capacitor.

Using $C = \varepsilon_0 \varepsilon_r A/d$, $C = 8.85 \times 10^{-12} \times 1 \times (30 \times 10^{-2})^2 / 1.0 \times 10^{-3}$

$= \mathbf{8.0 \times 10^{-10}\ F}$

Now it's your turn

2 A capacitor consists of two metal discs of diameter 15 cm separated by a sheet of polythene 0.25 mm thick. The relative permittivity of polythene is 2.3. Calculate the capacitance of the capacitor.

Capacitors in series and in parallel

In Figure 18.5, the two capacitors of capacitance C_1 and C_2 are connected in series.

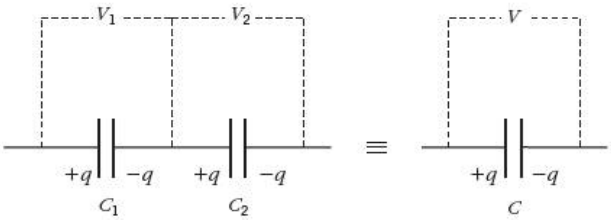


Figure 18.5 Capacitors in series

www.iranino.co

IRANINOVIN

We shall show that the combined capacitance C is given by

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2}$$

If the voltage across the equivalent capacitor is V and the charge stored on each plate is q, then $V = q/C$. The potential difference across the combination is the sum of the potential differences across the individual capacitors, $V = V_1 + V_2$, and each capacitor has charge q on each plate. A charge of $+q$ induced on one plate of one capacitor will induce a charge of $-q$ on the other plate of the capacitor. This will, in turn, induce a charge of $+q$ on one plate of the second capacitor and $-q$ on its other plate, since, initially, the capacitors were uncharged and charge is always conserved.
Since $V_1 = q/C_1$ and $V_2 = q/C_2$, then

$$q/C = q/C_1 + q/C_2$$

Dividing each side of the equation by q, we have

$$1/C = 1/C_1 + 1/C_2$$

A similar result applies for any number of capacitors connected in series.

The reciprocal of the combined capacitance equals the sum of the reciprocals of the individual capacitances in series.

Note that:
- For two identical capacitors in series, the combined capacitance is equal to half of the value of each one.
- For capacitors in series, the combined capacitance is always less than the value of the smallest individual capacitance.

In Figure 18.6, the two capacitors of capacitance C_1 and C_2 are connected in parallel. We shall show that the combined capacitance C is given by

$$C = C_1 + C_2$$

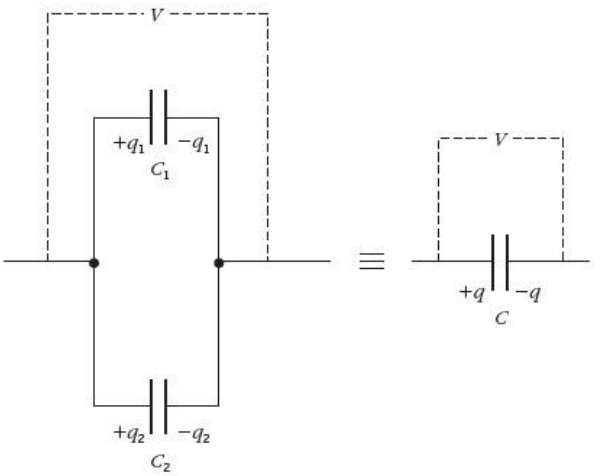


Figure 18.6 Capacitors in parallel

If the voltage across the equivalent capacitor is V and the charge stored on each plate is q, then $q = CV$. The total charge stored is the sum of the charges on the individual capacitors, $q = q_1 + q_2$, and there is the same potential difference V across each capacitor since they are connected in parallel. Since $q_1 = C_1V$ and $q_2 = C_2V$, then

$$CV = C_1V + C_2V$$

www.iranino.co

IRANINOVIN

Dividing each side of the equation by V, we have

$$C = C_1 + C_2$$

The same result applies for any number of capacitors connected in parallel.

The combined capacitance equals the sum of all the individual capacitances in parallel.

Note that the equation for capacitors in *series* is similar to the equation for resistors in *parallel*, and the equation for capacitors in parallel is similar to the equation for resistors in series (see Topic 20).

Examples

1 A 100 μF capacitor in parallel with a 50 μF capacitor is connected to a 12 V supply. Calculate:
(a) the total capacitance,
(b) the potential difference across each capacitor,
(c) the charge stored on one plate of each capacitor.

(a) Using the equation for capacitors in parallel, $C = C_1 + C_2 = 100 + 50 = \mathbf{150\,\mu F}$.
(b) The potential difference across each capacitor is the same as the potential difference across the supply. This is **12 V**.
(c) Using $Q = CV$, the charge stored on the 100 μF capacitor is $100 \times 10^{-6} \times 12 = \mathbf{1.2 \times 10^{-3}\,C}$. The charge stored on the 50 μF capacitor is $50 \times 10^{-6} \times 12 = \mathbf{6.0 \times 10^{-4}\,C}$.

2 A 100 μF capacitor in series with a 50 μF capacitor is connected to a 12 V supply. Calculate:
(a) the combined capacitance,
(b) the charge stored on one plate of each capacitor,
(c) the potential difference across each capacitor.

(a) Using the equation for capacitors in series, $1/C = 1/C_1 + 1/C_2 = 1/100 \times 10^{-6} + 1/50 \times 10^{-6} = 3 \times 10^4$.
Thus $C = 3.3 \times 10^{-5} = \mathbf{33\,\mu F}$.
(b) The charge stored by each capacitor is the same as the charge stored by the combination, so $Q = CV = 33 \times 10^{-6} \times 12 = \mathbf{4.0 \times 10^{-4}\,C}$.
(c) Using $V = Q/C$, the potential difference across the 100 μF capacitor is $(4.0 \times 10^{-4})/(100 \times 10^{-6}) = \mathbf{4.0\,V}$. The potential difference across the 50 μF capacitor is $4.0 \times 10^{-4}/50 \times 10^{-6} = \mathbf{8.0\,V}$.
Note that the two potential differences add up to the supply voltage.

Now it's your turn

3 (a) A 470 μF capacitor is connected to a 20 V supply. Calculate the charge stored on one plate of the capacitor.
(b) The capacitor in **a** is now disconnected from the supply and connected to an uncharged 470 μF capacitor.
(i) Explain why the capacitors are in parallel, rather than series, and why the total charge stored by the combination must be the same as the answer to **a**.
(ii) Calculate the capacitance of the combination.
(iii) Calculate the potential difference across each capacitor.
(iv) Calculate the charge stored on one plate of each capacitor.

18.2 Energy stored in a capacitor

When charging a capacitor, work is done by the battery to move charge on to the capacitor. Energy is transferred from the power supply and is stored as **electric potential energy** in the capacitor.

Camera flash units use a capacitor to store energy. The capacitor takes a few seconds to charge when connected to the battery in the camera. Then the energy is discharged very rapidly when the capacitor is connected to the flash-bulb to give a short but intense flash.

www.iranino.co

IRANINOVIN

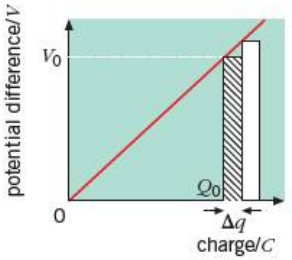


Figure 18.7 Graph of potential against charge for a capacitor

Since $Q = CV$, the charge stored on each plate of a capacitor is directly proportional to its potential (see Figure 18.7).

From the definition of potential, the work done (and therefore the energy transferred) is the product of the potential and the charge. That is,

$$W \text{ (and } E_p) = VQ$$

However, while more and more charge is transferred to the capacitor, the potential difference is increasing. Suppose the potential is V_0 when the charge stored is Q_0. When a further small amount of charge Δq is supplied, the energy transferred is given by

$$\Delta E_p = V_0 \Delta q$$

which is equal to the area of the hatched strip in Figure 18.7. Similarly, the energy transferred when a further charge Δq is transferred is given by the area of the next strip, and so on. If the amount of charge Δq is very small, the strips become very thin and their combined areas are just equal to the area between the line and the horizontal axis. Thus

The energy transferred from the battery when a capacitor is charged is given by the area under the graph line when charge (x-axis) is plotted against potential difference (y-axis).

Because the graph is a straight line through the origin, this area is just the area of the right-angled triangle formed by the line and the charge axis. Thus

$$E_p = \tfrac{1}{2}QV$$

This is the expression for the energy transferred from the battery in charging the capacitor. This is electric potential energy, and it is released when the capacitor is discharged. Since $C = Q/V$, this expression can be written in different forms.

$$E_p = \tfrac{1}{2}QV = \tfrac{1}{2}CV^2 = \frac{Q^2}{2C}$$

Examples

1 Calculate the energy stored by a 470 μF capacitor at a potential of 20 V.

Using $E_p = \frac{1}{2}CV^2$, $E_p = \frac{1}{2} \times 470 \times 10^{-6} \times (20)^2$ = **0.094 J.**

2 A camera flash-lamp uses a 5000 μF capacitor which is charged by a 9 V battery. The capacitor is then disconnected from the battery. Calculate the energy transferred when the capacitor is discharged through the lamp so that the final potential difference across its plates is 6.0 V.

Energy change = $\frac{1}{2}CV_1^2 - \frac{1}{2}CV_2^2$

= $\frac{1}{2} \times 5000 \times 10^{-6} \times (9^2 - 6^2)$

= **0.113 J**

Note: $(V_1^2 - V_2^2)$ is not equal to $(V_1 - V_2)^2$ – a common mistake amongst students!

Now it's your turn

4 A camera flash-lamp uses a 5000 μF capacitor which is charged by a 9 V battery. Calculate the energy transferred when the capacitor is fully discharged through the lamp.

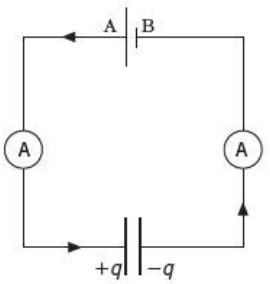


Figure 18.8 Charging a capacitor

www.iranino.com

Charging and discharging capacitors

The circuit in Figure 18.8 can be used to investigate what happens when a capacitor is connected to a battery. Start with the lead to battery terminal A connected, but that to B disconnected. When connection is first made to terminal B, both ammeters flick to the right and then return to zero, indicating a momentary pulse of current. If the process is repeated, nothing further happens. The capacitor has become charged.

During the initial connection, the ammeters have similar deflections. This tells us that the same charge has moved on to the left-hand plate as has been removed from the right-hand plate. The momentary current has caused the left-hand plate to acquire

IRANINOVIN

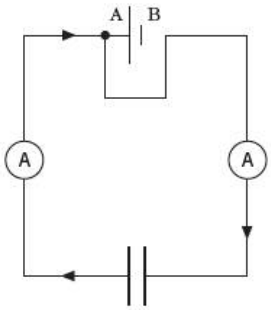


Figure 18.9 Discharging a capacitor

a charge of $+q$, while the right-hand plate has a charge of $-q$. The capacitor is now said to store a charge of q (although if we add up the charges on the left- and right-hand plates taking account of their sign, the net charge is zero). There can be no steady current because of the gap between the capacitor plates. The current stops when the potential difference across the capacitor is the same as the e.m.f. of the battery.

When the battery lead to terminal B is disconnected and joined to point A (Figure 18.9) so that the battery is no longer in circuit, both ammeters give similar momentary deflections to the left. This time, a current in the opposite direction has moved the charge of $+q$ from the left-hand plate to cancel the charge of $-q$ on the right-hand plate. The capacitor has become discharged.

Remember that in metal wires the current is carried by free electrons. These move in the opposite direction to that of the conventional current (see Topic 19). When the capacitor is charged, electrons move from the negative terminal of the battery to the right-hand plate of the capacitor, and from the left-hand plate to the positive terminal of the battery. When the capacitor is discharged, electrons flow from the negative right-hand plate of the capacitor to the positive left-hand plate.

The experiment described using the circuit in Figure 18.9 showed that there is a momentary current when a capacitor discharges. A resistor connected in series with the capacitor will reduce the current, so that the capacitor discharges more slowly.

The circuit shown in Figure 18.10 can be used to investigate more precisely how a capacitor discharges. When the two-way switch is connected to point A, the capacitor will charge up until the potential difference between its plates is equal to the e.m.f. V_0 of the supply. When the switch is moved to B, the capacitor will discharge through the resistor. When the switch makes contact with B, the current can be recorded at regular intervals of time as the capacitor discharges.

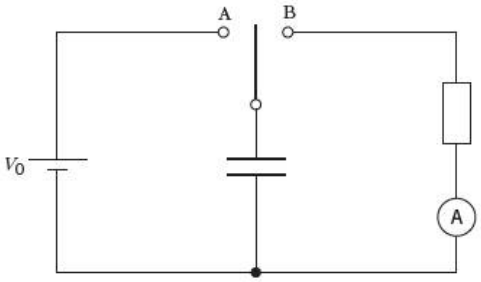


Figure 18.10 Circuit for investigating capacitor discharge

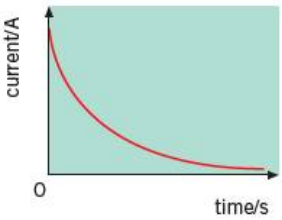


Figure 18.11 Graph of current against time for capacitor discharge

A graph of the discharge current against time is shown in Figure 18.11.

The current is seen to change rapidly at first, and then more slowly. More detailed analysis shows that the curve is **exponential**. We shall meet exponential changes again when we deal with the decay of radioactive substances (Topic 26).

All exponential decay curves have an equation of the form

$$x = x_0 e^{-kt}$$

where x is the quantity that is decaying (and x_0 is the value of x at $t = 0$), e to three decimal places is the number 2.718 (the root of natural logarithms) and k is a constant characteristic of the decay. A large value of k means that the decay is rapid, and a small value means a slow decay.

The solution for the discharge of a capacitor of capacitance C through a resistor of resistance R is of the form

$$Q = Q_0 e^{-t/CR}$$

This is the equation of the graph in Figure 18.12. Here Q_0 is the charge on the plates at time $t = 0$, and Q is the charge at time t.

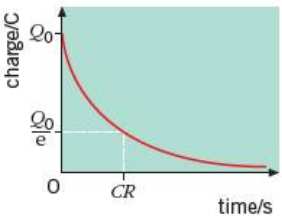


Figure 18.12 Graph of charge against time for capacitor discharge

www.iranino.co

IRANINOVIN

The graphs of Figures 18.11 and 18.12 have exactly the same shape, and thus the equation for the discharge of a capacitor may be written

$$I = I_0 e^{-t/CR}$$

Furthermore, since for a capacitor Q is proportional to V, then the solution for the discharge can be written as

$$V = V_0 e^{-t/CR}$$

Time constant

As time progresses, the exponential curve in Figure 18.12 gets closer and closer to the time axis, but never actually meets it. Thus, it is not possible to quote a time for the capacitor to discharge completely.

However, the quantity CR in the decay equation may be used to give an indication of whether the decay is fast or slow, as shown in Figure 18.13

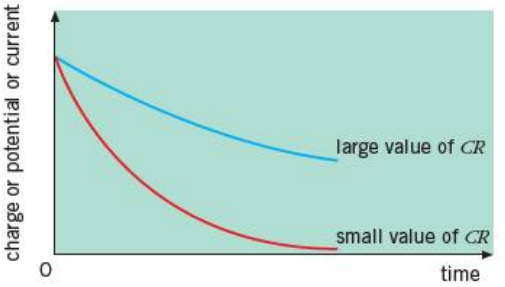


Figure 18.13 Decay curves for large and small time constants

CR is called the **time constant** *of the capacitor–resistor circuit.*

CR has the units of time, and is measured in seconds.

We can easily show that CR has units of time. From $C = Q/V$ and $R = V/I$, then $CR = Q/I$. Since charge Q is in coulombs and current I is in amperes, and one ampere is equal to one coulomb per second, Q/I is in seconds.

To find the charge Q on the capacitor plates after a time $t = CR$, we substitute in the exponential decay equation

$$Q = Q_0 e^{-CR/CR} = Q_0 e^{-1} = Q_0/e = Q_0/2.718$$

Thus

The time constant is the time for the charge to have decreased to 1/e (or 1/2.718) of its initial charge.

In one time constant the charge stored by the capacitor drops to roughly one-third of its initial value. During the next time constant it will drop by the same ratio, to about one-ninth of the value at the beginning of the decay.

Example

A 500 μF capacitor is connected to a 10 V supply, and is then discharged through a 100 kΩ resistor. Calculate:

(a) the initial charge stored by the capacitor,
(b) the initial discharge current,
(c) the value of the time constant,
(d) the charge on the plates after 100 s,
(e) the time at which the remaining charge is 2.5×10^{-3} C.

(a) From $Q = CV$, we have $Q = 500 \times 10^{-6} \times 10 = \mathbf{5.0 \times 10^{-3}\ C}$.
(b) From $I = V/R$, we have $I = 10/(100 \times 10^3) = \mathbf{1.0 \times 10^{-4}\ A}$.
(c) $CR = 500 \times 10^{-6} \times 100 \times 10^3 = \mathbf{50\ s}$.

www.iranino.co

IRANINOVIN

(d) After 50 s, the charge on the plates is $Q_0/e = 5 \times 10^{-3}/2.718 = 1.8 \times 10^{-3}$ C; after a further 50 s, the charge is $1.8 \times 10^{-3}/2.718 = \mathbf{6.8 \times 10^{-4}}$ **C**.

(e) Using $Q = Q_0e^{-t/CR}$, $2.5 \times 10^{-3} = 5.0 \times 10^{-3}e^{-t/50}$, or $0.50 = e^{-t/50}$. Taking natural logarithms of both sides, $-0.693 = -t/50$, or $t = \mathbf{35\ s}$.

Now it's your turn

5 A 5.0 μF capacitor is charged from a 12 V battery, and is then discharged through a 0.50 MΩ resistor. Calculate
(a) the initial charge on the capacitor,
(b) the charge on the capacitor 2.0 s after the discharge starts,
(c) the potential difference across the capacitor at this time.

Some uses of capacitors

Some uses of capacitors have already been outlined. These, and others, are listed below.

- Capacitors can be used to store charge on, perhaps, an isolated sphere. Alternatively, by building up the charge on a sphere, high potentials can be achieved (the van de Graaff generator).
- Capacitors are used to store electrical energy. Discharging the capacitor means that the energy can be released in a short time as in, for example, a flash gun.
- A capacitor allows the storage of separated charges. The total charge stored is, in fact, zero but a capacitor separates the positive and negative charges to allow charge to be stored.
- Capacitors may be used to prevent arcing. Rather than the charge creating a spark, the charge is stored in a capacitor. The capacitor is discharged gradually at a later time.
- When used in conjunction with a coil, the capacitor can form part of a circuit that produces electrical oscillations or, under different circumstances, can be used to tune a circuit to particular radio frequencies.
- As seen in Topic 24, capacitors may be used in a smoothing circuit to reduce the ripple on rectified current or voltage.

Summary

- A capacitor stores energy. Its capacitance C is given by $C = Q/V$, where Q is the charge on the capacitor when there is a potential difference V between its plates.
- A capacitor allows the storage of separated charges.
- The unit of capacitance, the farad (F), is one coulomb per volt.
- The energy stored in a charged capacitor is given by $E = \frac{1}{2}QV = \frac{1}{2}CV^2 = \frac{1}{2}Q^2/C$
- The equivalent capacitance C of capacitors connected in series is given by: $1/C = 1/C_1 + 1/C_2 + \ldots$
- The equivalent capacitance C of capacitors connected in parallel is given by: $C = C_1 + C_2 + \ldots$
- When a charged capacitor discharges, the charge on the plates decays exponentially. The equation for the decay is $Q = Q_0e^{-t/CR}$
- The time constant of the circuit, given by CR, is the time for the charge to decay to 1/e of its initial value.

www.iranino.co

IRANINOVIN

Examination style questions

1 A parallel-plate capacitor has rectangular plates of side 200 mm by 30 mm. The plates are separated by an air gap 0.50 mm thick. The permittivity of free space is 8.85×10^{-12} F m^{-1}.

a Calculate the capacitance of the capacitor.

b The capacitor is connected to a 12 V battery. Calculate:

i the charge on each plate,

ii the electric field between the plates.

2 Fig. 18.14 shows an arrangement of capacitors.

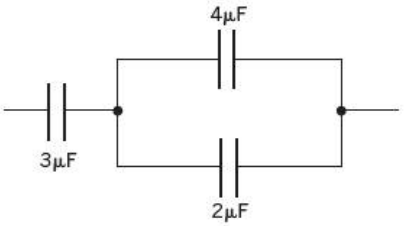


Fig. 18.14

a Calculate the capacitance of this arrangement.

b The 4 μF capacitor is disconnected. Calculate the new capacitance.

3 A 10 μF capacitor is charged from a 9.0 V battery.

a Calculate:

i the electric potential energy stored by the capacitor,

ii the charge stored by the capacitor.

b The charged capacitor is discharged through a 150 kΩ resistor. Calculate

i the initial discharge current,

ii the time constant,

iii the time taken for the current to fall to 3.6×10^{-5} A.

4 a i Define *capacitance*. *[1]*

ii A capacitor is made of two metal plates, insulated from one another, as shown in Fig. 18.15.

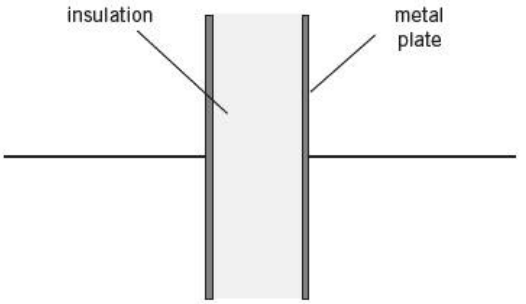


Fig. 18.15

Explain why the capacitor is said to store energy but not charge. *[4]*

b Three uncharged capacitors X, Y and Z, each of capacitance 12 μF, are connected as shown in Fig. 18.16.

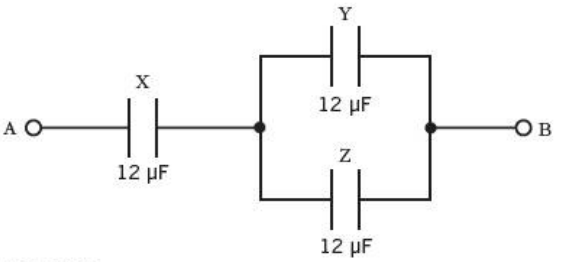


Fig. 18.16

A potential difference of 9.0 V is applied between points A and B.

i Calculate the combined capacitance of the capacitors X, Y and Z. *[2]*

ii Explain why, when the potential difference of 9.0 V is applied, the charge on one plate of capacitor X is 72 μC. *[2]*

iii Determine

- the potential difference across capacitor X, *[1]*
- the charge on one plate of capacitor Y. *[2]*

Cambridge International AS and A Level Physics, 9702/42 Oct/Nov 2012 Q 5

5 A capacitor consists of two metal plates separated by an insulator, as shown in Fig. 18.17.

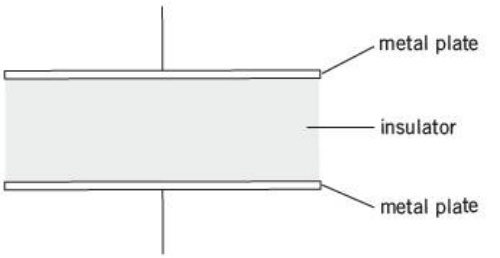


Fig. 18.17

The potential difference between the plates is V. The variation with V of the magnitude of the charge Q on one plate is shown in Fig. 18.18.

a Explain why the capacitor stores energy but not charge. *[3]*

www.iranino.co

IRANINOVIN

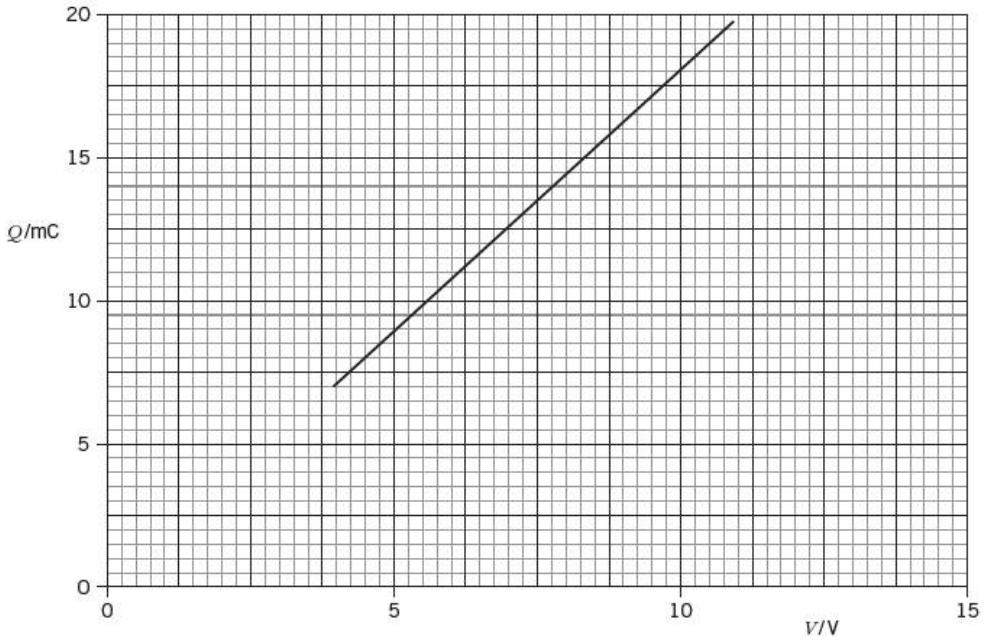


Fig. 18.18

b Use Fig. 18.18 to determine

i the capacitance of the capacitor. *[2]*

ii the loss in energy stored in the capacitor when the potential difference V is reduced from 10.0V to 7.5V. *[2]*

Cambridge International AS and A Level Physics, 9702/42 May/June 2011 Q 3

www.iranino.co

IRANINOVIN

19 & 20 Electronic sensors

By the end of this topic, you will be able to:

19.4 **(a)** show an understanding of the change in resistance with light intensity of a light-dependent resistor (LDR)
(b) sketch the temperature characteristic of a negative temperature coefficient thermistor
(c) show an understanding of the action of a piezo-electric transducer and its application in a simple microphone
(d) describe the structure of a metal-wire strain gauge
(e) relate extension of a strain gauge to change in resistance of the gauge

20.3 **(c)** understand that an electronic sensor consists of a sensing device and a circuit that provides an output that can be registered as a voltage
(d) explain the use of thermistors, light-dependent resistors and strain gauges in potential dividers to provide a potential difference that is dependent on temperature, illumination and strain respectively

Starting points

- An understanding of basic concepts of electric current.
- An understanding of basic electrical circuitry.

Introduction

Electronic sensors have become such everyday items that we hardly appreciate them – except when they fail to work! Some sensors are simple, such as those detecting temperature in a thermostat or detecting light levels in an automatic switch for a lamp. Other sensors may be more complex and used to control systems. For example, in the braking system of a car, an electronic circuit may be used to sense whether a wheel is skidding. If the car is skidding, then the circuit controls the brakes, not only to stop the skid, but also to prevent further skidding.

An electronic sensor may be thought to consist of three parts: a sensing device, a processing unit and an output device, as illustrated in Figure 19.1.



Figure 19.1 Block diagram for an electronic sensor

19.4 Sensing devices

The **sensing device** is the first stage of any electronic sensor and is the means by which whatever is to be detected or monitored is converted into an electrical property. This electrical property is frequently resistance. A physical property of the sensing device changes with a change in whatever it is monitoring.

www.iranino.co

IRANINOVIN

Sensing devices are available that can be used to detect many different types of change – for example, changes in light intensity, temperature, sound level, humidity, pressure, strain, or magnetic field. Some of the more common sensing devices will be considered.

The light-dependent resistor (LDR)

A light-dependent resistor (LDR) consists of two metal grids that intersect each other. The space between the grids is filled with a semiconductor material, for example, cadmium sulfide doped with copper, as shown in Figure 19.2.

When light is incident on the semiconductor material, the number of electrons in the semiconductor that are free to conduct increases. The higher the intensity of light on the LDR, the greater the number of electrons that can move freely. Hence, as the light intensity increases, the resistance of the LDR decreases. Figure 19.3 shows the variation with light intensity of the resistance of a typical LDR. The variation of resistance with light intensity is similar to that of the variation of resistance with temperature for the thermistor.



Figure 19.2 An LDR and its symbol

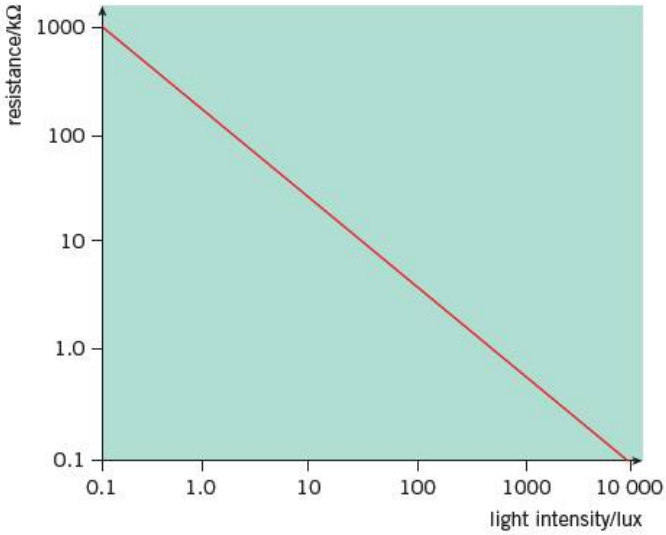


Figure 19.3 Resistance v. light intensity for an LDR

Note: Both light intensity, measured in lux, and resistance, measured in ohms, are plotted on logarithmic scales in Figure 19.3. The graph is a straight line but this does not mean that resistance is inversely proportional to light intensity. Data relating to light intensity and resistance for a typical LDR are shown in Table 19.1.

Table 19.1 Typical LDR data

light level	illumination/lux	LDR resistance/Ω
moonlight	0.1	1×10^6
normal room lighting	450	900
sunlight	28000	100

The lux is a unit that is used to measure the light power incident per unit area of a surface.

The thermistor

The resistance of most metals increases to a certain extent with rise in temperature. Negative temperature coefficient devices, often referred to as **thermistors**, are made from semiconductor material, usually the oxides of metals. The resistance of thermistors decreases significantly with rise in temperature. Thermistors are manufactured in various shapes and sizes, including rods, discs and beads. Figure 19.4a shows an example of a disc and bead thermistor.

Data relating to the temperature and resistance of a typical bead thermistor are shown in Table 19.2.

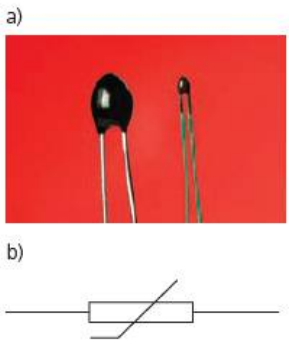


Figure 19.4 Thermistors and their symbol

www.iranino.co

IRANINOVIN

Table 19.2 Typical thermistor data

temperature/°C	thermistor resistance/Ω
1	3700
10	2500
20	1800
30	1300
40	900
50	660

The variation with temperature of a typical thermistor is shown in Figure 19.5. This variation is non-linear and is, in fact, approximately exponential over a limited range of temperature.

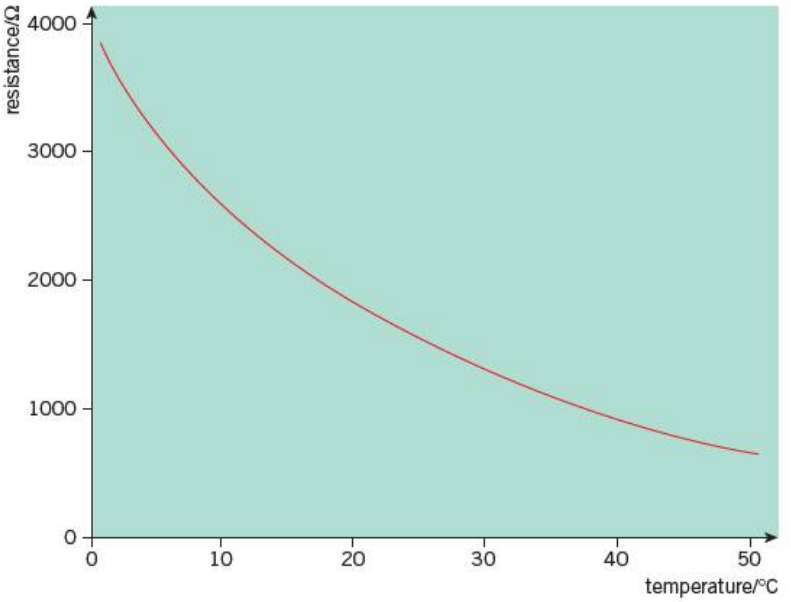


Figure 19.5 Resistance v. temperature for a thermistor

The piezo-electric transducer

A **transducer** is any device that converts energy from one form into another.

Piezo-electric crystals such as quartz have a complex ionic structure (see A level Topic 14). In the normal unstressed state of the crystal, the centre of charge of the positive ions coincides with the centre of charge of the negative ions. However, when pressure is applied to the crystal, the crystal changes shape by a small amount and the centres of the positive and the negative charge no longer coincide. A voltage is generated across the crystal. This is known as the **piezo-electric effect**.

The magnitude of the voltage that is generated depends on the magnitude of the pressure on the crystal, and its polarity depends on whether the crystal is compressed or expanded – that is, whether the pressure applied is greater or less than the ambient pressure.

A sound wave consists of a series of compressions and rarefactions. If a sound wave is incident on a piezo-electric crystal, then a voltage will be produced across the crystal that varies in a similar way to the variation in pressure of the sound wave. To detect the voltages, opposite faces of the crystal are coated with a metal and electrical connections are made to these metal films. Since the voltages generated are small, they are amplified. The crystal and its amplifier may be used as a simple microphone for converting sound signals into electrical signals. The symbol for a microphone is shown in Figure 19.6.

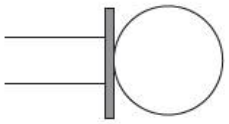


Figure 19.6 Symbol for a microphone

www.iranino.co

IRANINOVIN

The metal wire strain gauge

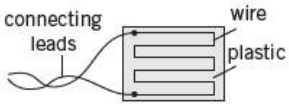


Figure 19.7 Metal wire strain gauge

In engineering, it is frequently necessary to test for the strains experienced in different parts of structures. This can be achieved using a metal wire **strain gauge**.

A strain gauge consists of a length of very fine metal wire sealed into a small rectangle of thin plastic, as shown in Figure 19.7.

In use, the strain gauge is attached firmly to the structure to be monitored. When the plastic is stretched, the fine wire will also be stretched. The wire is strained, causing its length to increase and its cross-sectional area to decrease. An increase in length and a decrease in cross-sectional area both result in an increase in the electrical resistance of the wire.

The resistance R of a uniform wire is related to its length L and area of cross-section A by the expression

$$R = \frac{\rho L}{A}$$

where ρ is the resistivity of the material of the wire (see AS Level Topic 19).

If the wire increases in length by only a small amount ΔL, then the change in the cross-sectional area may be assumed to be negligible. The new resistance of the wire is given by

$$(R + \Delta R) = \frac{\rho(L + \Delta L)}{A}$$

By subtracting these two expressions, the change in resistance ΔR is given by

$$\Delta R = \frac{\rho \Delta L}{A}$$

It is assumed that A is constant, and ρ is also a constant, so

$$\Delta R \propto \Delta L$$

Since the strain is proportional to the extension ΔL, then the strain is also proportional to the change in resistance.

Examples

1 Explain what is meant by a negative temperature coefficient thermistor.
It is an electrical device whose resistance decreases as its temperature rises.

2 Suggest why the resistance of a strain gauge changes when the object on to which it is fixed is strained.
The strain gauge consists of a grid of fine metal wire in a plastic envelope. When the envelope is stretched, the length of the wire increases and its area of cross-section decreases (slightly). Increase in length and decrease in area both cause the resistance to increase.

Now it's your turn

1 Draw a sketch graph to show the variation with temperature θ of the resistance R of a thermistor. Mark typical values on the axes of your graph.

2 (a) State what is meant by the piezo-electric effect.
(b) Describe how a quartz crystal may be used to provide an electrical signal that is dependent on the sound pressure incident on it.

20.3 The use of potential dividers

The change in the physical property of a sensing device must, in general, be processed in some way before the change can be displayed or measured. This processing is carried out by the processing unit, which is some form of electrical circuit that is connected to the sensing device and provides a voltage at its output. If the voltage is small, then it may be necessary to amplify it. This will be considered in Topic 21. The voltage can then be used to control an output device.

A potential divider circuit (see AS Level Topic 20) enables a change in resistance of a sensing device to be converted into a voltage change. In Figure 20.1, the sensing device of resistance S is connected in series with a resistor of resistance R.

www.iranino.co

IRANINOVIN

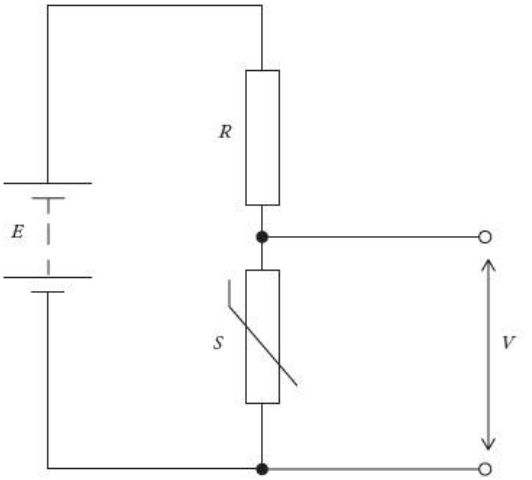


Figure 20.1 Potential divider circuit

The battery of e.m.f. E is assumed to have negligible internal resistance. The output voltage V is given by the expression

$$V = \frac{S}{S + R} \times E$$

The magnitude V of the potential difference (voltage) at any particular value S of the sensing device is dependent on the relative values of S and R. Note that, as the resistance S of the sensor increases, the output voltage V also increases.

By monitoring the potential difference (voltage) across the resistor of fixed resistance R, as shown in Figure 20.2, the output voltage will be given by

$$V = \frac{R}{S + R} \times E$$

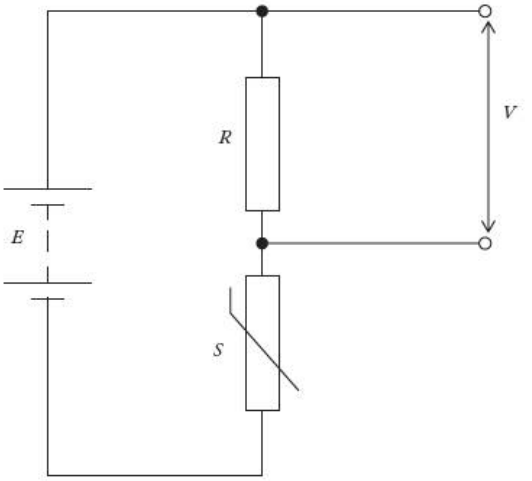


Figure 20.2 Alternative connections for a potential divider

The output voltage V will then decrease as the sensor resistance S increases.

Example

A potential divider consists of a battery of e.m.f. 6.00 V and negligible internal resistance connected in series with a resistor of resistance 120 Ω and a variable resistor of resistance 0 → 200 Ω. Determine the range of potential difference that can be obtained across the fixed resistor.

When the variable resistor is at 0 Ω, the p.d. across the fixed resistor = 6.0 V. When the variable resistor is at 200 Ω,

$$\text{p.d. across fixed resistor} = \frac{120}{120 + 200} \times 6.00 = 2.25 \text{ V}$$

The range is **2.25 V → 6.00 V**.

www.iranino.co

IRANINOVIN

Now it's your turn

3 A potential divider consists of a battery of e.m.f. 7.5V and negligible internal resistance connected in series with a resistor of resistance R and a variable resistor of resistance $0 \rightarrow 500\Omega$. Deduce how the potential divider may be arranged so as to provide a potential difference that may be varied between 0 and 3.0V.

Summary

- Sensing devices are used to detect changes in the environment or the properties of a body.
- In general, the change being monitored causes a change in an electrical property of the sensor.
- Frequently-used sensors include the light-dependent resistor (LDR), the thermistor, the piezo-electric transducer and the strain gauge.
- The resistance of an LDR decreases as the light intensity on it increases.
- The resistance of a negative temperature coefficient thermistor decreases with increasing temperature.
- A piezo-electric transducer converts variations in pressure (such as due to a sound wave) into variations in voltage.
- The strain on a metal wire strain gauge is proportional to its change in resistance.
- A potential divider may be used to convert the change in resistance of the sensor into a change in potential difference.

Examination style questions

1 Suggest suitable sensing devices to detect changes in
   a sunlight,
   b the width of a crack in a wall,
   c the weight of water in a tank,
   d the temperature of the oil in an engine.

2 Describe how a sensor may be used to count the number of mice that pass through a small hole in a wall.

3 a Suggest electrical sensing devices, one in each case, that may be used to monitor changes in
   i light intensity, [1]
   ii the width of a crack in a welded joint, [1]
   iii the intensity of an ultrasound beam. [1]

   b A student designs the circuit of Fig. 20.3 to detect changes in temperature in the range 0°C to 100°C.

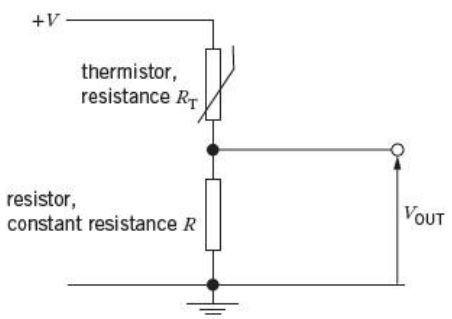


Fig. 20.3

The resistance of the thermistor is R_T and that of the resistor is R.

The student monitors the potential difference V_{OUT}.

State and explain

i whether V_{OUT} increases or decreases as the temperature of the thermistor increases, [3]

ii whether the change in V_{OUT} varies linearly with the change in temperature of the thermistor. [2]

Cambridge International AS and A Level Physics, 9702/43 May/June 2013 Q 9

www.iranino.co

IRANINOVIN

21 Electronics

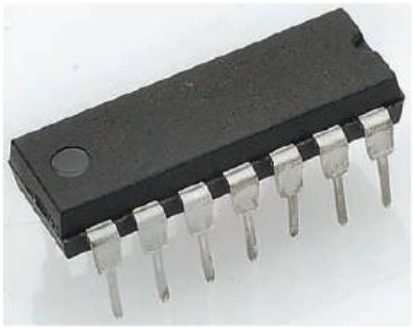
By the end of this topic, you will be able to:

21.1 **(a)** recall the main properties of an ideal operational amplifier (op-amp)

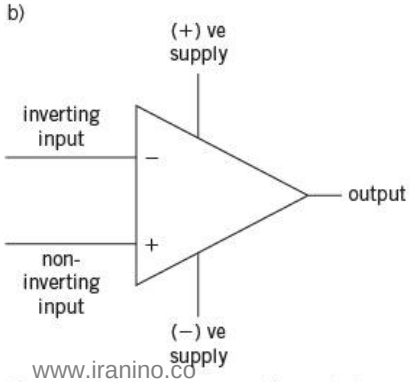
21.2 **(a)** deduce, from the properties of an ideal operational amplifier, the use of an operational amplifier as a comparator
(b) understand the effects of negative feedback on the gain of an operational amplifier
(c) recall the circuit diagrams for both the inverting and the non-inverting amplifier for single signal input
(d) understand the virtual earth approximation and derive an expression for the gain of inverting amplifiers
(e) recall and use expressions for the voltage gain of inverting and of non-inverting amplifiers

21.3 **(a)** understand that an output device may be required to monitor the output of an op-amp circuit
(b) understand the use of relays in electronic circuits
(c) understand the use of light-emitting diodes (LEDs) as devices to indicate the state of the output of electronic circuits
(d) understand the need for calibration where digital or analogue meters are used as output devices

a)



b)



www.iranino.co

Figure 21.1 An op-amp and its symbol

Starting points

- Sensing circuits can be devised that will monitor environmental conditions.
- The output of these sensing circuits can be registered as a voltage, but this voltage may be small.

21.1 The ideal operational amplifier (op-amp)

Once the output of the sensing device is in the form of a voltage, it may require further change (processing) before it is used to control an output device. The basis of many circuits used for processing the sensor voltage is the **operational amplifier** or **op-amp**.

An op-amp is an integrated circuit that contains about 20 transistors together with resistors and capacitors. It is referred to as an 'integrated circuit' because all the components are formed on a small slice of a semiconductor (e.g. silicon) with connections to enable the op-amp to be connected into a circuit. The whole of the integrated circuit is encapsulated.

The symbol for an op-amp and some of its connections are shown in Figure 21.1.

It is referred to as an 'operational amplifier' because the circuit can easily be made to carry out different operations. These operations include:

- acting as a switch when a voltage reaches a certain level
- amplifying direct voltages
- amplifying alternating voltages
- comparing two voltages and giving an output that depends on the result of the comparison.

IRANINOVIN

The op-amp has two inputs, called the inverting input (–) and the non-inverting input (+). For many applications, the positive and the negative power supplies are ±6V or ±9V.

Properties of an ideal op-amp

The ideal operational amplifier (op-amp) has the following properties:

- **infinite input impedance** (no current enters or leaves either of the inputs). Think about connecting an input to one of the resistors in a potential divider. The input would be in parallel with the resistor and, therefore, the resistance in the potential divider circuit would be changed. Connecting the op-amp to the potential divider would, thus, change the potential differences across the components in the potential divider. However, if the op-amp connection (the input) has infinite resistance (or, more strictly, for alternating voltages, infinite impedance) then connecting it to the potential divider will not affect the potential differences.
- **zero output impedance** (the whole of the output voltage is seen across the load connected to the output). If the output connection had some resistance, then the voltage produced in the op-amp (the output voltage) would be divided between the output and the resistor connected to the output. Zero output resistance (or zero output impedance) means that all of the output voltage will be seen across the resistor connected to the output.
- **infinite open-loop gain** (this means that when there is only a very small input voltage, the amplifier will **saturate** and the output will have the same value as the supply voltage). The voltage gain, or simply **gain**, of an amplifier is a measure of how many times the output voltage is greater than the input voltage.

$$\text{voltage gain} = \frac{\text{output voltage}}{\text{input voltage}}$$

When there are no components connected between the output and the input of the op-amp, the gain that is measured is said to be the **open-loop gain**. The output of the op-amp cannot be greater than the supply voltage (from energy conservation) and so, if a *very* small voltage is applied to the input and the gain is infinite, then the output will be at the supply voltage. The output cannot be any greater, even for a larger input signal. The amplifier is said to be saturated.

- **infinite bandwidth** (all frequencies are amplified equally). If an alternating voltage is applied to the input, then the output will have the same frequency but a larger amplitude. The range of frequencies that are amplified by the same amount (the input signals of different frequencies that all have the same gain) is known as the bandwidth.
- **infinite slew rate**. When the input signal is changed, then the output signal will also change. The slew rate is a measure of the time delay between the changes to the input and output. A high slew rate implies a short time delay. With an infinite slew rate there is no delay.

In reality, op-amps are not ideal. The input impedance is in the range 10^6–$10^{12}\,\Omega$ and the output impedance is about $10^2\,\Omega$. The open-loop gain is not infinite but is the order of 10^5 for direct voltages. There is a finite bandwidth and the slew rate is about $10\,\text{V}\,\mu\text{s}^{-1}$.

21.2 Operational amplifier circuits

The operational amplifier as a comparator

When an op-amp is used (incorporated) in an electrical circuit, it is usually connected to a dual power supply so that the output voltage can be either positive or negative. Such a power supply can be represented as two sets of batteries, as shown in Figure 21.2.

www.iranino.co

IRANINOVIN

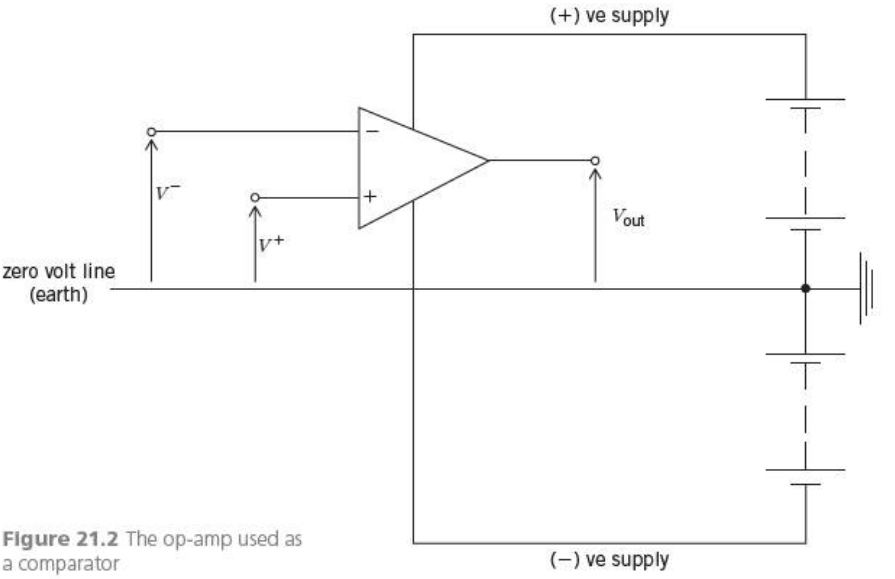


Figure 21.2 The op-amp used as a comparator

The link between the two batteries is referred to as the 'zero volt line' or 'earth' and it provides the reference from which the voltages on the inputs and the output are measured. The output V_{out} of this circuit is given by

$$V_{out} = A_0(V^+ - V^-)$$

where V^+ is the voltage at the non-inverting input, V^- is the voltage at the inverting input, and A_0 is the open-loop gain of the amplifier.

Consider the case where the non-inverting input V^+ is 0.95 V and the inverting input V^- is 0.94 V, and $A_0 = 10^5$. The supply voltages are at ±6 V. By substituting into the equation, the output voltage V_{out} is given by $V_{out} = 10^5 \times (0.95 - 0.94) = +1000$ V. From energy considerations, this answer is impossible since the output voltage can never exceed the power supply voltage. The amplifier is **saturated** and the output voltage will be 6 V. So,

if $V^+ > V^-$ *the output is* $+V_{supply}$
if $V^- > V^+$ *the output is* $-V_{supply}$

The circuit of Figure 21.2 is called a *comparator* because it compares the voltages applied to the non-inverting and the inverting inputs and then gives an output that depends on whether $V^+ > V^-$ or $V^- > V^+$.

When a circuit incorporating an op-amp is used as a comparator, it is usual to connect each of the two inputs to a potential divider, as shown in Figure 21.3.

The two resistors of equal resistance R provide an input voltage of 3 V at the inverting input. When the light-dependent resistor (LDR) is in darkness, its resistance is greater than 10 kΩ. The voltage at the non-inverting input will be greater than 3 V and the output will be at +6 V. In daylight, the resistance of the LDR will be less than 10 kΩ and the voltage at the non-inverting input will be less than 3 V. The output will be at −6 V.

The output, therefore, depends on the level of light illumination. This output can be made to operate an output device such as an LED (see page 306). By changing the values of the resistors in the potential divider connected to the inverting input, the voltage at which the circuit switches from +6 V to −6 V can be altered. Thus, the level of illumination at which the circuit switches can be changed. In practice, one of these two resistors would be a variable resistor.

The circuit shown in Figure 21.3 switches from +6 V to −6 V when the level of illumination increases. By changing round the connections to the two inputs, the output could be made to switch from +6 V to −6 V when the level of illumination decreases.

www.iranino.co

IRANINOVIN

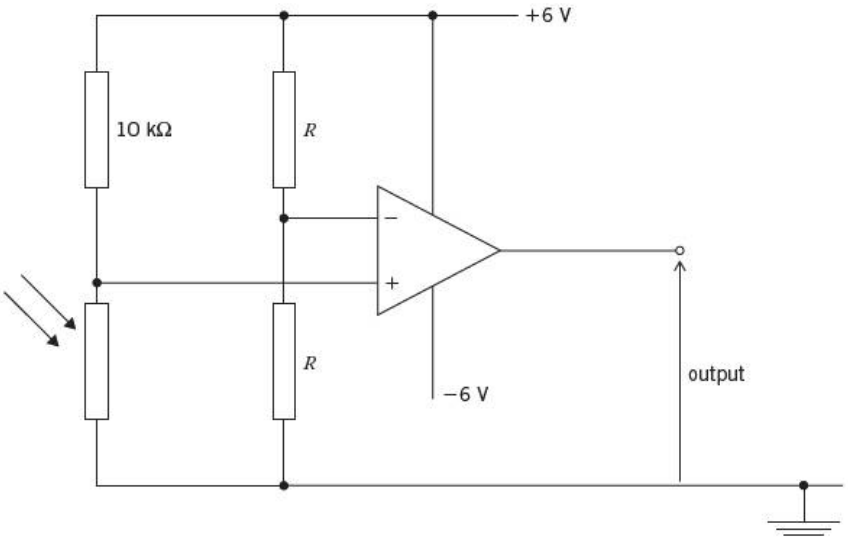


Figure 21.3 The op-amp used as a comparator to monitor illumination

Other devices could also be fitted into the comparator circuit. For example, a thermistor could be used so that the circuit provides a warning for either high or low temperatures.

Examples

1 State what is meant by a comparator circuit.
It is a circuit that usually incorporates an op-amp. The output voltage of the circuit is either positive or negative, depending on which of the two inputs to the circuit is at the higher potential.

2 Draw a diagram of a comparator that will give a positive output when the intensity of light incident on an LDR is high and a negative output when the intensity is low.

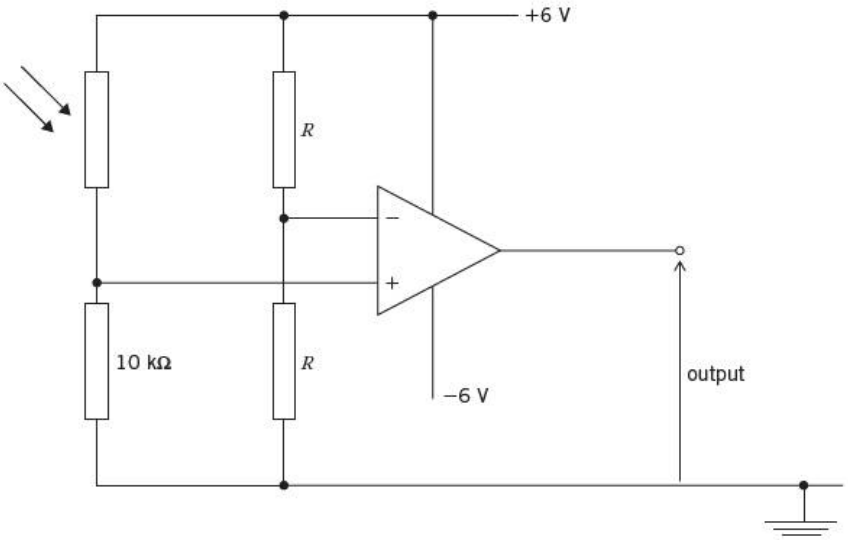


Figure 21.4

Now it's your turn

1 A thermistor has a resistance of 2500 Ω at 10 °C. Design a circuit that will switch from a positive output to a negative output when the thermistor temperature rises above 10 °C.

2 The resistance of the thermistor in question 1 is 1800 Ω at 20 °C. Show how your circuit in question 1 may be modified so that the comparator can be used either for temperatures rising above 10 °C *or* for temperatures rising above 20 °C.

www.iranino.co

IRANINOVIN

Operational amplifiers and feedback

Feedback is a process whereby a fraction of the output of any device is fed back to the input, so as to assist in the control of the device. Much of the movement of humans is controlled by feedback. If a person wishes to pick up an object, for example, then the person stretches out a hand while, at the same time, looking at the hand and the object. The visual signal from the eye is fed back to the brain to provide control for the hand. This feedback is a continuous process of refining the position of the hand relative to the object.

For an amplifier circuit, the basic arrangement is as shown in Figure 21.5.

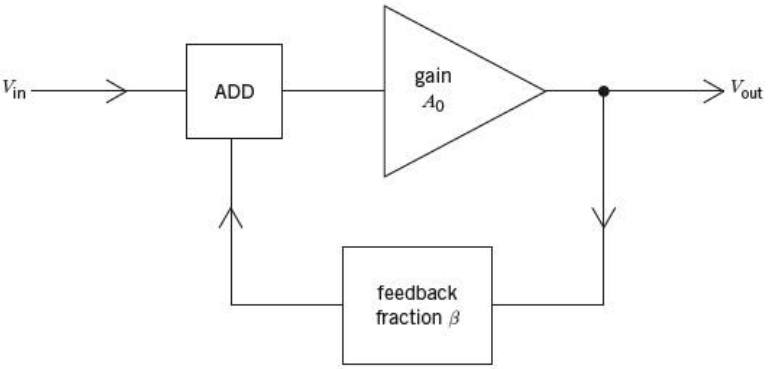


Figure 21.5 An amplifier circuit with feedback

A fraction β of the output signal V_{out} is fed back and added to the input signal V_{in}. The amplifier has open-loop gain A_0 and it amplifies whatever voltage there is at its input. So,

$$V_{out} = A_0 \times (\text{the amplifier input})$$

$$= A_0 \times (V_{in} + \beta V_{out})$$

and this gives

$$V_{out}(1 - A_0\beta) = A_0 \times V_{in}$$

The overall **voltage gain** (or simply **gain**) V_{out}/V_{in} of the amplifier is given by

$$\frac{V_{out}}{V_{in}} = \frac{A_0}{(1 - A_0\beta)}$$

If the fraction β is negative, then $(1 - A_0\beta)$ is greater than unity and the overall gain of the amplifier circuit is less than the open-loop gain of the operational amplifier itself. This process is known as **negative feedback**.

Negative feedback seems, at first sight, to defeat the object of an amplifier. However, it does have a number of benefits:

- increased bandwidth (the range of frequencies for which the gain is constant)
- less distortion of the output
- greater stability.

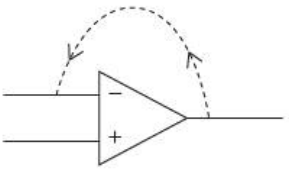


Figure 21.6 Negative feedback with an op-amp

Negative feedback can be achieved by feeding part of the output of the op-amp back to the *inverting* input, as illustrated in Figure 21.6.

The inverting amplifier

A circuit for an inverting amplifier incorporating an op-amp is shown in Figure 21.7. For simplicity, the power supplies are not shown.

www.iranino.co

IRANINOVIN

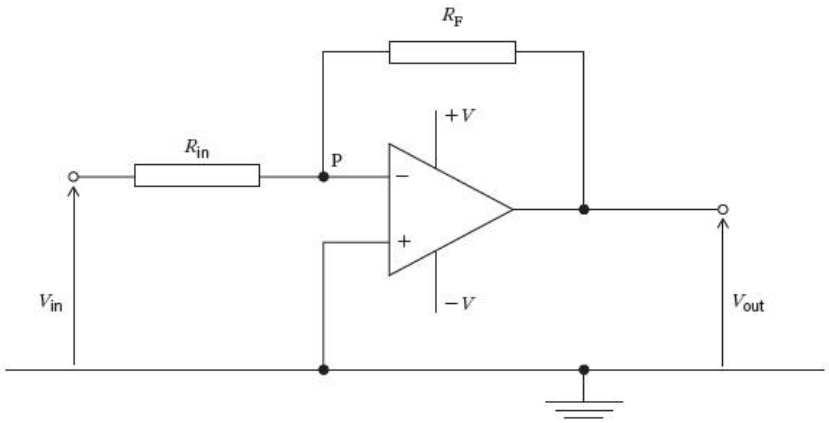


Figure 21.7 An inverting amplifier

Negative feedback is achieved by means of the resistor R_F. The resistors R_{in} and R_F act as a potential divider between the input and the output of the op-amp.

Since the open-loop gain of the op-amp is very large, then the input voltages at the non-inverting (+) and at the inverting (–) inputs must be almost the same. The non-inverting input is connected to the zero-volt line (the earth line) and, therefore, this input is zero. This means that the inverting input must be very nearly zero volts (or earth) and so point P in Figure 21.7 is known as a **virtual earth**.

The input impedance of the op-amp is very large, which means there is no current in either the inverting or the non-inverting inputs. So all the current from, or to, the input signal to the circuit must pass through the feedback resistor to the output. This is illustrated in Figure 21.8

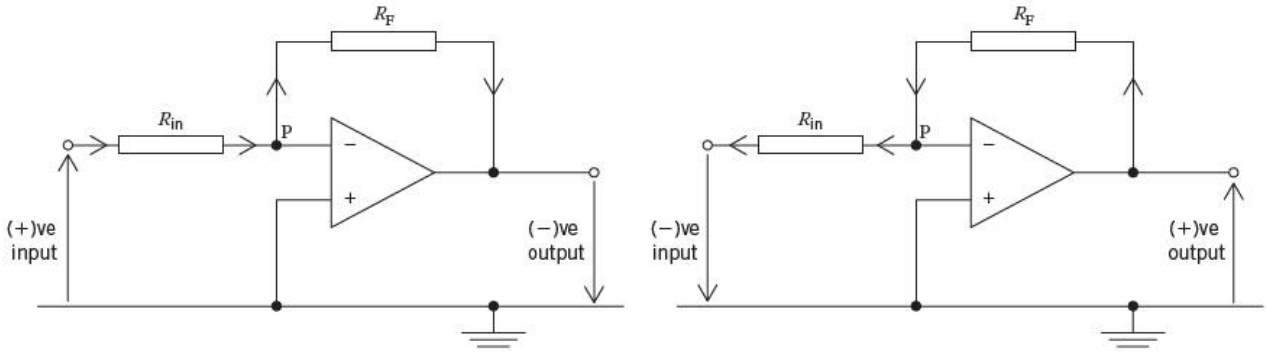


Figure 21.8 Feedback current

It can be seen that, because point P is a virtual earth, then a positive input signal gives rise to a negative output and vice versa. Consequently, the output is the inverse of the input and the amplifier is referred to as an **inverting amplifier**.

Referring to Figure 21.7, since the input impedance is infinite,

current in R_{in} = current in R_F.

So

$$\frac{\text{p.d. across } R_{in}}{R_{in}} = \frac{\text{p.d. across } R_F}{R_F}$$

The potential at P is zero and therefore

$$\frac{V_{in} - 0}{R_{in}} = \frac{0 - V_{out}}{R_F}$$

The overall voltage gain of the amplifier circuit is given by

$$\frac{V_{out}}{V_{in}} = -\frac{R_F}{R_{in}}$$

www.iranino.co

IRANINOVIN

The negative sign indicates that there is a phase difference of π rad between the input and the output. That is, the input is negative when the output is positive, and the input is positive when the output is negative.

The non-inverting amplifier

Figure 21.9 is a circuit for a non-inverting amplifier.

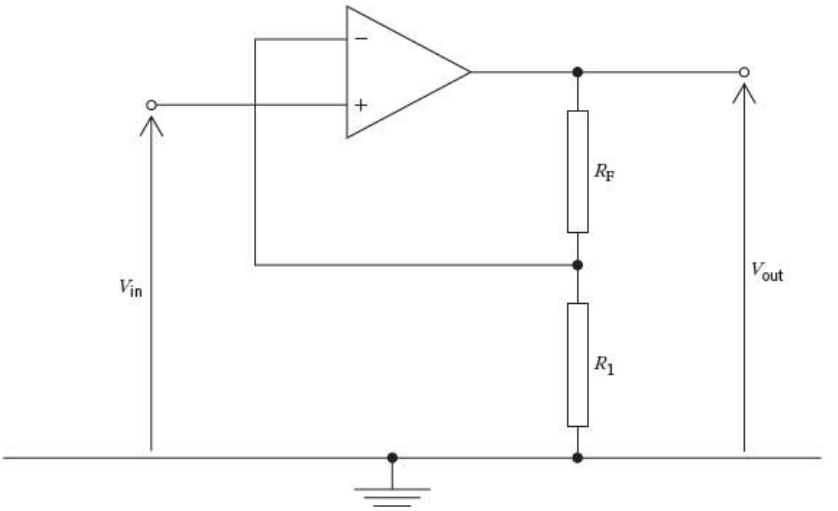


Figure 21.9 A non-inverting amplifier

The input voltage V_{in} is applied directly to the non-inverting input. Negative feedback is achieved by means of the resistors R_F and R_1. If the amplifier has not saturated, the overall voltage gain of the amplifier is given by

$$\frac{V_{out}}{V_{in}} = 1 + \frac{R_F}{R_1}$$

The non-inverting amplifier produces an output that is in phase with the input. That is, if the input is positive, the output is positive, and if the input is negative, the output is negative.

Examples

1 What is meant by an inverting amplifier?
An electrical circuit where the output potential difference (voltage) is π rad out of phase with the input voltage.

2 The magnitude of the gain of an inverting amplifier is 25. The supply voltage to the op-amp is ±9.0 V and the non-inverting input is at earth potential. Calculate the output voltage of the amplifier circuit for an input voltage at the inverting input of:
(a) +40 mV,
(b) −80 mV,
(c) −1.2 V.
$V_{out} = -V_{in} \times$ gain
(a) **−1.0 V**
(b) **+2.0 V**
(c) **+9.0 V (amplifier is saturated)**

www.iranino.co

IRANINOVIN

Now it's your turn

The circuit for an amplifier incorporating an op-amp is shown in Figure 21.10.

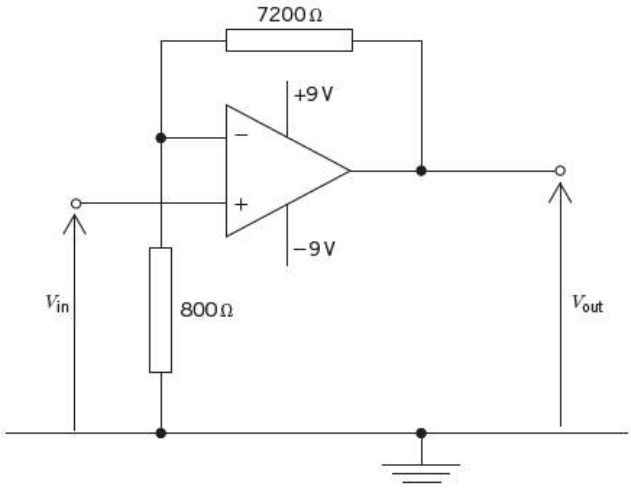


Figure 21.10

3 **(a)** State whether the amplifier is an inverting or a non-inverting amplifier.
(b) Calculate the gain of the amplifier circuit.
(c) Calculate the maximum input potential such that the amplifier does not saturate.

21.3 Output devices

The circuits incorporating an operational amplifier produce an output voltage. If this voltage is connected across some form of resistor, there will be a current from the output of the op-amp to the resistor.

This output current must not be larger than about 25 mA. With larger currents the op-amp would be destroyed. In fact, op-amps frequently have an output resistor as part of the integrated circuit so that, even if the output is 'shorted', the op-amp will not be damaged.

The sensing circuit may be required to switch on or off an appliance that requires a large current; for example, an electric motor. The switching on or off of a large current by means of a small current can be achieved using a relay.

The relay

A **relay** is an electromagnetic switch. When a small current passes through the coil of an electromagnet, the electromagnet operates a switch. This switch is used to switch on, or off, a much larger current. The symbol for a relay is shown in Figure 21.11.

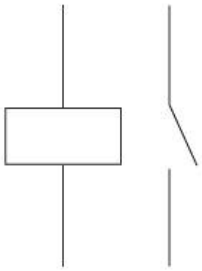


Figure 21.11
Symbol for a relay

Figure 21.12 shows one way in which a relay may be connected to the output of an op-amp.

The output of the op-amp may be positive or negative. However, diode D_1 conducts only when the output is positive with respect to earth. Therefore, the relay coil is energised only when the output of the op-amp is positive.

When a current is switched off in an electromagnet, an e.m.f. is induced in the coil of the electromagnet. This e.m.f. could be large enough to damage the op-amp. A diode D_2 is connected across the coil to protect the op-amp from this e.m.f. Note that the diode is connected so that, when the output of the op-amp is positive, the diode D_2 is reverse-biased so that the current is in the coil, not the diode.

Relays are designed so that they can be used to switch on or switch off currents when the coil of the relay is energised.

www.iranino.co

IRANINOVIN

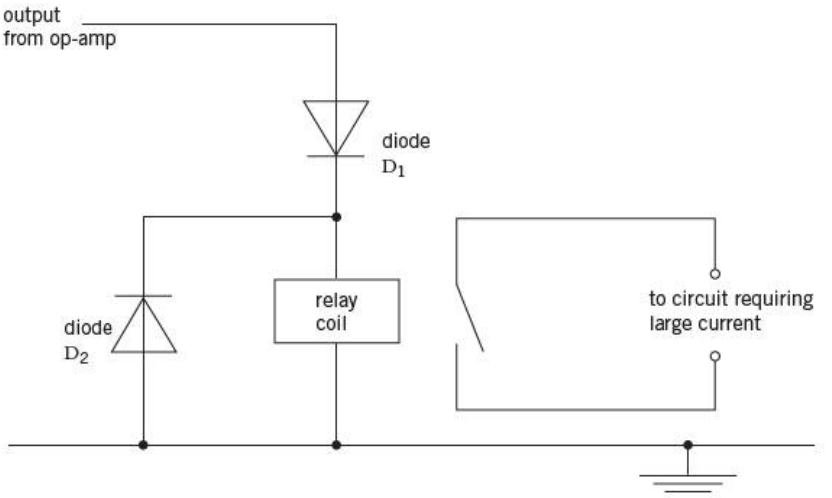


Figure 21.12 The connection of a relay to an op-amp

The light-emitting diode (LED)

There are many uses of sensing devices where the purpose is to indicate whether something is too hot or too cold, switched on or off, etc. In these situations, all that is required for the output of the sensor is some form of visible indicator; that is, a lamp. Filament lamps not only dissipate a comparatively large amount of power but also, they are not very robust and the filament is likely to break after being switched on and off a number of times. The **light-emitting diode (LED)** is a semiconductor device that is robust, reliable and dissipates much less power than a filament lamp. LEDs are available in different colours including red, green, yellow and amber. The symbol for an LED is shown in Figure 21.13.

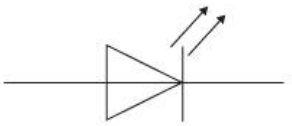


Figure 21.13 Symbol for a light-emitting diode (LED)

The LED emits light only when it is forward-biased. A resistor is frequently connected in series with an LED so that, when the LED is forward-biased, the current in the diode is not so large that it would damage it. A typical maximum current for an LED is 20 mA. Also, the LED may be damaged if the reverse-bias voltage exceeds about 5 V.

Figure 21.14 shows two LEDs connected to the output of an op-amp to indicate whether the output is positive or negative.

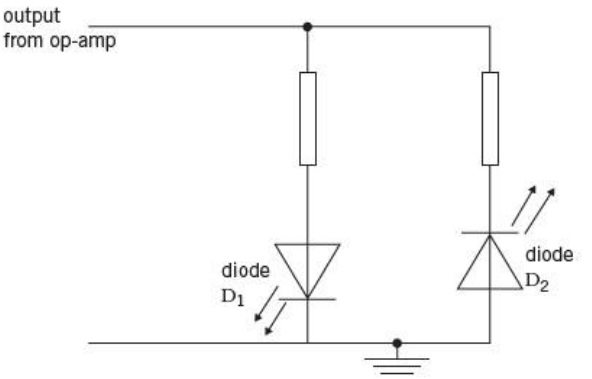


Figure 21.14 LEDs connected to indicate the state of an op-amp output

When the output is positive with respect to earth, diode D_1 will be forward-biased and will conduct, emitting light. Diode D_2 will be reverse-biased and will not emit light. When the polarity of the output changes, diode D_2 will be forward-biased and will conduct, emitting light. D_1 will be reverse-biased and will not emit light. The LEDs can be chosen so that they emit light of different colours.

www.iranino.co

IRANINOVIN

Digital and analogue meters as output devices

LEDs are used to indicate the polarity of the output and may be used when the op-amp saturates (for example, using the op-amp as a comparator to detect whether a temperature is above or below a specific value). When the output of the op-amp does not saturate, the output voltage can be used to indicate the magnitude of whatever is being sensed (for example, the level of fuel in a tank). A digital or analogue voltmeter, connected between the op-amp output and earth, will indicate the output voltage.

The output voltage will be proportional to the input to the processing unit. It will also be dependent on the magnitude of whatever is being sensed. However, the voltmeter reads a potential difference, not the quantity that is being measured and therefore the voltmeter reading is unlikely to vary linearly with the change in this quantity. The voltmeter must be calibrated. The calibration of meters and the use of calibration curves is discussed in Topic 2.

Summary

- A comparator incorporating an operational amplifier (op-amp) enables a particular voltage level to be monitored, indicating whether the voltage is above or below a specified value.
- Low-value potential differences may be amplified using an amplifier incorporating an op-amp.
- Negative feedback reduces the gain of an amplifier but provides stability and increased bandwidth.
- The gain of an inverting amplifier is $-R_F/R_{in}$
- The gain of a non-inverting amplifier is $(1 + R_F/R_1)$
- A relay is used to control a large-value current/voltage by means of the low current/voltage output of a processing unit.
- A light-emitting diode (LED) is used to indicate the state (high or low, on or off) of the output of a processing unit.
- Digital and analogue meters may be used to measure the output of a processing unit. Meters need to be calibrated in order to measure the quantity that is being monitored.

Examination style questions

In all questions, assume that any operational amplifier (op-amp) is ideal and that the power supply to the op-amp is a dual ±9.0 V supply.

1 **a** State four properties of an ideal operational amplifier (op-amp).

b Fig. 21.15 shows an amplifier incorporating an op-amp.

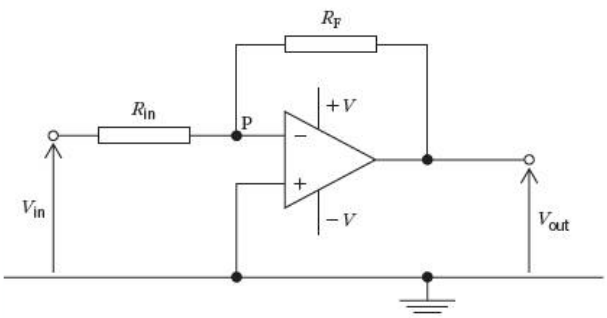


Fig. 21.15

i The point P is referred to as a *virtual earth*. State what is meant by this.

ii Derive an expression, in terms of the resistances R_F and R_{in}, for the gain of the amplifier.

c The resistances R_F and R_{in} are 12 kΩ and 630 Ω respectively. Calculate:

i the gain of the amplifier,

ii the maximum input voltage such that the amplifier does not saturate.

2 A thermistor T is included in the circuit shown in Fig. 21.16. At 6 °C, the thermistor has a resistance of 3.00 kΩ.

a Calculate the resistance of resistor R such that the output of the op-amp will change polarity at 6 °C.

b Suggest and explain the direction of change of the polarity as the temperature is gradually increased from a low value.

www.iranino.co

IRANINOVIN

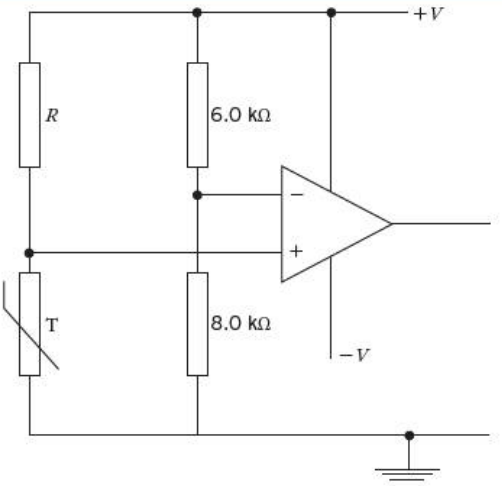


Fig. 21.16

c Two light-emitting diodes, one red and the other green, are to be used to indicate the temperature of the thermistor. The red LED is to emit light when the thermistor is below the temperature stated in **a** and the green LED is to emit light when the temperature is above that stated in **a**. Sketch the circuit including the arrangement of the LEDs.

3 **a** Negative feedback may be used in amplifier circuits. State

i what is meant by *negative feedback*, *[2]*

ii two effects of negative feedback on an amplifier incorporating an operational amplifier (op-amp). *[2]*

b Fig. 21.17 is a circuit for an amplifier that is used with a microphone.

The output potential difference V_{OUT} is 4.4 V when the potential at point P is 62 mV.

Determine

i the gain of the amplifier, *[1]*

ii the resistance of the resistor R. *[2]*

c The maximum potential produced by the microphone at point P on Fig. 21.17 is 95 mV. The power supply for the operational amplifier may be either +/–5 V or +/–9 V.

State which power supply should be used. Justify your answer quantitatively. *[3]*

Cambridge International AS and A Level Physics, 9702/42 May/June 2010 Q 9

4 **a** Describe the structure of a metal wire strain gauge. You may draw a diagram if you wish. *[3]*

b A strain gauge S is connected into the circuit of Fig. 21.18.

The operational amplifier (op-amp) is ideal.

The output potential V_{OUT} of the circuit is given by the expression

$$V_{OUT} = \frac{R_F}{R} \times (V_2 - V_1).$$

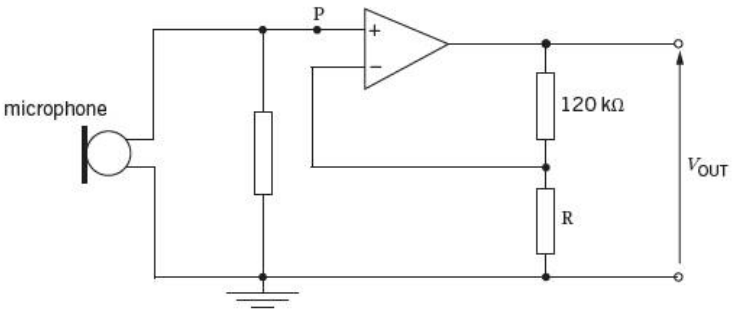


Fig. 21.17

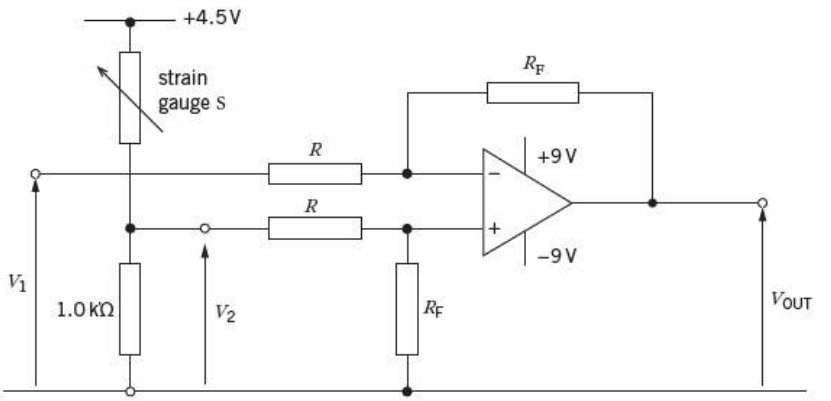


Fig. 21.18

www.iraninovin.com

IRANINOVIN

i State the name given to the ration R_f/R. [1]

ii The strain gauge S has resistance 125 Ω when not under strain.
Calculate the magnitude of V_1 such that, when the strain gauge S is not strained, the output V_{OUT} is zero. [3]

iii In a particular test, the resistance of S increases to 128 Ω. V_1 is unchanged.
The ratio R_f/R is 12.
Calculate the magnitude of V_{OUT}. [2]

Cambridge International AS and A Level Physics, 9702/42 May/June 2011 Q 9

5 a State two effects of negative feedback on the gain of an amplifier incorporating an operational amplifier (op-amp). [2]

b An incomplete circuit diagram of a non-inverting amplifier using an ideal op-amp is shown in Fig. 21.19.

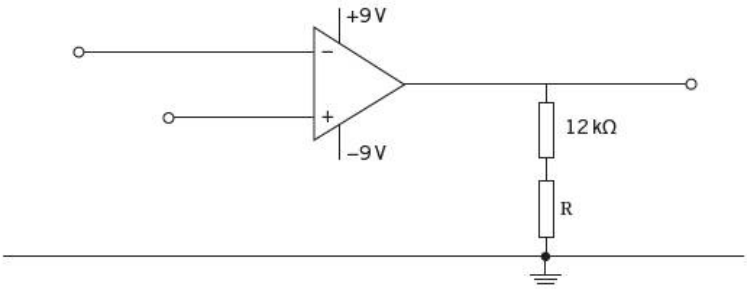


Fig. 21.19

i Complete a copy of the circuit diagram of Fig. 21.19. Label the input and the output. [2]

ii Calculate the resistance of resistor R so that the non-inverting amplifier has a voltage gain of 15. [2]

c On a copy of Fig. 21.20, draw a graph to show the variation with input potential V_{IN} of the output potential V_{OUT}.

You should consider input potentials in the range 0 to +1.0 V. [2]

d The output of the amplifier circuit of Fig. 21.19 may be connected to a relay.

State and explain one purpose of a relay. [2]

Cambridge International AS and A Level Physics, 9702/42 Oct/Nov 2011 Q 9

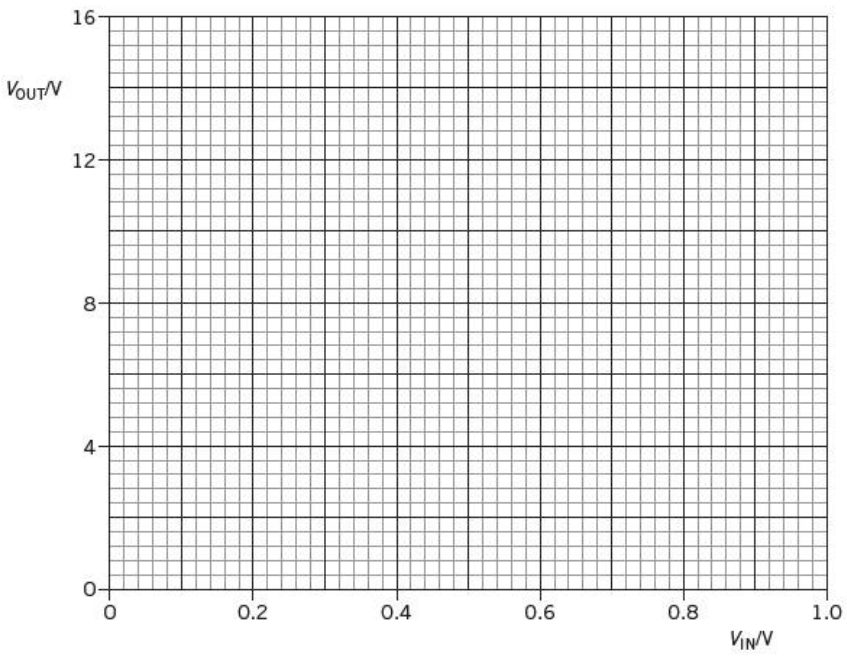


Fig. 21.20

www.iranino.co

IRANINOVIN

6 A student designs an electronic sensor that is to be used to switch on a lamp when the light intensity is low. Part of the circuit is shown in Fig. 21.21.

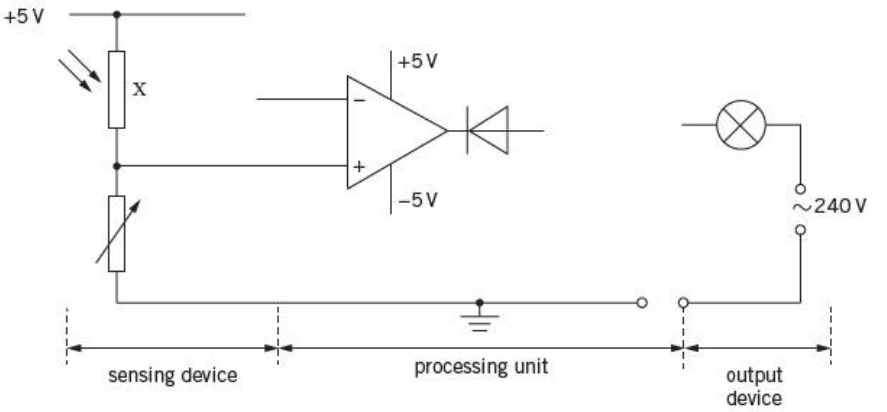


Fig. 21.21

a State the name of the component labelled X on Fig. 21.21. *[1]*

b On a copy of Fig. 21.21, draw the symbols for

i two resistors to complete the circuit for the sensing device, *[2]*

ii a relay to complete the circuit for the processing unit. *[2]*

c **i** State the purpose of the relay. *[1]*

ii Suggest why the diode is connected to the output of the operational amplifier (op-amp) in the direction shown. *[2]*

Cambridge International AS and A Level Physics, 9702/41 May/June 2012 Q 10

www.iranino.co

IRANINOVIN

22 Magnetic fields

By the end of this topic, you will be able to:

22.1 **(a)** understand that a magnetic field is an example of a field of force produced either by current-carrying conductors or by permanent magnets
(b) represent a magnetic field by means of field lines;

22.2 **(a)** appreciate that a force might act on a current-carrying conductor placed in a magnetic field
(b) recall and solve problems using the equation $F = BIL \sin\theta$, with directions as interpreted by Fleming's left-hand rule
(c) define magnetic flux density and the tesla
(d) understand how the force on a current-carrying conductor can be used to measure the flux density of a magnetic field using a current balance

22.3 **(a)** predict the direction of the force on a charge moving in a magnetic field
(b) recall and solve problems using $F = Bqv \sin\theta$
(c) derive the expression $V_H = BI/ntq$ for the Hall voltage
(d) describe and analyse qualitatively the deflection of beams of charged particles by uniform electric and uniform magnetic fields
(e) explain how electric and magnetic fields can be used in velocity selection
(f) explain the main principles of one method for the determination of v and e/m_e for electrons

22.4 **(a)** sketch flux patterns due to a long straight wire, a flat circular coil and a long solenoid
(b) understand that the field due to a solenoid is influenced by the presence of a ferrous core
(c) explain the forces between current-carrying conductors and predict the direction of the forces
(d) describe and compare the forces on mass, charge and current in gravitational, electric and magnetic fields, as appropriate

22.5 **(a)** explain the main principles behind the use of nuclear magnetic resonance imaging (NMRI) to obtain diagnostic information about internal structures
(b) understand the function of the non-uniform magnetic field, superimposed on the large constant field, in diagnosis using NMRI

Starting points

- Identify (ferro-)magnetic materials.
- The law of magnets – like poles repel, unlike poles attract.

22.1 Concept of a magnetic field

Some of the properties of magnets have been known for many centuries. The ancient Greeks discovered an iron ore called lodestone which, when hung from a thread, would come to rest always pointing in the same direction. This was the basis of the magnetic compass which has been in use since about 1500BC as a means of navigation.

The magnetic compass is dependent on the fact that a freely suspended magnet will come to rest pointing north–south. The ends of the magnet are said to be poles. The pole pointing to the north is referred to as the north-seeking pole (the north pole or N-pole) and the other, the south-seeking pole (the south pole or S-pole). It is now known that a compass behaves in this way because the Earth is itself a magnet.

www.iranino.co

IRANINOVIN

Magnets exert forces on each other. These forces of either attraction or repulsion are used in many children's toys, in door catches and 'fridge magnets'. The effects of the forces may be summarised in the law of magnets.

- *Like poles repel.*
- *Unlike poles attract.*

The law of magnets implies that around any magnet, there is a region where a magnetic pole will experience a force. This region is known as a **magnetic field**.

A magnetic field is a region of space where a magnetic pole experiences a force.

Magnetic fields are not visible but they may be represented by lines of magnetic force or **magnetic field lines**, or magnetic 'flux'. A simple way of imagining magnetic field lines is to think of one such line as the direction in which a free magnetic north pole would move if placed in the field. Magnetic field lines may be plotted using a small compass (a plotting compass) or by the use of iron filings and a compass (Figure 22.1).

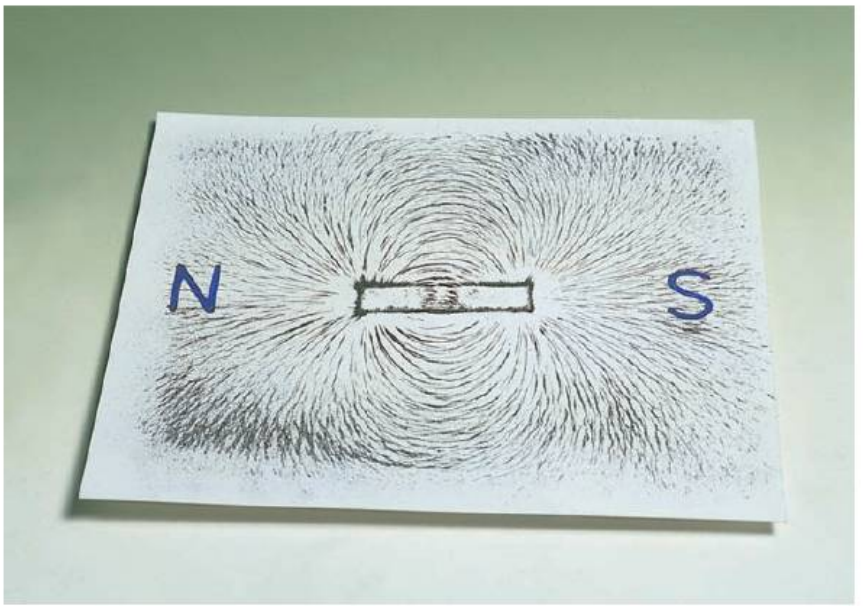


Figure 22.1 The iron filings line up with the magnetic field of the bar magnet which is under the sheet of paper. A plotting compass will give the direction of the field.

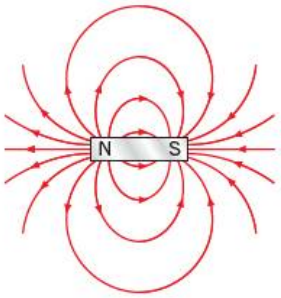


Figure 22.2 Magnetic field pattern of a bar magnet

The magnetic field lines of a bar magnet are shown in Figure 22.2. Effects due to the Earth's magnetic field have not been included since the Earth's field is relatively weak and would be of importance only at some distance from the magnet. It is important to realise that, although the magnetic field has been drawn in two dimensions, the actual magnetic field is three-dimensional.

For any magnetic field:
- the magnetic field lines start at a north pole and end at a south pole
- the magnetic field lines are smooth curves which never touch or cross
- the strength of the magnetic field is indicated by the distance between the lines – closer lines mean a stronger field.

It can be seen that these properties are very similar to those for electric field lines (Topic 17).

Figure 22.3 illustrates the magnetic field pattern between the north pole of one magnet and the south pole of another. This pattern is similar to that produced between the poles of a horseshoe magnet.

www.iranino.co

IRANINOVIN

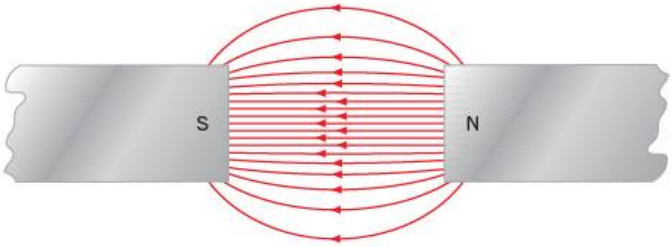


Figure 22.3 Magnetic field pattern between a north and south pole

Figure 22.4 shows the magnetic field between the north poles of two magnets. The magnetic field due to one magnet opposes that due to the other. The field lines cannot cross and consequently there is a point X, known as a **neutral point**, where there is no resultant magnetic field because the two fields are equal in magnitude but opposite in direction.

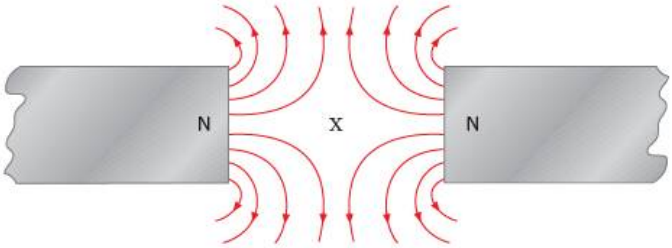


Figure 22.4 Magnetic field pattern between two north poles

Example

A circular magnet is made with its north pole at the centre, separated from the surrounding circular south pole by an air gap. Draw the magnetic field lines in the gap (Figure 22.5).

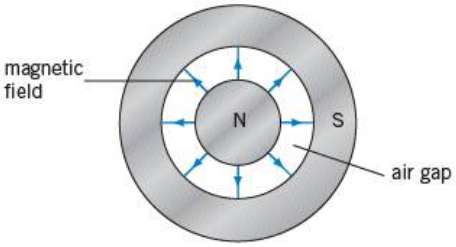


Figure 22.5

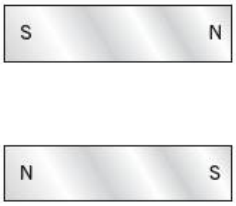


Figure 22.6

Now it's your turn

1 Draw a diagram to illustrate the magnetic field between the south poles of two magnets.
2 Two bar magnets are placed on a horizontal surface (Figure 22.6)
Draw the two magnets, and sketch the magnetic field lines around them. On your diagram, mark the position of any neutral points (points where there is no resultant magnetic field).



Figure 22.7 Hans Christian Oersted

www.iranino.co

Magnetic effect of an electric current

The earliest discovery of the magnetic effect of an electric current was made in 1820 by Oersted, a Danish physicist.

He noticed that a compass was deflected when brought near to a wire carrying an electric current. It is now known that all electric currents produce magnetic fields. The size and shape of the magnetic field depends on the size of the current and the shape (configuration) of the conductor through which the current is travelling. The shape of the fields will be considered in more detail on pages 324–325.

IRANINOVIN

22.2 Force on a current-carrying conductor

We have seen that the plotting of lines of magnetic force gives the direction and shape of the magnetic field. Also, the distance between the lines indicates the strength of the field. However, the strength of the magnetic field was not defined. Physics is the science of measurement and, consequently, the topic would not be complete without defining and measuring magnetic field strength. Magnetic field strength is defined through a study of the force on a current-carrying conductor – the motor effect.

The motor effect

The interaction of the magnetic fields produced by two magnets causes forces of attraction or repulsion between the two. A current-carrying conductor produces a magnetic field around the conductor (see page 324). If a conductor is placed between magnet poles and a current is passed through the conductor, the magnetic fields of the current-carrying conductor and the magnet will interact, causing forces between them. This is illustrated in Figure 22.8.

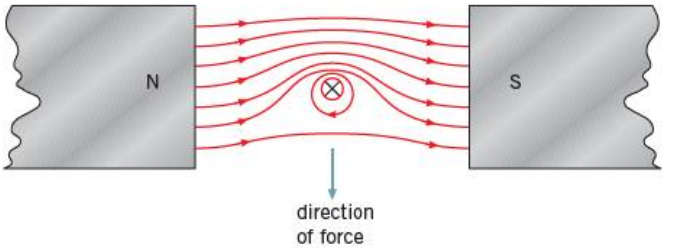


Figure 22.8 The interacting magnetic fields of a current-carrying conductor and the poles of a magnet

The existence of the force may be demonstrated with the apparatus shown in Figure 22.9.

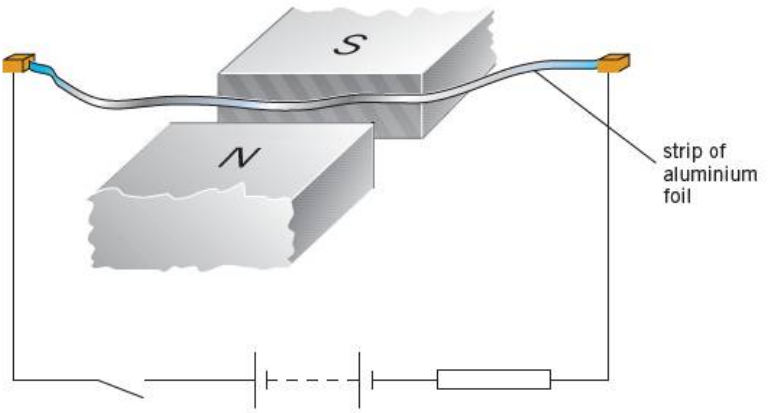


Figure 22.9 Demonstrating the motor effect acting on a piece of aluminium

The strip of aluminium foil is held loosely between the poles of a horseshoe magnet so that the foil is at right angles to the magnetic field. When the current is switched on, the foil jumps and becomes taut, showing that a force is acting on it. The direction of the force, known as the **electromagnetic force**, depends on the directions of the magnetic field and of the current. This phenomenon, when a current-carrying conductor is at an angle to a magnetic field, is called the **motor effect**.

The direction of the force relative to the directions of the current and the magnetic field may be predicted using **Fleming's left-hand rule**. This is illustrated in Figure 22.10.

www.iranino.co

IRANINOVIN

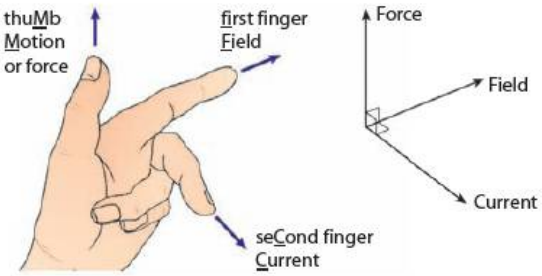


Figure 22.10 Fleming's left-hand rule

*If the first two fingers and thumb of the left hand are held at right angles to one another with the **F**irst finger in the direction of the **F**ield and the se**C**ond finger in the direction of the **C**urrent, then the thu**M**b gives the direction of the force or **M**otion.*

Note that, if the conductor is held fixed, motion will not be seen but, nevertheless, there will be an electromagnetic force.

The magnitude of the electromagnetic force may be investigated using the 'current balance' apparatus shown in Figure 22.11.

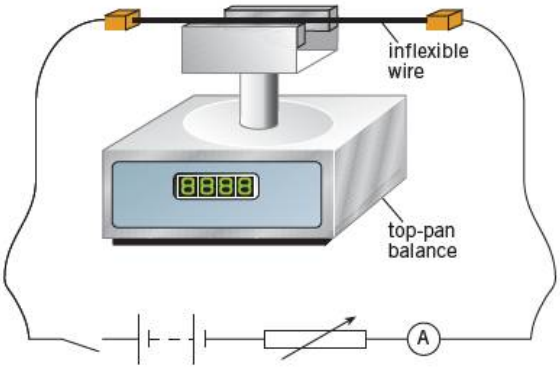


Figure 22.11 Apparatus to investigate the magnitude of the electromagnetic force

A horseshoe magnet is made by placing two flat magnadur magnets on a U-shaped iron core. A length of inflexible wire is firmly fixed between the poles of the magnet. When the current is switched on, the reading on the top-pan balance changes. This change in reading is a measure of the electromagnetic force. Variation of the current leads to the conclusion that the electromagnetic force F is proportional to the current I. By using magnadur magnets of different lengths (but of similar strengths), the force F is found to be proportional to the length L of wire in the magnetic field. Finally, by varying the angle θ between the wire and the direction of the magnetic field, the force F is found to be proportional to sin θ. Hence the expression

$$F \propto IL \sin \theta$$

is derived. The expression can be re-written as

$$F = BIL \sin \theta$$

where B is a constant.

The constant B depends on the strength of the magnet and, if stronger magnets are used, the constant has a greater value. The equation can, therefore, be used as the defining equation for magnetic field strength.

The equation can be rewritten as

$$\frac{F}{L} = BI \sin \theta$$

www.iranino.co

IRANINOVIN

So,

For a long straight conductor carrying unit current at right angles to a uniform magnetic field, the magnetic flux density B is numerically equal to the force per unit length of the conductor.

The **magnetic flux density** B is measured in **tesla (T)**. An alternative name for this unit is **weber per square metre (Wb m^{-2})**.

One tesla is the uniform magnetic flux density which, acting normally to a long straight wire carrying a current of 1 ampere, causes a force per unit length of 1 N m^{-1} on the conductor.

Since force is measured in newtons, length in metres and current in amperes, it can be derived from the defining equation for the tesla that the tesla may also be expressed as N m^{-1} A^{-1}. The unit involves force which is a vector quantity and thus magnetic flux density is also a vector.

When using the equation $F = BIL \sin \theta$, it is sometimes helpful to think of the term $B \sin \theta$ as being the component of the magnetic flux density which is at right angles (or normal) to the conductor (see Figure 22.12).

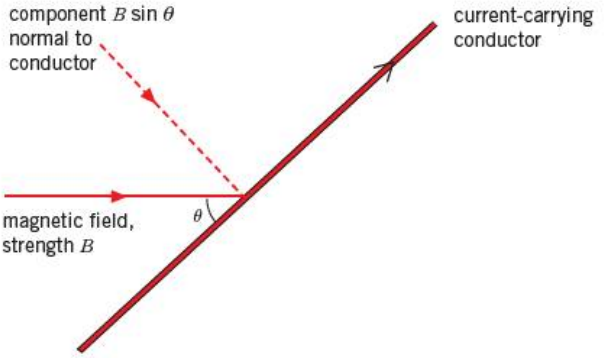


Figure 22.12 $B \sin \theta$ is the component of the magnetic field which is normal to the conductor

The tesla is a large unit for the measurement of flux density. A strong magnet may have a flux density between its poles of a few teslas. The magnetic flux density due to the Earth in the UK is about 44 μT at an angle of 66° to the horizontal.

Example

The horizontal component of the Earth's magnetic flux density is 1.8×10^{-5} T. The current in a horizontal cable is 150 A. Calculate, for this cable:

(a) the maximum force per unit length,

(b) the minimum force per unit length.

In each case, state the angle between the cable and the magnetic field.

(a) force per unit length = $F/L = BI \sin \theta$

Force per unit length is a maximum when $\theta = 90°$ and $\sin \theta = 1$.

Force per unit length = $1.8 \times 10^{-5} \times 150 = 2.7 \times 10^{-3}$ N m^{-1}

Maximum force per unit length = **2.7×10^{-3} N m^{-1}** when the cable is at right angles to the field.

(b) Force per unit length is a minimum when $\theta = 0$ and $\sin \theta = 0$.

Minimum force per unit length = **0** when the cable is along the direction of the field.

Now it's your turn

3 The effective length of the filament in a light bulb is 3.1 cm and, for normal brightness, the current in the filament is 0.25 A. Calculate the maximum electromagnetic force on the filament when in the Earth's field of flux density 44 μT.

www.iranino.co

IRANINOVIN

4 A straight conductor carrying a current of 6.5 A is situated in a uniform magnetic field of flux density 4.3 mT. Calculate the electromagnetic force per unit length of the conductor when the angle between the conductor and the field is:
(a) 90°,
(b) 45°.

Force between parallel conductors

A current-carrying conductor has a magnetic field around it. If a second current-carrying conductor is placed parallel to the first, this second conductor will be in the magnetic field of the first and, by the motor effect, will experience a force.

By similar reasoning, the first conductor will also experience a force. By Newton's third law (see Topic 4), these two forces will be equal in magnitude and opposite in direction. The effect can be demonstrated using the apparatus in Figure 22.13. It can be seen that, if the currents are in the same direction, the pieces of foil move towards one another (the pinch effect). If the currents are in opposite directions, the pieces of foil move apart. An explanation for the effect can be found by reference to Figure 22.14.

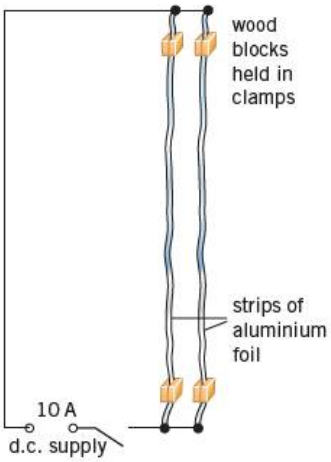


Figure 22.13 Apparatus to demonstrate the force between parallel current-carrying conductors

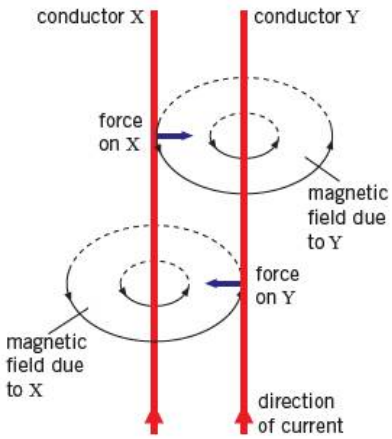


Figure 22.14 Diagram to illustrate the force between parallel current-carrying conductors

The current in conductor X causes a magnetic field and the field lines are concentric circles (see page 324). These field lines will be at right angles to conductor Y and so, using Fleming's left-hand rule, there will be a force on Y in the direction of X. Using similar reasoning, the force on X due to the magnetic field of Y is towards Y. Reversing the direction of the current in one conductor will reverse the directions of the two forces and, thus, when the currents are in opposite directions, the conductors tend to move apart.

The force per unit length on each of the conductors depends on the magnitude of the current in each conductor and also their separation. Since force can be expressed in terms of base units, then it is possible to use the effect to define the ampere in terms of SI units. The definition may be given as follows.

Consider two long straight parallel conductors of negligible area of cross-section, situated one metre apart in a vacuum. Then, if the current in each conductor is 1 ampere, the force per metre length acting on each conductor is 2.0×10^{-7} N.

It may not be necessary to learn the definition of the ampere, but it is important to realise how the ampere is defined in terms of SI units and that the definition does not involve the property of a substance.

Measurement of flux density

The magnetic flux density of a horseshoe magnet may be determined using the apparatus shown in Figure 22.11.

www.iranino.co

IRANINOVIN

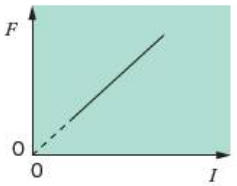


Figure 22.15 Graph of force against current to determine magnetic flux density

The wire is held normal to the magnetic field. The values of the force F for corresponding values of current I are determined and a graph of force F (y-axis) against current I is plotted (Figure 22.15).

Since $F/I = BL \sin\theta$ and $\theta = 90°$, then the gradient of the graph (F/I) is equal to BL. Measuring L, then the magnetic flux density B can be found.

22.3 Force on a moving charged particle in a magnetic field

An electric current is charge in motion. Since charge is always associated with particles, then the current in a conductor is a movement of charged particles. If a current-carrying conductor is placed in a magnetic field, it may experience a force depending on the angle between the field and the conductor. The force arises from the force on the individual moving charged particles in the conductor.

It has been shown that a conductor of length L carrying a current I at an angle θ to a uniform magnetic field of flux density B experiences a force F given by

$$F = BIL \sin\theta$$

If there are n charged particles in a length L of the conductor, each carrying a charge q, which pass a point in the conductor in time t, then the current in the conductor is given by

$$I = \frac{nq}{t}$$

and the speed v of charged particles is given by $v = \frac{L}{t}$. Hence,

$$F = B\frac{nq}{t}L \sin\theta$$

and

$$F = Bnqv \sin\theta$$

This force is the force on n charged particles. Therefore,

The force on a particle of charge q moving at speed v at an angle θ to a uniform magnetic field of flux density B is given by
$F = Bqv \sin\theta$

The direction of the force will be given by Fleming's left-hand rule. However, it must be remembered that the second finger is used to indicate the direction of the conventional current. If the particles are positively charged, then the second finger is placed in the same direction as the velocity. However, if the particles are negatively charged (e.g. electrons), the finger must point in the opposite direction to the velocity.

Consider a positively charged particle of mass m carrying charge q and moving with velocity v as shown in Figure 22.16.



Figure 22.16 Path of a charged particle in a magnetic field

The particle enters a uniform magnetic field of flux density B which is normal to the direction of motion of the particle. As the particle enters the field, it will experience a force normal to its direction. This force will not change the speed of the particle but it

www.iranino.co

will change its direction of motion. As the particle moves through the field, the force will remain constant, since the speed has not changed, and it will always be normal to the direction of motion. The particle will, therefore, move in an arc of a circle (see Topic 7).

The force $F = Bqv \sin \theta$ (in this case, $\sin \theta = 1$), provides the centripetal force for the circular motion. Hence,

centripetal force = $mv^2/r = Bqv$

Re-arranging,

$$\text{radius of path} = r = \frac{mv}{Bq}$$

The importance of this equation is that, if the speed of the particle and the radius of its path are known, then the specific charge, i.e. the ratio of charge to mass, can be found. Then, if the charge on the particle is known, its mass may be calculated. The technique is also used in nuclear research to try to identify some of the fundamental particles. The tracks of charged particles are made visible in a bubble chamber (Figure 22.17). Analysing these tracks gives information as to the sign of the charge on the particle and its specific charge.

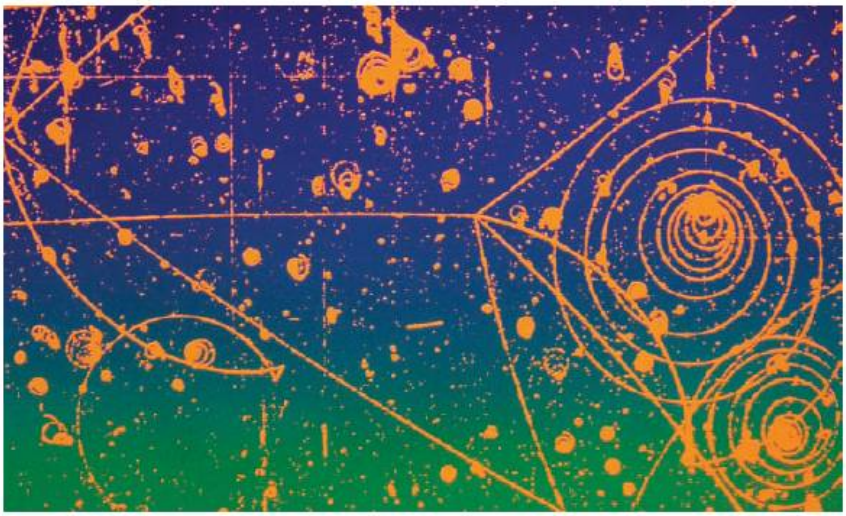


Figure 22.17 Tracks of particles produced in a bubble chamber

Specific charge of the electron: the fine-beam tube

Specific charge is the name given to the ratio of the charge q on a particle and its mass m.

specific charge = q/m

Specific charge can give us information about a particle and, if the charge on the particle is known, then the mass of the particle can be determined.

The determination of the charge on the electron has already been outlined in AS Topic 17. Determination of the specific charge on the electron enables us to obtain a value for its mass.

We have seen that a particle of mass m and charge q moving with speed v at right angles to a uniform magnetic field of flux density B experiences a force F given by

$F = Bqv$

The direction of this force is given by Fleming's left-hand rule and is always normal to the velocity, giving rise to circular motion

$Bqv = mv^2/r$

Re-arranging the terms,

$q/m = v/Br$

www.iranino.co

IRANINOVIN

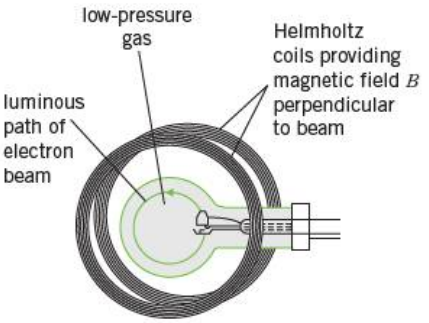


Figure 22.18 Fine-beam tube

The ratio charge/mass (e/m_e) for an electron – its specific charge – may be determined using a fine-beam tube, as shown in Figure 22.18.

The path of electrons is made visible by having low-pressure gas in the tube and, thus, the radius of the orbit may be measured. By accelerating the electrons through a known potential difference V, their speed v on entry into the region of the magnetic field may be calculated (see A Level Topic 17).

$$eV = \tfrac{1}{2}m_e v^2$$

The magnetic field is provided by a pair of current-carrying coils (Helmholtz coils, page 325).

Combining the equations $e/m_e = v/Br$ and $eV = \tfrac{1}{2}m_e v^2$ then

specific charge on electron $e/m_e = 2V/B^2r^2$.

Values for the charge e and mass m_e are usually given as

charge $e = 1.60 \times 10^{-19}$ C
mass $m_e = 9.11 \times 10^{-31}$ kg

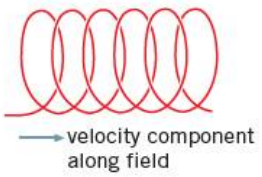


Figure 22.19 Electrons moving in a helix

It is of interest to rotate the fine-beam tube slightly, so that the velocity of the electrons is not normal to the magnetic field. In this case, the path of the electrons is seen to be a helix (rather like the coils of a spring). The component of the velocity normal to the field gives rise to circular motion. However, there is also a component of velocity along the direction of the field. There is no force on the electrons resulting from this component of velocity. Consequently, the electrons execute circular motion *and* move in a direction normal to the plane of the circle. The circle is 'pulled out' into a helix (Figure 22.19). The helical path is an important aspect of the focusing of electron beams by magnetic fields in an electron microscope (Figure 22.20).



Figure 22.20 A laboratory technician using an electron microscope

www.iranino.com

IRANINOVIN

Example

Electrons are accelerated to a speed of 8.4×10^6 m s^{-1}. They then pass into a region of uniform magnetic flux of flux density 0.50 mT. The path of the electrons in the field is a circle with a radius 9.6 cm. Calculate:

(a) the specific charge of the electron,

(b) the mass of the electron, assuming the charge on the electron is 1.6×10^{-19} C.

(a) $e/m = v/Br$

$= 8.4 \times 10^6/(0.50 \times 10^{-3} \times 0.096)$

$= \mathbf{1.8 \times 10^{11}\ C\ kg^{-1}}$

(b) $e/m = 1.8 \times 10^{11} = 1.6 \times 10^{-19}/m$

$m = \mathbf{9.1 \times 10^{-31}\ kg}$

Now it's your turn

5 Electrons are accelerated through a potential difference of 220 V. They then pass into a region of uniform magnetic flux of flux density 0.54 mT. The path of the electrons is normal to the magnetic field. Given that the charge on the electron is 1.6×10^{-19} C and its mass is 9.1×10^{-31} kg, calculate:

(a) the speed of the accelerated electron,

(b) the radius of the circular path in the magnetic field.

Velocity selection of charged particles

We have seen (AS Level Topic 17) that when particles of mass m and charge $+q$ enter an electric field of field strength E, there is a force F_E on the particle given by

$$F_E = qE$$

If the velocity of the particles before entry into the field is v at right angles to the field lines (Figure 22.21), the particles will follow a parabolic path as they pass through the field.

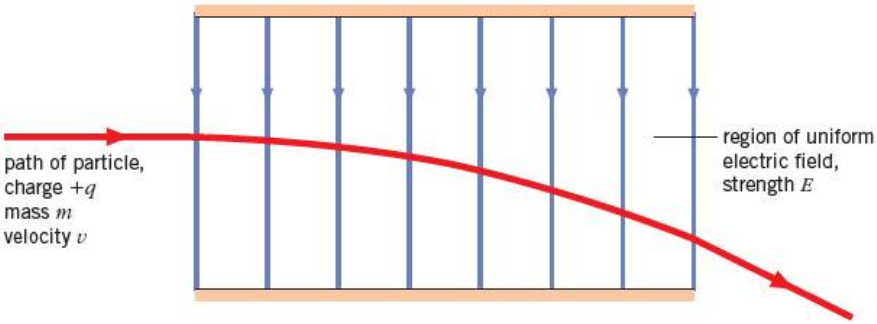


Figure 22.21

Now suppose that a uniform magnetic field acts in the same region as the electric field. If this field acts downwards into the plane of the page, then, by Fleming's left-hand rule (see page 314–315), a force will act on the charged particle in the direction opposite to the force due to the electric field. The magnitude F_B of this force is given by

$$F_B = Bqv$$

where B is the flux density of the magnetic field.

If the magnitude of one of the two fields is adjusted, then a situation can arise where the two forces, F_E and F_B, are equal in magnitude but opposite in direction. Thus

$$Bqv = qE$$

and

$$v = E/B$$

For the value of the velocity given by E/B, the particles will not be deflected, as shown in Figure 22.22.

www.iranino.co

IRANINOVIN

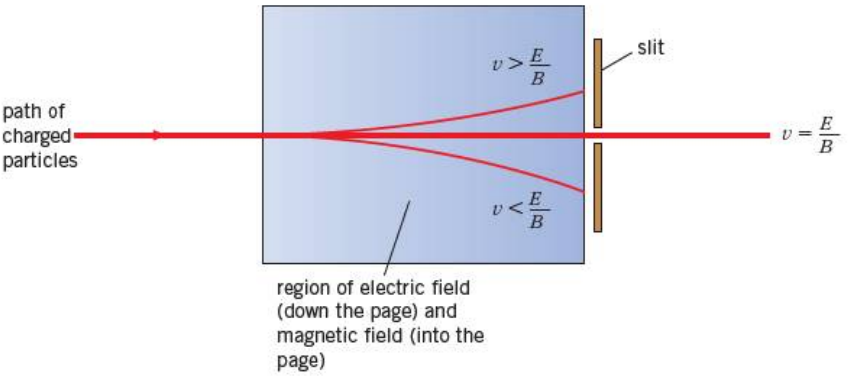


Figure 22.22 Velocity selector

Particles with any other velocities will be deflected. If a parallel beam of particles enters the field then all the particles passing undeviated through the slit will have the same velocity. Note that the mass does not come into the equation for the speed and so, particles with a different mass (but the same charge) will all pass undeviated through the region of the fields if they satisfy the condition $v = E/B$.

The arrangement shown in Figure 22.22 is known as a velocity selector. Velocity selectors are very important in the study of ions. Frequently, the production of the ions gives rise to different speeds but to carry out investigations on the ions, the ions must have one speed only.

Example

It is required to select charged ions which have a speed of 4.2×10^6 m s^{-1}. The electric field strength in the velocity selector is 3.2×10^4 V m^{-1}. Calculate the magnetic flux density required.

$v = E/B$

$B = 3.2 \times 10^4 / 4.2 \times 10^6$

$= \mathbf{7.6 \times 10^{-3}\ T}$

Now it's your turn

6 Singly charged ions pass undeviated through a velocity selector. The electric field strength in the selector is 3.6×10^4 V m^{-1} and the magnetic flux density is 8.5 mT. Calculate the selected speed of the ions.

The Hall effect

Consider a thin slice of a conductor which is normal to a magnetic field, as illustrated in Figure 22.23.

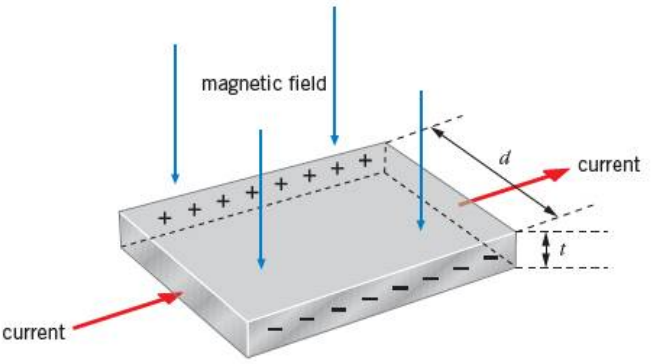


Figure 22.23 The Hall effect

When there is a current in the conductor in the direction shown, charge carriers (electrons in a metal) will be moving at right angles to the magnetic field. They will

www.iranino.co

experience a force which will tend to make them move to one side of the conductor. A potential difference known as the **Hall voltage** V_H will develop across the conductor. The Hall voltage does not increase indefinitely but reaches a constant value when the force on the charge carrier due to the magnetic field is equal to the force due to the electric field set up as a result of the Hall voltage.

If the distance between the two faces with the potential difference V_H is d (see Figure 22.23), then the electric field strength E between these two faces will be given by $E = V_H/d$ (see AS Level Topic 17). The force F_E acting on each charged carrier will be

$$F_E = q \times (V_H/d)$$

where q is the charge on the charge carrier.

The force F_B on the charge carrier due to the magnetic field of flux density B is given by $F_B = Bqv$ where v is the speed of the charge carriers.

When the electric field has been established, charge carriers will pass undeviated through the slice and $F_E = F_B$,

$$q \times (V_H/d) = Bqv,$$

$$V_H/d = Bv$$

Now, the speed v of the charge carriers is given, in terms of the current I in the slice, by the expression

$I = Anvq$ (see AS Level Topic 19)

where A is the area of cross-section of the slice and n is the number density of the charge carriers (number of charge carriers per unit volume).

Since the area A is equal to td (see Figure 22.23), then

$V_H/d = B \times (I/tdnq)$

and

Hall voltage $V_H = BI/ntq$

The size of the Hall voltage depends on the material of the conductor, the current in the sample and the magnetic flux density. The number density of charge carriers is very large in metals and, consequently, the Hall voltage is very small. However, with semiconductors, the number density is very much reduced and, therefore, detectable and measurable Hall voltages are possible. In fact, the Hall effect is used to study semiconductor materials.

If the current is kept constant, then the Hall voltage across a sample will be proportional to the magnetic flux density. The Hall effect provides a means by which flux densities may be measured, using a Hall probe (see Topic 2).

Example

A Hall probe is placed at right angles to a uniform magnetic field. A Hall voltage of 100 mV is measured.

The probe is adjusted so that the angle its plane makes with the magnetic field is 40°.

Calculate the value of the Hall voltage.

The component of the magnetic flux density B that is normal to the plane is B sin 40°.

Since Hall voltage is proportional to the magnetic flux density normal to the plane of the probe, then

Hall voltage = 100 sin 40° = **64 mV**

Now it's your turn

7 A piece of aluminium foil is 0.1 mm thick. The current in the foil is 5.0 A. Aluminium has 6.0×10^{28} free electrons per cubic metre.
A uniform magnetic field of flux density 0.1 T is normal to the slice. Calculate the Hall voltage that is generated.

www.iranino.co

IRANINOVIN

22.4 Magnetic fields due to currents

The magnetic field due to a long straight wire may be plotted using the apparatus illustrated in Figure 22.24. Note that the current must be quite large (about 5A). Iron filings are sprinkled on to the horizontal board and a plotting compass is used to determine the direction of the field.

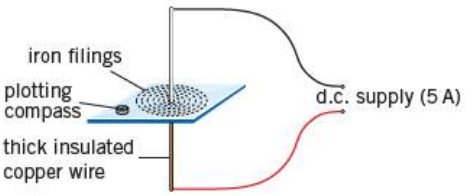


Figure 22.24 Apparatus to plot the magnetic field due to a long wire

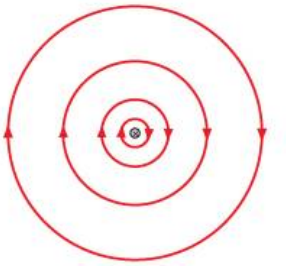


Figure 22.25 Magnetic field pattern due to a long straight wire

Figure 22.25 shows the field pattern due to a long straight current-carrying wire. The lines are concentric circles centred on the middle of the wire. The separation of the lines increases with distance from the wire, indicating that the field is decreasing in strength. The direction of the magnetic field may be found using the right-hand rule as illustrated in Figure 22.26.

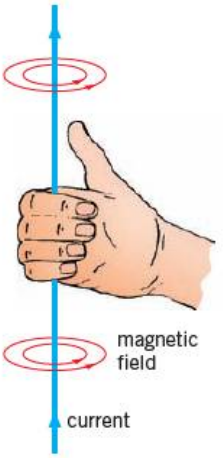


Figure 22.26 The right-hand rule

Imagine holding the conductor in the right hand with the thumb pointing in the direction of the current. The direction of the fingers gives the direction of the magnetic field.

Similar apparatus to that in Figure 22.24 may be used to investigate the shapes of the magnetic field due to a flat coil and to a solenoid (a long coil). Figure 22.27 illustrates the magnetic field pattern due to a flat coil. The field has been drawn in a plane normal to the coil and through its centre.

A solenoid may be thought of as being made up of many flat coils placed side-by-side. The magnetic field pattern of a long solenoid (that is, a coil which is long in comparison with its diameter) is shown in Figure 22.28.

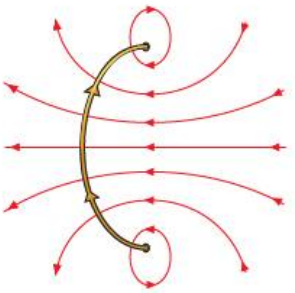


Figure 22.27 Magnetic field pattern due to a flat coil

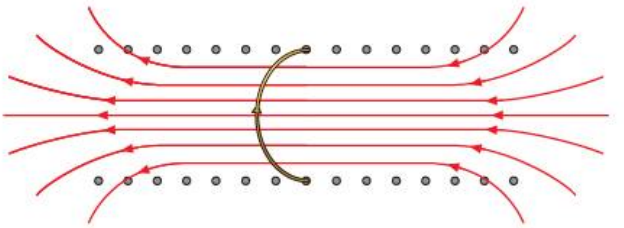


Figure 22.28 Magnetic field pattern due to a solenoid

The field lines are parallel and equally spaced over the centre section of the solenoid, indicating that the field is uniform. The field lines spread out towards the ends. The strength of the magnetic field at each end is one half that at the centre. The direction of the magnetic field in a flat coil and in a solenoid may be found using the **right-hand grip rule**, as illustrated in Figure 22.29.

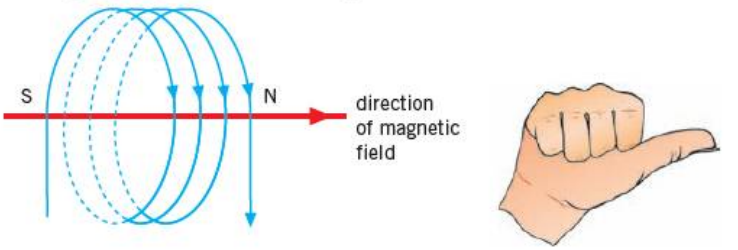


Figure 22.29 The right-hand grip rule

www.iranino.co

IRANINOVIN

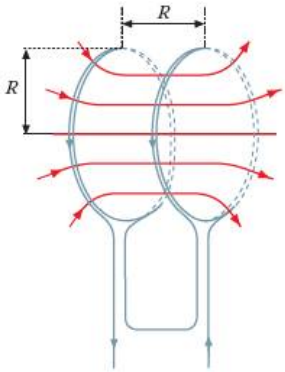


Figure 22.30 Magnetic field in Helmholtz coils

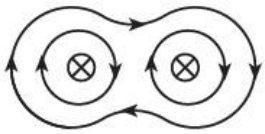


Figure 22.31

Grasp the coil or solenoid in the right hand with the fingers pointing in the direction of the current. The thumb gives the direction of the magnetic field.

The magnetic north end of the coil or solenoid is the end from which the lines of magnetic force are emerging. Note the similarities and, more importantly, the differences between this rule and the right-hand rule for the long straight wire (Figure 22.26).

Uniform magnetic fields are of importance in the study of charged particles. A uniform field is produced in a solenoid but this field is inside the solenoid and consequently, it may be difficult to make observations and to take measurements. This problem is overcome by using **Helmholtz coils**. These are two identical flat coils, with the same current in each. The coils are positioned so that their planes are parallel and separated by a distance equal to the radius of either coil. The coils and their resultant magnetic field are illustrated in Figure 22.30.

Example

Two long straight wires, of circular cross-section, are each carrying the same current directly away from you, down into the page. Draw the magnetic field due to the two current-carrying wires.

The solution is shown in Figure 22.31.

Now it's your turn

8 Draw magnetic field patterns, one in each case to represent
   (a) a uniform field,
   (b) a field which is decreasing in strength in the direction of the field,
   (c) a field which is increasing in strength along the direction of the field.

9 Draw a diagram of the magnetic field due to two long straight wires when the currents in the two wires are in opposite directions.

Electromagnets and their uses

The strength of the magnetic field due to a flat coil or a solenoid may be increased by winding the coil on a bar of soft iron. The bar is said to be the **core** of the coil. The iron is referred to as being 'soft' because it can be magnetised and demagnetised easily. With such a core (ferrous core), the strength of the magnetic field may be increased by up to 1000 times. With ferrous alloys (iron alloyed with cobalt or nickel), the field may be 10^4 times stronger. Magnets such as these are called **electromagnets**. Electromagnets have many uses because, unlike a permanent magnet, the magnetic field can be switched off by switching off the current in the coil.

Comparing the effects of fields

We can summarise the effects of the different sorts of fields on masses, charges and current-carrying conductors. Although, as we have seen, there are close analogies between gravitational and electric fields, there are some ways in which they behave very differently.

Start with the effect of a gravitational field on a mass. Because masses always attract each other, a mass placed in a gravitational field will always move in the direction of the field, from a position of higher potential to a lower potential. For a field produced by a point mass, the field strength obeys an inverse square law relationship, and the potential obeys a reciprocal relationship with distance from the source of the field.

Electric fields are like gravitational fields in that, for a field produced by a point charge, the field strength is also given by the inverse square law, and the potential by a reciprocal relationship. However, we can have both positive charges and negative charges. A positive electric charge moves in the direction of the field, from a position of higher potential to a lower potential (just like a mass in a gravitational field). But a negative charge does just the opposite, against the direction of the field and from a low potential to a high potential.

What about electric charges in a magnetic field? A stationary charge is unaffected, but a moving charge experiences a force F given by $F = Bqv \sin \theta$. The direction of the force is given by Fleming's left-hand rule (for positive charges).

www.iranino.co

IRANINOVIN

Finally, a current-carrying conductor in a magnetic field does not experience a force if the conductor is parallel to the field direction, but for all other directions it experiences a force given by $F = BIl \sin \theta$. The direction of the force is again given by Fleming's left-hand rule.

Example

A charged particle has mass 6.7×10^{-27} kg and charge $+3.2 \times 10^{-19}$ C. It is travelling at speed 2.5×10^{8} m s^{-1} when it enters a region of space where there is a uniform magnetic field of flux density 1.6T at right angles to its direction of motion. Calculate

(a) the gravitational force on the particle,
(b) the force on the particle due to the magnetic field,
(c) the radius of its orbit in the field.

(a) gravitational force $= mg = 6.7 \times 10^{-27} \times 9.81$

$= \mathbf{6.6 \times 10^{-26}\,N}$ (negligible when compared with the force due to the magnetic field)

(b) Force $= Bqv \sin \theta$

$= 1.6 \times 3.2 \times 10^{-19} \times 2.5 \times 10^{8} \times 1$

$= \mathbf{1.3 \times 10^{-10}\,N}$

(c) Centripetal force is provided by the electromagnetic force

$$mv^2/r = Bqv$$

$$6.7 \times 10^{-27} \times (2.5 \times 10^{8})^2/r = 1.3 \times 10^{-10}$$

$$r = \mathbf{3.2\,m}$$

Now it's your turn

10 An electron and an α-particle travelling at the same speed both enter the same region of uniform magnetic flux which is at right angles to their direction of motion. State and explain any differences between the two paths in the field.

22.5 The use of (nuclear) magnetic resonance imaging



Figure 22.32 Magnetic imaging of body structures requires the use of a very large uniform magnetic field.

Nuclear magnetic resonance

Many atomic nuclei have a property that is known as 'spin'. The 'spin' causes the nuclei to behave as if they were small magnets. Such nuclei have an odd number of protons and/or neutrons. Examples of these nuclei are hydrogen, carbon and phosphorus.

When a magnetic field is applied to these nuclei, they tend to line up along the field. However, this alignment is not perfect and the nuclei rotate about the direction of the magnetic field, due to their spin. The motion can be modelled as the motion of a top spinning about the direction of a gravitational field. This rotation is known as **precession**.

www.iranino.co

IRANINOVIN

The spinning about the direction of the magnetic field has a frequency known as the frequency of precession or the **Larmor frequency**. The frequency depends on the nature of the nucleus and the magnitude of the magnetic flux density.

If a pulse of electromagnetic radiation of the same frequency as the Larmor frequency is incident on precessing nuclei, the nuclei will resonate, absorbing energy. This frequency is in the radio-frequency (RF) band with a wavelength shorter than about 10 cm. A short time after the incident pulse has ended, the nuclei will return to their equilibrium state, emitting RF radiation. The short time between the end of the RF pulse and the re-emitting of the radiation is known as the **relaxation time**. The whole process is referred to as **nuclear magnetic resonance**.

In practice, there are two relaxation processes and it is the time between these two that forms the basis of **nuclear magnetic resonance imaging (NMRI)**, often abbreviated to **magnetic resonance imaging (MRI)**. Since hydrogen is abundant in body tissues and fluids, hydrogen is the nucleus used in MRI.

Magnetic resonance imaging (MRI)

A schematic diagram of a magnetic resonance scanner is shown in Figure 22.33.

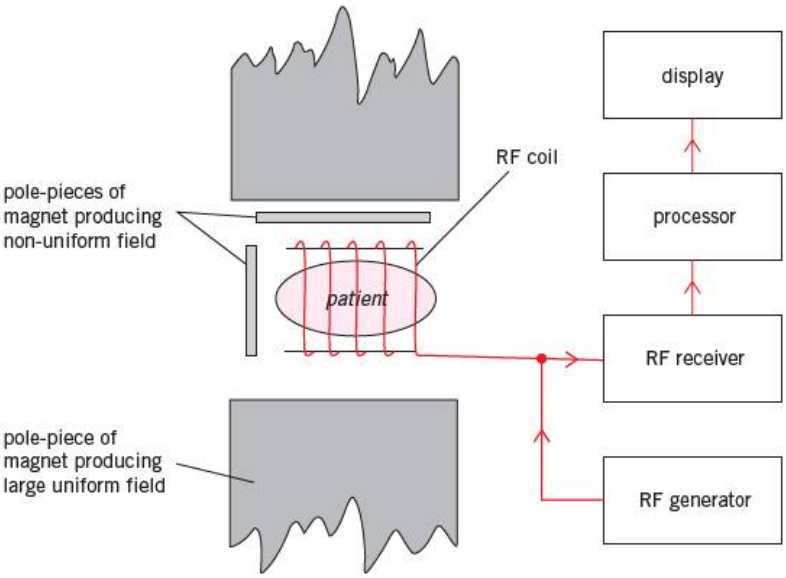


Figure 22.33 Schematic diagram of a magnetic resonance scanner

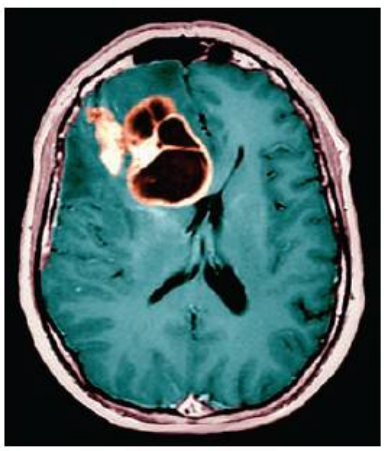


Figure 22.34 MRI scan showing a brain tumour

The patient who is to be investigated is positioned in the scanner (Figure 22.33), between the poles of a large magnet that produces a very large uniform magnetic field (in excess of 1 tesla). The magnetic field causes all the hydrogen nuclei within the person to precess with the same Larmor frequency. In order that the hydrogen nuclei in only one small part of the body may be detected, a non-uniform magnetic field is also applied across the patient. This non-uniform field is accurately calibrated and results in a different magnitude of magnetic flux density at each point in the body of the patient. Since the Larmor frequency is dependent on the strength of the magnetic field, the Larmor frequency will also be different in each part of the patient. The particular value of magnetic flux density, together with the radio-frequency that is emitted, enables the hydrogen nuclei in the part under investigation to be located.

Radio-frequency (RF) pulses are produced in coils near the patient. These pulses pass into the patient. The emitted pulses produced as a result of de-excitation of the hydrogen nuclei are picked up by the coils. These signals are processed and displayed so as to construct an image of the number density of hydrogen atoms in the patient. As the non-uniform magnetic field is changed, atoms in different parts of the patient's body are detected and displayed. An image of a cross-section through the patient may be produced. One such image is shown in Figure 22.34.

A whole series of images of sections through the patient can be produced and stored in a computer memory. As with CT scanning (Topic 25) this enables a

www.iranino.co

IRANINOVIN

three-dimensional image to be generated. Sections through the patient may be viewed from many different angles.

Example

The use of magnetic resonance imaging is now a common diagnostic tool in medicine. However, there may be disadvantages for some patients. Suggest why:

(a) some people may find the scanner to be claustrophobic

(b) there are additional problems when scanning young children rather than adults.

In order to produce a magnetic field of large magnitude through the patient, the patient must be placed inside a large cylindrical magnet. This can mean that the patient feels trapped.

The total scan takes some time because separate scans of many slices through the patient must be taken. The patient must be very still whilst this is done and this is difficult for children.

Now it's your turn

11 During an MRI scan, the patient is placed in a uniform magnetic field of large magnitude on to which is superimposed a highly calibrated non-uniform magnetic field. Explain the purpose of each of these two magnetic fields.

Summary

- A magnetic field is a region of space where a magnet will experience forces.
- A magnetic field can be represented by magnetic field lines. Magnetic field lines never touch or cross.
- The separation of magnetic field lines indicates the strength of the magnetic field.
- The direction of the magnetic field is given by the direction in which a free magnetic north pole would move, if placed in the field.
- There is a force on a current-carrying conductor whenever it is at an angle to a magnetic field.
- The direction of the force is given by Fleming's left-hand rule – place the first two fingers and thumb of the left hand at right angles to each other, first finger in the direction of the magnetic field, second finger in the direction of the current, then the thumb gives the direction of the force.
- The magnitude of the force F on a conductor of length L carrying a current I at an angle θ to a magnetic field of flux density B is given by the expression $F = BIL \sin \theta$.
- Magnetic flux density (field strength) is measured in tesla (T). $1\,\text{T} = 1\,\text{Wb m}^{-2}$.
- The force F on a particle with charge q moving at speed v at an angle θ to a magnetic field of flux density B is given by the expression: $F = Bqv \sin \theta$.
- The direction of the force is given by Fleming's left-hand rule – place the first two fingers and thumb of the left hand at right angles to each other, first finger in the direction of the magnetic field, second finger in the direction of motion of positive charge, then the thumb gives the direction of the force.
- The path of a charged particle, moving at constant speed in a plane at right angles to a uniform magnetic field, is circular.
- The Hall voltage V_H is given by the expression $V_H = BI/ntq$.
- An electric current gives rise to a magnetic field, the strength and direction of which depends on the size of the current and the shape of the current-carrying conductor.
- The direction of the field due to a straight wire is given by the right-hand rule.
- The direction of the field in a solenoid is given by the right-hand grip rule.
- The field of a solenoid may be increased in strength by a ferrous core; this is the principle of an electromagnet.
- The force between parallel current-carrying conductors provides a means by which the ampere may be defined in terms of SI units.
- Nuclear magnetic resonance imaging (NMRI) – shortened to magnetic resonance imaging (MRI) – monitors the concentration of hydrogen nuclei in the body.
- Hydrogen nuclei have 'spin' which causes them to precess in an applied magnetic field.
- The frequency of precession (the Larmor frequency) depends on the magnetic flux density applied to the nuclei.
- A radio-frequency pulse at the Larmor frequency causes the nuclei to resonate.
- On de-excitation, the nuclei give off radio-frequency waves that can be detected and analysed.

www.iranino.co

IRANINOVIN

Examination style questions

1 A stiff straight wire has a mass per unit length of 45 g m^{-1}. The wire is laid on a horizontal bench and a student passes a current through it to try to make it lift off the bench. The horizontal component of the Earth's magnetic field is 18 μT in a direction from south to north and the acceleration of free fall is 10 m s^{-2}.

a i State the direction in which the wire should be laid on the bench.

ii Calculate the minimum current required.

b Suggest whether the student is likely to be successful with this experiment.

2 The magnetic flux density B at a distance r from a long straight wire carrying a current I is given by the expression $B = (2.0 \times 10^{-7}) \times I/r$, where r is in metres and I is in amperes.

a Calculate:

i the magnetic flux density at a point distance 2.0 cm from a wire carrying a current of 15 A,

ii the force per unit length on a second wire, also carrying a current of 15 A, which is parallel to, and 2.0 cm from, the first wire.

b Suggest why the force between two wires is demonstrated in the laboratory using aluminium foil rather than copper wires.

3 A small horseshoe magnet is placed on a balance and a stiff wire is clamped in the space between its poles. The length of wire between the poles is 5.0 cm. When a current of 3.5 A is passed through the wire, the reading on the balance increases by 0.027 N.

a Calculate the magnetic flux density between the poles of the magnet.

b State three assumptions which you have made in your calculation.

4 Fig. 22.35 shows the track of a particle in a bubble chamber as it passes through a thin sheet of metal foil. A uniform magnetic field is applied into the plane of the paper.

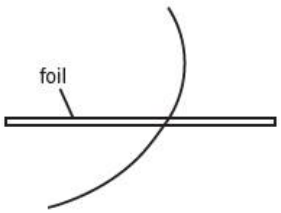


Fig. 22.35

State with a reason

a in which direction the particle is moving,

b whether the particle is positively or negatively charged.

5 a A uniform magnetic field has constant flux density B. A straight wire of fixed length carries a current I at an angle θ to the magnetic field, as shown in Fig. 22.36.

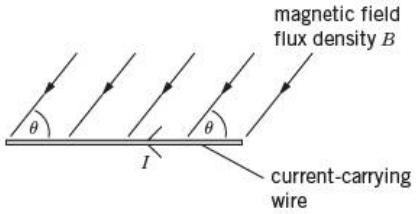


Fig. 22.36

i The current I in the wire is changed, keeping the angle θ constant.
On a copy of Fig. 22.37, sketch a graph to show the variation with current I of the force F on the wire.

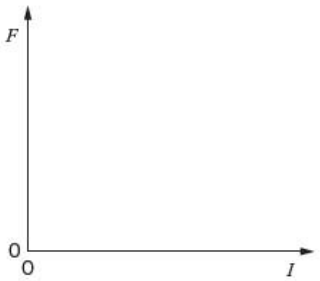


Fig. 22.37

[2]

ii The angle θ between the wire and the magnetic field is now varied. The current I is kept constant.
On a copy of Fig. 22.38, sketch a graph to show the variation with angle θ of the force F on the wire.

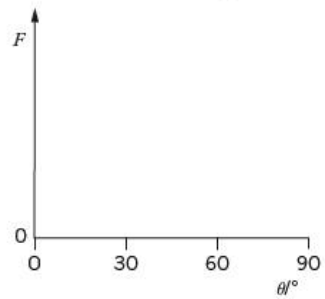


Fig. 22.38

[3]

b A uniform magnetic field is directed at right angles to the rectangular surface PQRS of a slice of a conducting material, as shown in Fig. 22.39.

www.iranino.co

IRANINOVIN

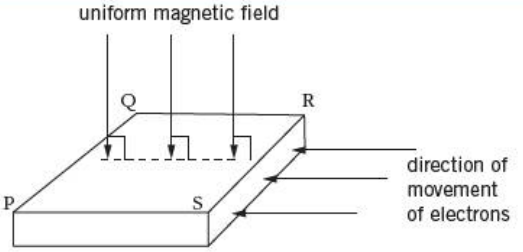


Fig. 22.39

Electrons, moving towards the side SR, enter the slice of conducting material. The electrons enter the slice at right angles to side SR.

i Explain why, initially, the electrons do not travel in straight lines across the slice from side SR to side PQ. *[2]*

ii Explain to which side, PS or QR, the electrons tend to move. *[2]*

Cambridge International AS and A Level Physics, 9702/42 May/June 2010 Q 6

6 Positive ions are travelling through a vacuum in a narrow beam. The ions enter a region of uniform magnetic field of flux density B and are deflected in a semi-circular arc, as shown in Fig. 22.40.

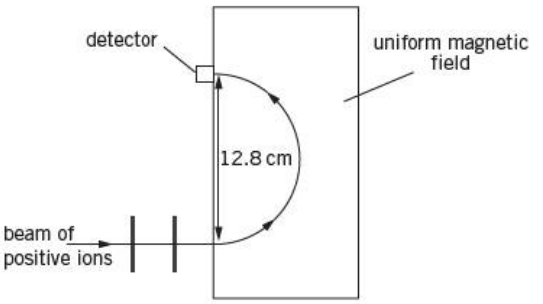


Fig. 22.40

The ions, travelling with speed 1.40×10^5 m s^{-1}, are detected at a fixed detector when the diameter of the arc in the magnetic field is 12.8 cm.

a By reference to Fig. 22.40, state the direction of the magnetic field. *[1]*

b The ions have mass 20 u and charge $+1.6 \times 10^{-19}$ C. Show that the magnetic flux density is 0.454 T. Explain your working. *[3]*

c Ions of mass 22 u with the same charge and speed as those in **b** are also present in the beam.

i On a copy of Fig. 22.40, sketch the path of these ions in the magnetic field of magnetic flux density 0.454 T. *[1]*

ii In order to detect these ions at the fixed detector, the magnetic flux density is changed.
Calculate this new magnetic flux density. *[2]*

Cambridge International AS and A Level Physics, 9702/42 Oct/Nov 2010 Q 5

7 a i State the condition for a charged particle to experience a force in a magnetic field. *[2]*

ii State an expression for the magnetic force F acting on a charged particle in a magnetic field of flux density B. Explain any other symbols you use. *[2]*

b A sample of a conductor with rectangular faces is situated in a magnetic field, as shown in Fig. 22.41.

The magnetic field is normal to face ABCD in the downward direction.

Electrons enter face CDHG at right angles to the face. As the electrons pass through the conductor, they experience a force due to the magnetic field.

i On a copy of Fig. 22.41, shade the face to which the electrons tend to move as a result of this force. *[1]*

ii The movement of the electrons in the magnetic field causes a potential difference between two faces of the conductor.
Using the lettering from Fig. 22.41, state the faces between which this potential difference will occur. *[1]*

c Explain why the potential difference in **b** causes an additional force on the moving electrons in the conductor. *[2]*

Cambridge International AS and A Level Physics, 9702/41 Oct/Nov 2012 Q 6

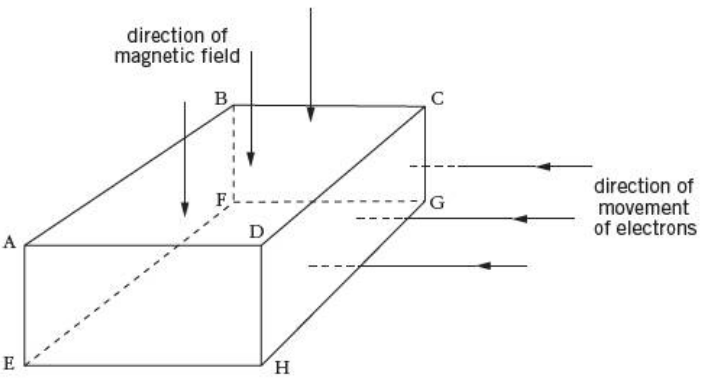


Fig. 22.41

www.iranino.co

IRANINOVIN

8 a Define the *tesla*. *[2]*

b Two long straight vertical wires X and Y are separated by a distance of 4.5 cm, as illustrated in Fig. 22.42.

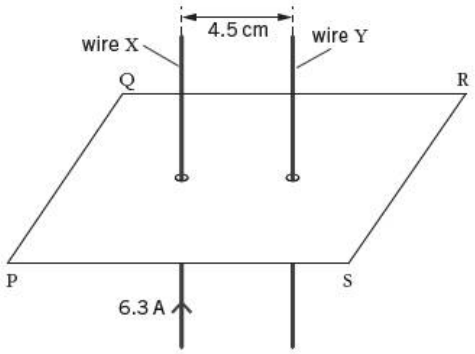


Fig. 22.42

The wires pass through a horizontal card PQRS.

The current in wire X is 6.3 A in the upward direction. Initially, there is no current in wire Y.

i On a copy of Fig. 22.42, sketch, in the plane PQRS, the magnetic flux pattern due to the current in wire X. Show at least four flux lines. *[3]*

ii The magnetic flux density B at a distance x from a long straight current-carrying wire is given by the expression

$$B = \frac{\mu_0 I}{2\pi x}$$

where I is the current in the wire and μ_0 is the permeability of free space.

Calculate the magnetic flux density at wire Y due to the current in wire X. *[2]*

iii A current of 9.3 A is now switched on in wire Y. Use your answer in **ii** to calculate the force per unit length on wire Y. *[2]*

c The currents in the two wires **b iii** are not equal.

Explain whether the force per unit length on the two wires will be the same, or different. *[2]*

Cambridge International AS and A Level Physics, 9702/42 May/June 2013 Q 5

www.iranino.co

IRANINOVIN

23 Electromagnetic induction

By the end of this topic, you will be able to:

23.1
- **(a)** define magnetic flux and the weber
- **(b)** recall and solve problems using $\Phi = BA$
- **(c)** define magnetic flux linkage
- **(d)** infer from appropriate experiments on electromagnetic induction
  - that a changing magnetic flux can induce an e.m.f. in a circuit
  - that the direction of the induced e.m.f. opposes the change producing it
  - the factors affecting the magnitude of the induced e.m.f.
- **(e)** recall and solve problems using Faraday's law of electromagnetic induction and Lenz's law
- **(f)** explain simple applications of electromagnetic induction

Starting points

- Magnets create a magnetic field around them.
- Magnetic fields are produced by current-carrying conductors and coils.

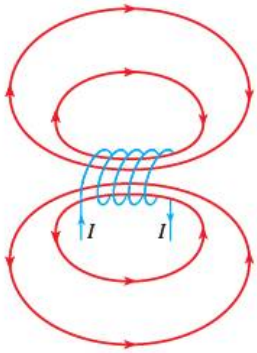


Figure 23.1 Continuous magnetic field lines

23.1 Magnetic flux and electromagnetic induction

In Topic 22, we examined the pattern of the magnetic field in the region of a straight wire and various coils. However, the patterns are not all complete. All magnetic field lines should be continuous, as illustrated in Figure 23.1.

Early experimenters thought that there was a flow of something along these lines and this gave rise to the idea of a magnetic flux, since 'flux' means 'flow'. Magnetic flux density (field strength) varies and this is shown by the spacing of the lines. Magnetic flux density may be considered as the number of lines of magnetic force per unit area of an area at right angles to the lines. Magnetic flux Φ can be thought of as the total number of lines.

Magnetic flux *is the product of the magnetic flux density and the area normal to the lines of flux.*

For a uniform magnetic field of flux density B which makes an angle θ with an area A, the magnetic flux Φ is given by the expression $\Phi = BA \sin \theta$.

The unit of magnetic flux density, the tesla, was defined in Topic 22.

The unit of magnetic flux is the **weber** (Wb). One weber is equal to one tesla metre-squared i.e. T m^2.

The concept of magnetic flux is used frequently when studying electromagnetic induction.

Electromagnetic induction

The link between electric current and magnetic field was discovered by Oersted (1820). In 1831, Henry in the United States and Faraday in England demonstrated that an

www.iranino.co

IRANINOVIN

e.m.f. could be induced by a magnetic field. The effect was called **electromagnetic induction**.

Electromagnetic induction is now easy to demonstrate in the laboratory because sensitive meters are available. Figure 23.2 illustrates apparatus which may be used for this purpose.

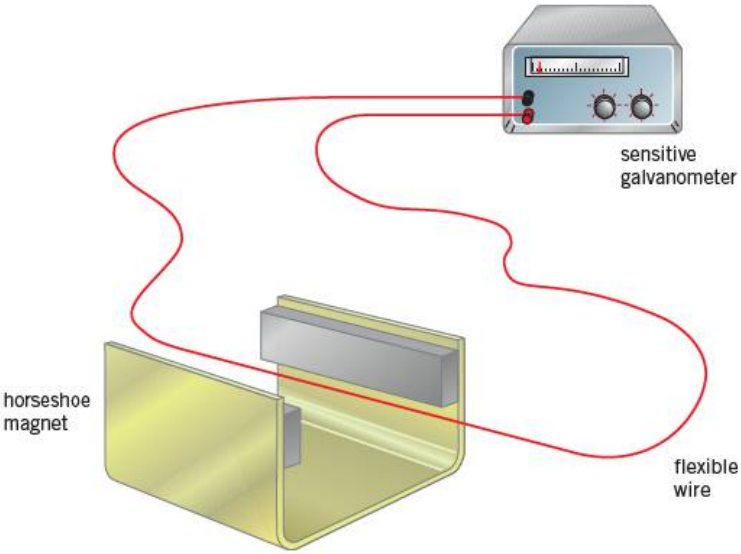


Figure 23.2 Apparatus to demonstrate electromagnetic induction

The galvanometer detects very small currents but it is important to realise that what is being detected are small electromotive forces (e.m.f.s). The current arises because there is a complete circuit incorporating an e.m.f. The following observations can be made.

- An e.m.f. is induced when
  - the wire is moved through the magnetic field, across the face of the pole-pieces
  - the magnet is moved so that the wire passes across the face of the pole-pieces.
- An e.m.f. is *not* induced when
  - the wire is held stationary between the pole-pieces
  - the magnet is moved so that the pole-pieces move along the length of the wire
  - the wire moves lengthways so that it does not change its position between the poles of the magnet.

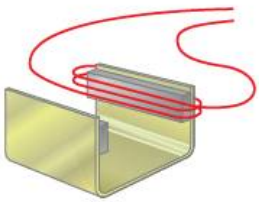


Figure 23.3 Wire wound to form a loop of several turns

These observations lead to the conclusion that an e.m.f. is induced whenever lines of magnetic flux are cut. The cutting may be caused by a movement of either the wire or the magnet. The magnitude of the e.m.f. is also observed to vary.

- The magnitude of the e.m.f.
  - increases as the speed at which the wire is moved increases
  - increases as the speed at which the magnet is moved increases
  - increases if the wire is made into a loop with several turns (see Figure 23.3)
  - increases as the number of turns on the loop increases.

It can be concluded that the magnitude of the induced e.m.f. depends on the rate at which magnetic flux lines are cut. The rate may be changed by varying the rate at which the flux lines are cut by a single wire or by using different numbers of turns of wire. The two factors are taken into account by using the term **magnetic flux linkage** ($N\Phi$). Change in magnetic flux linkage $\Delta(N\Phi)$ is equal to the product of the change in magnetic flux $\Delta\Phi$ and the number of turns N of a conductor involved in the change in flux.

change in magnetic flux linkage $\Delta(N\Phi) = N\Delta\Phi$

www.iranino.co

IRANINOVIN

The experimental observations are summarised in **Faraday's law of electromagnetic induction**.

The e.m.f. induced is proportional to the rate of change of magnetic flux linkage.

The experimental observations made with the apparatus of Figure 23.2 and Figure 23.3 have been concerned with the magnitude of the e.m.f. However, it is noticed that the direction of the induced e.m.f. changes and that the direction is dependent on the direction in which the magnetic flux lines are being cut. The direction of the induced e.m.f. or current in a wire moving through a magnetic field at right angles to the field may be determined using **Fleming's right-hand rule**. This rule is illustrated in Figure 23.4.

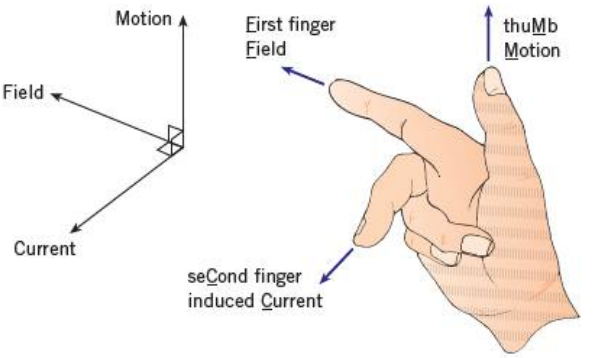


Figure 23.4 Fleming's right-hand rule

*If the first two fingers and thumb of the right hand are held at right angles to one another, the **F**irst finger in the direction of the magnetic **F**ield and the thu**M**b in the direction of **M**otion, then the se**C**ond finger gives the direction of the induced e.m.f. or **C**urrent.*

An explanation for the direction of the induced e.m.f. can be found by reference to the motor effect and conservation of energy.

Figure 23.5 shows a wire being moved downwards through a magnetic field.

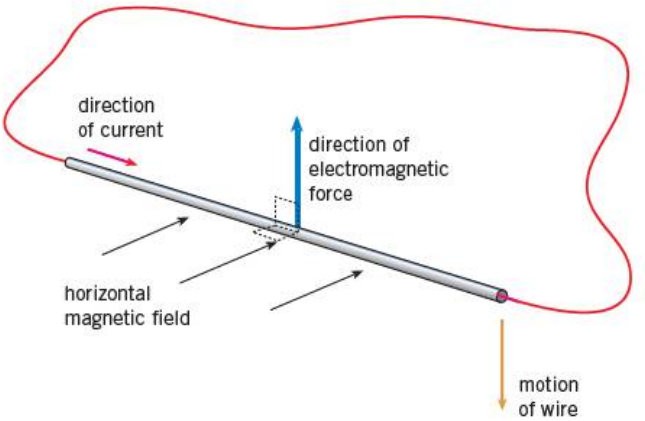


Figure 23.5

Since the wire is in the form of a continuous loop, the induced e.m.f. gives rise to a current, and the direction of this current can be found using Fleming's right-hand rule. This current is at right angles to the magnetic flux and, by the motor effect, there will be a force on the wire. Using Fleming's left-hand rule (see Topic 22), the force is upwards when the wire is moving downwards. Reversing the direction of motion of the wire causes a current in the opposite direction and, hence, the electromagnetic force would once again oppose the motion. This conclusion is not surprising when conservation of energy is considered. An electric current is a form of energy and this

www.iranino.co

IRANINOVIN

energy must have been converted from some other form. Movement of the wire against the electromagnetic force means that work has been done in overcoming this force and it is this work which is seen as electrical energy. Anyone who has ridden a bicycle with a dynamo will realise that work has to be done to light the lamp!

This application of conservation of energy is summarised in **Lenz's law**.

The direction of the induced e.m.f. is such as to cause effects to oppose the change producing it.

Faraday's law of electromagnetic induction and Lenz's law may be summarised using the equation

$$E = \frac{-\mathrm{d}(N\Phi)}{\mathrm{d}t}$$

(see the Maths Note below) where E is the e.m.f. induced by a rate of change of flux linkage of $\mathrm{d}(N\Phi)/\mathrm{d}t$. The minus sign indicates that the induced e.m.f. causes effects to oppose the change producing it.

For a small change $\Delta(N\Phi)$ in flux linkage that occurs in time Δt (or where the flux linkage changes linearly with time), then the induced e.m.f. E is given by

$$E = -\Delta(N\Phi)/\Delta t$$

Maths Note

The shorthand way of expressing the rate of change of a quantity x with time t is

rate of change of x with t is $\frac{\mathrm{d}x}{\mathrm{d}t}$

This represents a mathematical operation known as *differentiation*. It is achieved by finding the gradient of the graph of x against t.

You will come across this notation here, in connection with the rate of change of magnetic flux linkage, and in one of the equations for radioactive decay (see Topic 26).

Example

The uniform flux density between the poles of a magnet is 0.080 T. A small coil of area of cross-section 6.5 cm^2 has 250 turns and is placed with its plane normal to the magnetic field. The coil is withdrawn from the field in a time of 0.26 s.

Determine:

(a) the magnetic flux through the coil when it is between the poles of the magnet,
(b) the change in magnetic flux linkage when the coil is removed from the field,
(c) the average e.m.f. induced in the coil whilst it is being withdrawn.

(a) magnetic flux $\Phi = BA \sin\theta$

$= 0.080 \times 6.5 \times 10^{-4}$

$= \mathbf{5.2 \times 10^{-5}\,Wb}$

(b) change in flux linkage $= (N\Phi)_{\text{FINAL}} - (N\Phi)_{\text{INITIAL}}$

$= 0 - (250 \times 5.2 \times 10^{-5})$

$= \mathbf{-1.3 \times 10^{-2}\,Wb}$

(the sign indicates that the flux linkage is decreasing)

(c) $\text{induced e.m.f.} = \frac{\text{change in flux linkage}}{\text{time taken}}$

$= (1.3 \times 10^{-2})/0.26$

$= \mathbf{0.050\,V}$

Now it's your turn

1 An aircraft has a wingspan of 17 m and is flying horizontally in a northerly direction at a speed of 94 m s^{-1}. The vertical component of the Earth's magnetic field is 40 μT in a downward direction.

www.iranino.co

IRANINOVIN

(a) Calculate:
  (i) the area swept out per second by the wings,
  (ii) the magnetic flux cut per second by the wings,
  (iii) the e.m.f. induced between the wingtips.
(b) State which wing-tip will be at the higher potential.

2 A current-carrying solenoid produces a uniform magnetic flux of density 4.6×10^{-2} T along its axis. A small circular coil of radius 1.2 cm has 350 turns of wire and is placed on the axis of the solenoid with its plane normal to the axis. Calculate the average e.m.f. induced in the coil when the current in the solenoid is reversed in a time of 85 ms.

3 A metal disc is made to spin at 15 revolutions per second about an axis through its centre normal to the plane of the disc. The disc has radius 24 cm and spins in a uniform magnetic field of flux density 0.15 T, parallel to the axis of rotation. Calculate:
(a) the area swept out each second by a radius of the disc,
(b) the flux cut each second by a radius of the disc,
(c) the e.m.f. induced in the disc.

Applications of electromagnetic induction

Eddy current damping

The conversion of mechanical energy to electrical energy may be shown by spinning a metal disc in a magnetic field, as illustrated in Figure 23.6.

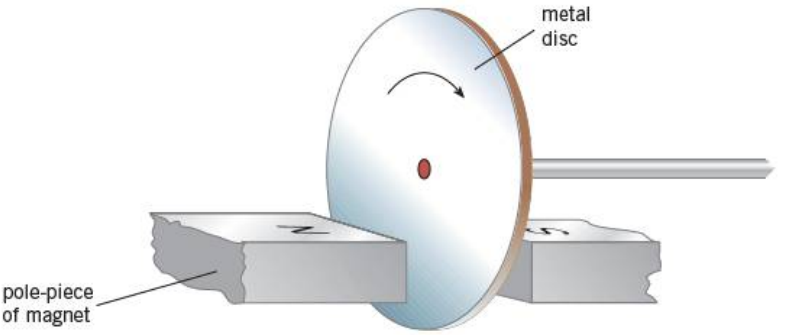


Figure 23.6 Apparatus to demonstrate eddy current damping

An e.m.f. is induced between the rim of the disc and the axle. The apparatus illustrated is the basis of a means by which a direct e.m.f. may be generated.

The disc is seen to slow down much more rapidly with the magnet in place than when it has been removed. As the disc spins, it cuts through the flux lines of the magnet. This cutting becomes more obvious if the radius of the disc is considered. As the radius rotates, it will cut flux. An e.m.f. will be induced in the disc but, because the rate of cutting of flux varies from one part of the disc to another, the e.m.f. will have different magnitudes in different regions of the disc. The disc is metal and, therefore, electrons will move between regions within the disc that have different, e.m.f. values. Currents are induced in the disc. Since these currents vary in magnitude and direction, they are called **eddy currents**. The eddy currents cause heating in the disc and the dissipation of the energy of rotation of the disc is referred to as **eddy current damping**.

If the permanent magnet in Figure 23.6 is replaced by an electromagnet, the spinning disc will be slowed down whenever there is a current in the electromagnet. This is the principle behind electromagnetic braking. The advantage over conventional brakes is that there is no physical contact with the spinning disc. This makes such brakes very useful for trains travelling at high speeds. However, the disadvantage is that, as the disc slows down, the induced eddy currents will be smaller and, therefore, the braking will be less efficient. This system would be useless as the parking brake on a car!

www.iranino.co

IRANINOVIN

e.m.f. induced between two coils

A current-carrying solenoid or coil is known to have a magnetic field. Consider the apparatus illustrated in Figure 23.7.

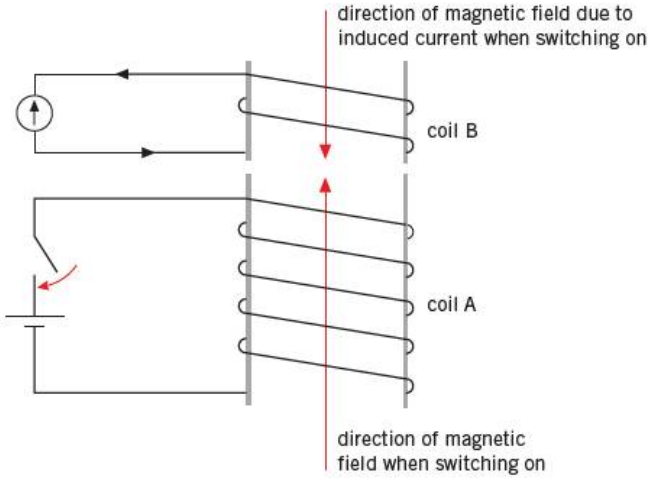


Figure 23.7

As the current in coil A is being switched on, the magnetic field in this coil grows. The magnetic field links with the turns on coil B and, as a result, there is a change in flux linkage in coil B and an e.m.f. is induced in this coil. Coil B forms part of a complete circuit and hence there is a current in coil B. The direction of this current can be determined using Lenz's law.

The change which brought about the induction of a current was a *growth* in the magnetic flux in coil A. The induced current in coil B will give rise to a magnetic field in coil B and this field will, by Lenz's law, try to oppose the growth of the field in coil A. Consequently, since the field in coil A is vertically upwards (the right-hand grip rule), the field in coil B will be vertically downwards and the induced current will be in an anticlockwise direction through the meter.

When the current in coil A is switched off, the magnetic field in coil A will *decay*. The magnetic field in coil B due to the induced current must try to prevent this decay and hence it will be vertically upwards. The induced current has changed in direction.

The magnitude of the induced e.m.f. can be increased by inserting a soft iron core into the coil (but be careful not to damage the meter as any induced e.m.f. will be very much greater) or by increasing the number of turns on the coils or by switching a larger current in coil A.

It is important to realise that an e.m.f. is induced only when the magnetic flux in coil A is changing; that is, when the current in coil A is changing. A steady current in coil A will not give rise to an induced e.m.f. An e.m.f. may be induced continuously in coil B if an alternating current is provided for coil A. This is the principle of the **transformer** (see Topic 24).

www.iranino.co

IRANINOVIN

Summary

- Magnetic flux is the product of flux density and area normal to the flux.
- The direction of the induced current in a conductor moving through a magnetic field is given by Fleming's right-hand rule. That is, if the first two fingers and thumb of the right hand are held at right angles to each other, the first finger in the direction of the magnetic field and the thumb in the direction of motion, then the second finger gives the direction of the induced e.m.f. or current.
- Faraday's law of electromagnetic induction states that the e.m.f. induced is proportional to the rate of change of magnetic flux linkage.
- Lenz's law states that the direction of the induced e.m.f. is such as to cause effects to oppose the change producing it.
- Faraday's law of electromagnetic induction and Lenz's law may be summarised using the equation $E = -\mathrm{d}(N\Phi)/\mathrm{d}t$ where E is the e.m.f. induced by a rate of change of flux linkage of $\mathrm{d}(N\Phi)/\mathrm{d}t$. The sign indicates the relative direction of the e.m.f. and the change in flux linkage.

Examination style questions

1 A coil is constructed by winding 400 turns of wire on to a cylindrical iron core. The mean radius of the coil is 3.0 cm. It is found that the flux density B in the core due to a current I in the coil is given by the expression

$B = 2.2\,I$

where B is in tesla.

a Calculate, for a current of 0.64 A in the coil:

i the magnetic flux density in the core,

ii the magnetic flux in the core,

iii the flux linkage of the coil.

b The current in the coil is switched off in a time of 0.011 s. Calculate the e.m.f. induced and state where this e.m.f. will be observed.

c Hence suggest why, when switching off a large electromagnet, the current is reduced gradually rather than switched off suddenly.

2 A flat coil contains 250 turns of insulated wire and has a mean radius of 1.5 cm. The coil is placed in a region of uniform magnetic flux of flux density 85 mT such that there is an angle θ between the plane of the coil and the flux lines. The coil is withdrawn from the magnetic field in a time of 0.30 s.

Calculate the average e.m.f. induced in the coil for an angle θ of:

a zero,

b 90°,

c 35°.

3 The magnetic flux density B at the centre of a long solenoid carrying a current I is given by the expression $B = 2.5\,I$, where B is in millitesla and I is in amperes. The cross-section of the solenoid has a diameter of 2.5 cm. A length of wire is wound tightly round the centre of the solenoid to form a small coil of 50 turns, as illustrated in Fig. 23.8.

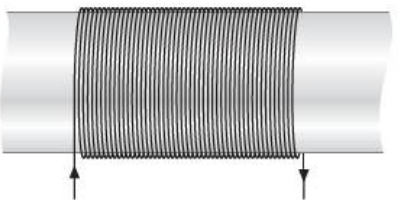


Fig. 23.8

a Calculate, for a current of 1.8 A in the solenoid:

i the flux density in the solenoid,

ii the magnetic flux linkage of the small coil.

b The current is reversed in direction in 0.031 s. Calculate the average e.m.f. induced in the small coil.

4 Two coils A and B are placed as shown in Fig. 23.9.

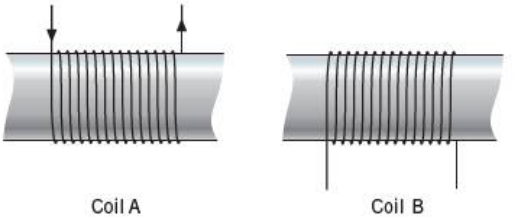


Fig. 23.9

The current in coil A is switched on and then off so that the variation with time t of the current I is as shown in Fig. 23.10.

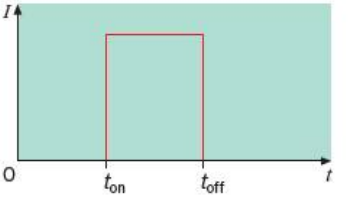


Fig. 23.10

www.iranino.co

IRANINOVIN

Sketch a graph to show the variation with time t of the e.m.f. E induced in coil B.

5 A bar magnet is suspended vertically from the free end of a helical spring, as shown in Fig. 23.11.

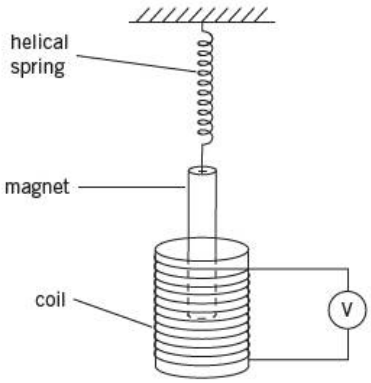


Fig. 23.11

One pole of the magnet is situated in a coil. The coil is connected in series with a high-resistance voltmeter.

The magnet is displaced vertically and then released.

The variation with time t of the reading V of the voltmeter is shown in Fig. 23.12.

a **i** State Faraday's law of electromagnetic induction. *[2]*

ii Use Faraday's law to explain why

- there is a reading on the voltmeter, *[1]*
- this reading varies in magnitude, *[1]*
- the reading has both positive and negative values. *[1]*

b Use Fig. 23.12 to determine the frequency f_0 of the oscillations of the magnet. *[2]*

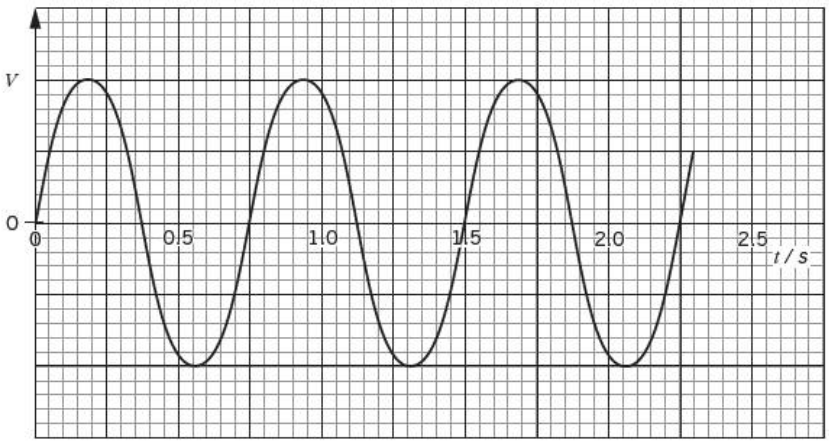


Fig. 23.12

c The magnet is now brought to rest and the voltmeter is replaced by a variable frequency alternating current supply that produces a constant r.m.s. current in the coil.

The frequency of the supply is gradually increased from $0.7 f_0$ to $1.3 f_0$, where f_0 is the frequency calculated in **b**.

On the axes of a copy of Fig. 23.13, sketch a graph to show the variation with frequency f of the amplitude A of the new oscillations of the bar magnet.

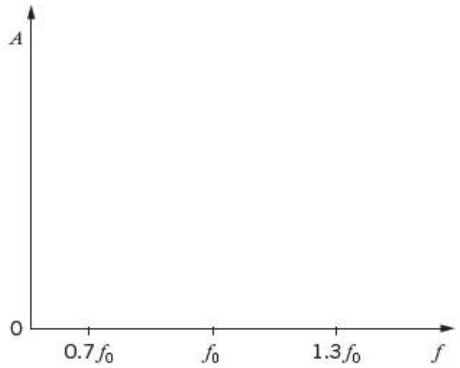


Fig. 23.13

[2]

Cambridge International AS and A Level Physics, 9702/42 May/June 2011 Q 5

www.iranino.co

IRANINOVIN

6 **a** Define the *tesla*. [2]

b A long solenoid has an area of cross-section of 28 cm^2, as shown in Fig. 23.14.

A coil C consisting of 160 turns of insulated wire is wound tightly around the centre of the solenoid.

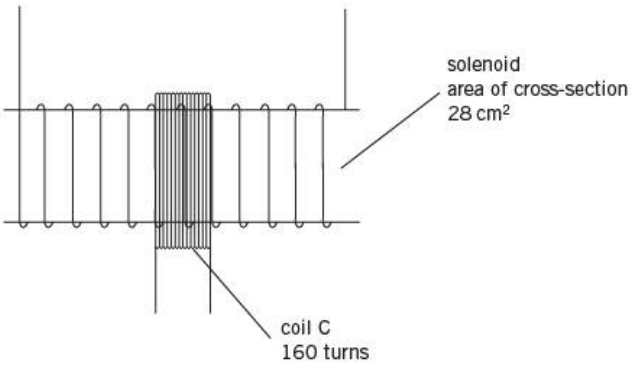


Fig. 23.14

The magnetic flux density B at the centre of the solenoid is given by the expression

$$B = \mu_0 nI$$

where I is the current in the solenoid, n is a constant equal to $1.5 \times 10^3 \text{m}^{-1}$ and μ_0 is the permeability of free space.

Calculate, for a current of 3.5 A in the solenoid,

i the magnetic flux density at the centre of the solenoid, [2]

ii the flux linkage in the coil C. [2]

c **i** State Faraday's law of electromagnetic induction. [2]

ii The current in the solenoid in **b** is reversed in direction in a time of 0.80 s. Calculate the average e.m.f. induced in coil C. [2]

Cambridge International AS and A Level Physics, 9702/43 May/June 2013 Q 5

www.iranino.co

IRANINOVIN

24 Alternating currents

By the end of this topic, you will be able to:

24.1 (a) understand and use the terms period, frequency, peak value and root-mean-square value as applied to an alternating current or voltage
(b) deduce that the mean power in a resistive load is half the maximum power for a sinusoidal alternating current
(c) represent a sinusoidally alternating current or voltage by an equation of the form $x = x_0 \sin \omega t$
(d) distinguish between r.m.s. and peak values and recall and solve problems using the relationship $I_{rms} = \frac{I_0}{\sqrt{2}}$ for the sinusoidal case

24.2 (a) understand the principle of a simple laminated iron-cored transformer and recall and solve problems using $\frac{N_s}{N_p} = \frac{V_s}{V_p} = \frac{I_p}{I_s}$ for an ideal transformer
(b) understand sources of energy loss in a practical transformer

24.3 (a) appreciate the scientific and economic advantages of alternating current and of high voltages for the transmission of electrical energy

24.4 (a) distinguish graphically between half-wave and full-wave rectification
(b) explain the use of a single diode for the half-wave rectification of an alternating current
(c) explain the use of four diodes (bridge rectifier) for the full-wave rectification of an alternating current
(d) analyse the effect of a single capacitor in smoothing, including the effect of the value of capacitance in relation to the load resistance

Starting points

- Power is dissipated in a resistor and the magnitude of the power is given by the expression I^2R or VI or V^2/R.
- Alternating voltages may be changed using a transformer.
- A diode is a device that allows current to move in one direction only.

24.1 Characteristics of alternating currents

Up to this point in our studies, we have dealt with systems in which a battery is connected to a circuit and there is a steady current in one direction. You will know this sort of current as a direct current, abbreviated to d.c. However, the domestic electricity supply, produced by generators at a power station, is one which uses **alternating current (a.c.)**. An alternating current or voltage reverses its direction regularly and is usually sinusoidal, as shown in Figure 24.1.

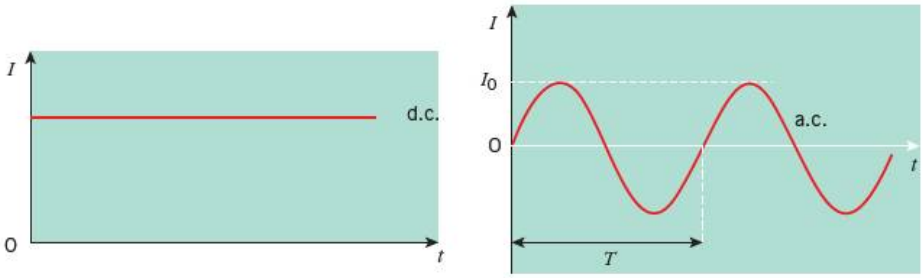


Figure 24.1 Direct and alternating currents

www.iranino.co

IRANINOVIN

We can represent the current and voltage by the equations

$$I = I_0 \sin \omega t$$
$$V = V_0 \sin \omega t$$

The time T taken for one complete cycle of the alternating current is the period of the current $T = 2\pi/\omega$. The reciprocal of the period is the frequency f. That is, $f = 1/T$ and $f = \omega/2\pi$. The frequency is the number of complete cycles per unit time. The unit of frequency is the hertz (Hz) where 1 Hz = 1 cycle per second. The **peak value** of the current or voltage is I_0 or V_0, the amplitude of the oscillating current or voltage. Sometimes the term peak-to-peak value is used; this means $2I_0$ or $2V_0$, or twice the amplitude.

It is clear from Figure 24.1 that the average value of an alternating current is zero. However, this does not mean that when an a.c. source is connected to a resistor, no power is generated in the resistor. An alternating current in a wire can be thought of as electrons moving backwards and forwards, and passing on their energy by collision. The power generated in a resistance R is given by the usual formula

$$P = I^2R$$

but here the current I must be written as

$$I = I_0 \sin \omega t$$

Thus

$$P = I_0^2 R \sin^2 \omega t$$

Since I_0^2 and $\sin^2 \omega t$ are always positive, we see that the power P is also always positive.

The expression we have just derived for P gives the power at any instant. What is much more useful is the average or mean power. This is the quantity which must be used in assessing the power generated in the resistor. Because I_0 and R are constants, the average value of P will depend on the average value of $\sin^2 \omega t$, which is ½. So the average power $\langle P \rangle$ delivered to the resistor is

$$\langle P \rangle = \tfrac{1}{2} I_0^2 R = \tfrac{1}{2} V_0^2 / R$$

This is half the maximum power. We could use the average value of the square of the current or the voltage in these relations, since

$$\langle I^2 \rangle = \tfrac{1}{2} I_0^2 \text{ and } \langle V^2 \rangle = \tfrac{1}{2} V_0^2$$

The square root of $\langle I^2 \rangle$ is called the **root-mean-square**, or **r.m.s.**, value of the current, and similarly for the voltage. The numerical relations are

$$I_{rms} = \sqrt{\langle I^2 \rangle} = I_0/\sqrt{2} = 0.707 I_0$$
$$V_{rms} = \sqrt{\langle V^2 \rangle} = V_0/\sqrt{2} = 0.707 V_0$$

The r.m.s. values are useful because they represent the *effective* values of current and voltage in an a.c. circuit. A direct current with a value of I equal to the r.m.s current I_{rms} of an a.c. circuit will produce exactly the same heating effect in a resistor. In specifying a domestic supply voltage, it is the r.m.s. value that is quoted, not the peak value.

The r.m.s. value of the alternating current or voltage is that value of the direct current or voltage that would produce thermal energy at the same rate in a resistor.

Example

A 1.5 kW heater is connected to the domestic supply, which is quoted as 240 V. Calculate the peak current in the heater, and its resistance.

The r.m.s version of the power/current/voltage equation is $I_{rms} V_{rms}$ = mean power.

This gives $I_{rms} = 1.5 \times 10^3/240 = 6.3$ A.

www.iranino.co

IRANINOVIN

The peak current $I_0 = \sqrt{2}\, I_{rms} = \mathbf{8.8\,A}$

The resistance $R = V_{rms}/I_{rms} = 240/6.3 = \mathbf{38\,\Omega}$

Now it's your turn

1 A heater of resistance 40 Ω is connected to a domestic supply, quoted as 240 V.
   (a) Calculate the average power in the resistor.
   (b) What are the maximum and minimum values of the instantaneous power in the resistor?

24.2 Transformers

In Topic 23, we saw that an e.m.f. can be induced in one coil when the current changes in another coil. If this current is alternating, then the e.m.f. induced is alternating. This is the principle of the **transformer**.

A simple transformer is illustrated in Figure 24.2.

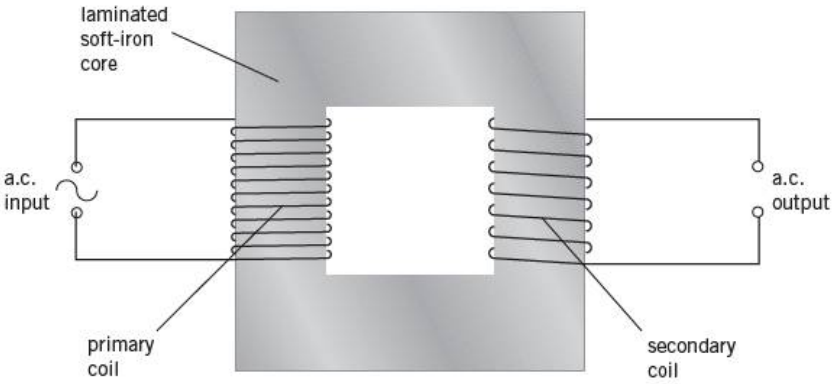


Figure 24.2 Simple transformer

Two coils of insulated wire are wound on to a laminated soft-iron core. An alternating e.m.f. is applied across one coil (the primary coil) and an e.m.f. is induced in the secondary coil. The ratio of the output e.m.f. to the applied e.m.f. is dependent on the ratio of the number of turns on the two coils. The applied e.m.f. and the induced e.m.f. across the secondary coil have the same frequency but are not in phase.

The alternating current in the primary coil gives rise to an alternating magnetic flux in the core that is in phase with the current in the primary coil. This magnetic flux in the core threads through the secondary coil. Since the flux is changing, by Faraday's law of electromagnetic induction, an e.m.f. will be induced in the secondary coil. The induced e.m.f. is proportional to the rate of change of magnetic flux linkage and, thus, this induced e.m.f. will be alternating at the same frequency as the current in the primary coil. There will be a phase angle between the two e.m.f.s.

Consider a transformer with N_p turns and N_s turns on the primary and secondary coils respectively, with currents in the coils of I_p and I_s when the e.m.f.s across the coils are V_p and V_s.

For a transformer which is 100% efficient (an **ideal** transformer)

$$\textit{input power} = \textit{output power}$$

$$V_p I_p = V_s I_s$$

and

$$\frac{N_s}{N_p} = \frac{V_s}{V_p} = \frac{I_p}{I_s}$$

When V_s is greater than V_p, there are more turns on the secondary coil than on the primary coil and the transformer is said to be a **step-up transformer**.

www.iranino.co

IRANINOVIN



Figure 24.3 Transformers are produced in many shapes and sizes.

Conversely, if V_s is less than V_p, the secondary coil has fewer turns than the primary and the transformer is said to be a **step-down transformer**.

In practice, the transformer will not be 100% efficient due to power losses. Some sources of these losses are:

- loss of magnetic flux between the primary and secondary coils. The core is designed for maximum flux linkage. The iron core forms a closed loop and the primary and secondary coils are wound on each other.
- resistive heating in the primary and secondary coils. The coils are made of a low resistivity metal.
- heating of the core due to eddy currents.
- heating of the core due to repeated magnetisation and demagnetisation.

The use of soft iron reduces heating due to repeated magnetisation and demagnetisation of the core. However, since iron (an electrical conductor) is used as the material for the core so that the magnetic flux linkage is large, eddy currents cannot be prevented. Laminating the core, i.e. building up the core from thin strips of soft iron which are electrically insulated from one another, reduces energy losses as thermal energy due to eddy currents.

Note: laminating the core cannot prevent eddy currents from forming but it can reduce the energy losses by reducing the sizes of these currents.

Example

A transformer is to be used with an alternating supply of 240 V to power a lamp rated as 12 V, 3 A. The secondary coil of the transformer has 80 turns. Assuming that the transformer is ideal, calculate, for the primary coil,

(a) the number of turns,
(b) the current.

(a) For an ideal transformer,

$V_s/V_p = N_s/N_p$

$12/240 = 80/N_p$

$\mathbf{N_p = 1600 \text{ turns}}$

(b) For an ideal transformer,

$V_s/V_p = I_p/I_s$

$12/240 = I_p/3$

$\mathbf{I_p = 0.15\,A}$

Now it's your turn

2 The primary coil of an ideal transformer has 1200 turns and is connected to a 240 V alternating supply. The transformer is to be used to step down the voltage to 9.0 V.
   (a) Calculate the number of turns on the secondary coil.
   (b) An appliance, rated as 9.0 V, 1.5 A is connected to the secondary coil of the transformer.
   Calculate the current in the primary coil.

3 An ideal transformer is to be used to step up a voltage from 220 V to supply a current of 15 mA at 6.6 kV. The primary coil has 250 turns.
   Calculate:
   (a) the ratio of the number of turns on the secondary coil to that on the primary coil,
   (b) the number of turns on the secondary coil,
   (c) the current in the primary coil.

24.3 Transmission of electrical energy

Transformers play an essential part in the transmission of electrical energy. Generating stations are often situated a long way from the cities they serve. Hydroelectric plants must be placed at the dam providing the water supply, and nuclear generators need a plentiful supply of cooling water and there is also a safety factor to consider. Fossil-fuel plants are also placed some distance from towns, mainly because of air pollution concerns and because land is cheaper away from towns. This means that, whatever the means of generating the electrical energy, it is likely to need to be transmitted over long distances.

www.iranino.co

IRANINOVIN

There is always a power loss in transmission lines, simply due to loss of energy due to heating of the cables (the I^2R effect, see page 150). However, it turns out that this loss can be reduced if the power is transmitted at a high voltage. Look at the following example.

Example

A small town, several kilometres from an electricity generating station, requires 120 kW of power on average. The total resistance of the transmission lines is 0.40 Ω. Calculate the power loss if the transmission is made at a voltage of **(a)** 240 V, **(b)** 24 kV.

First calculate the current I in the transmission lines in each case.

In **(a)**, $I = P/V = 1.2 \times 10^5/240 = 500$ A.

In **(b)**, the same calculation gives $I = 5.0$ A.

Now find the power loss in the transmission lines from $P = I^2R$.

In **(a)**, it is $(500)^2 \times 0.40 =$ **100 kW**, or more than 80% of the power required by the town.

In **(b)**, the power loss is $5.0^2 \times 0.40 =$ **10 W**, less than one-hundredth of 1% of the power requirement.

The advantage of transmitting power at the higher voltage is now obvious. There is much less power loss in the transmission lines.

The advantage of using alternating current rather than direct current is that an alternating supply is so efficiently stepped up to a high voltage for transmission using a transformer. Transformers used in transmission networks have very high efficiencies, often over 98%.

Before transmission from the generating station, the alternating voltage is stepped up to, in many instances, in excess of 200 kV. After transmission, the voltage is stepped down. The final voltage will depend on the needs of the customer.

24.4 Rectification

It is sometimes necessary to convert an alternating current into a direct current. This is because most electronic devices require direct current, whereas the domestic supply is alternating. The conversion can be done by a process known as **rectification**.

Suppose a single diode (see AS Level Topic 19) is connected into the a.c. circuit of Figure 24.4.

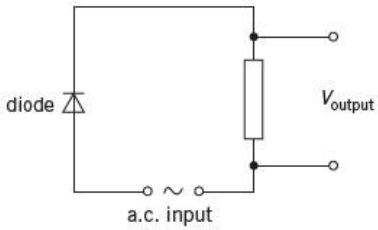


Figure 24.4 Single-diode circuit for half-wave rectification

We know that the diode allows current to flow in one direction only. This means that the output voltage across the resistor will consist only of the positive half-cycles of the input voltage, as shown in Figure 24.5. The diode has rejected the negative part of the input, producing a unidirectional voltage which fluctuates considerably, rather than a constant direct voltage. Nevertheless, we have achieved what is called **half-wave rectification**.

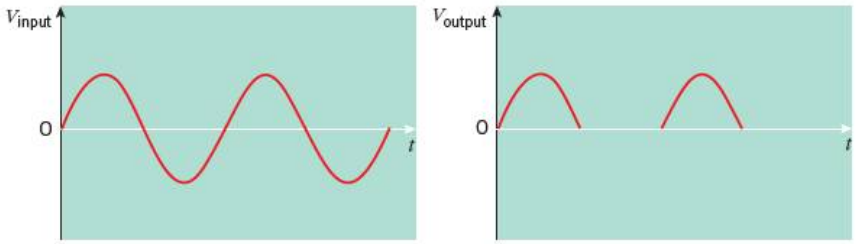


Figure 24.5 Half-wave rectification

It is more satisfactory to make use of the negative half-cycles of the input and reverse their polarity, as shown in Figure 24.6. This process is called **full-wave rectification**.

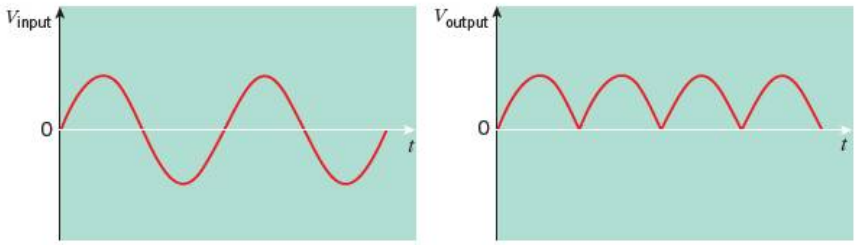


Figure 24.6 Full-wave rectification

www.iranino.co

IRANINOVIN

One circuit used for full-wave rectification is illustrated in Figure 24.7.

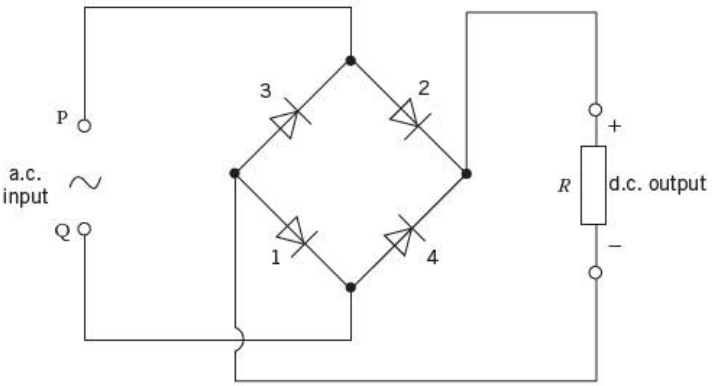


Figure 24.7 Four-diode (bridge) circuit for full-wave rectification

It uses four diodes arranged in a diamond pattern and is referred to as a **bridge rectifier** circuit. The input terminals are P and Q. If P is positive during the first half-cycle, diodes 1 and 2 on opposite sides of the diamond will conduct. In the next half-cycle Q is positive, and diodes 3 and 4 conduct. Thus the resistor acting as a load will always have its upper terminal positive and its lower terminal negative.

The circuit has produced a unidirectional voltage, but the output is still not a good approximation to the steady direct voltage, as required by most electronic equipment. We can improve the situation by inserting a capacitor across the output terminals of the bridge circuit as in Figure 24.8.

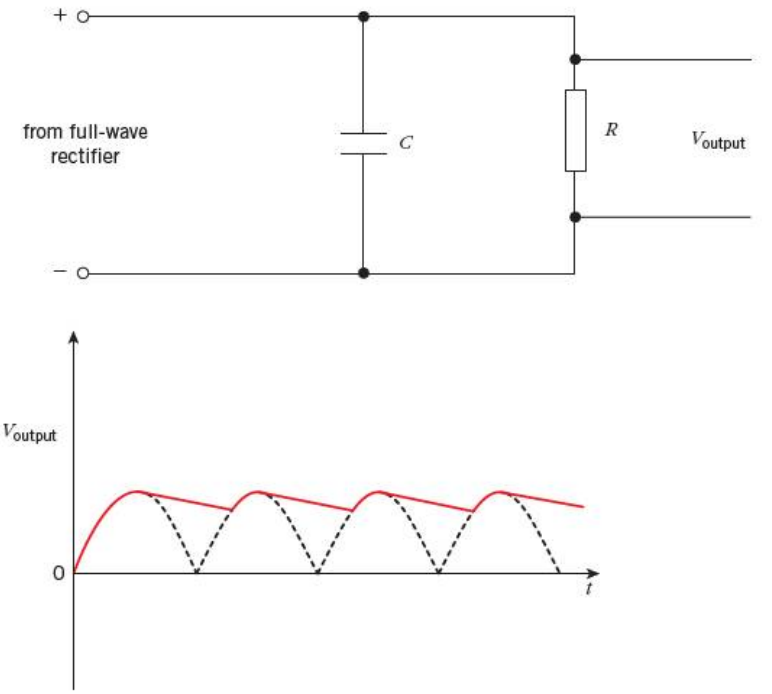


Figure 24.8 Smoothing by capacitor

The capacitor charges up on the rising part of the half-cycle, and then discharges through the resistor as the output voltage falls. The effect is to reduce the fluctuations in the unidirectional output. This process is called **smoothing**.

The charge and discharge of a capacitor was discussed in Topic 18. The important factor is the time constant of the resistor–capacitor circuit. If the product of the capacitance C and the load resistance R is much larger than the half-period of the

www.iranino.co

IRANINOVIN

original supply to the rectifier circuit, the ripple on the direct voltage or current will be small. Reducing the time constant will increase the ripple, as illustrated in Figure 24.9.

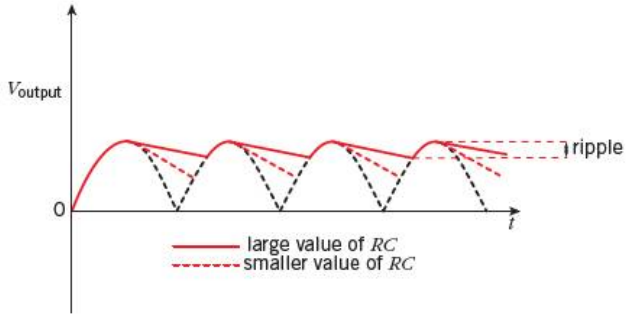


Figure 24.9 Magnitude of the ripple

Summary

- Alternating current or voltage is represented by an equation of the form $x = x_0 \sin \omega t$.
- Peak and root-mean-square (r.m.s.) values of sinusoidal current or voltage are related by an equation of the form $x_0 = \sqrt{2}\, x_{rms}$.
- For a sinusoidal input, mean power in a resistive load is one half of the peak power.
- For an ideal transformer, $\frac{N_s}{N_p} = \frac{V_s}{V_p} = \frac{I_p}{I_s}$.
- Electrical power is transmitted at high voltages to reduce thermal energy losses in cables.
- A single diode gives half-wave rectification: negative half-cycles are blocked.
- A bridge circuit of four diodes can give full-wave rectification.
- A capacitor across the output reduces the fluctuations of the rectified output voltage.

Examination style questions

1 A stereo system has two output channels, each connected to a loudspeaker of effective resistance 8.0Ω. Each channel can deliver a maximum average power output of 50W to its speaker. Calculate the r.m.s. voltage and the r.m.s. current fed to one speaker at this maximum power. Assume that the loudspeaker can be treated as a simple resistance.

2 The half-wave rectifier circuit of Fig. 24.4 is used to rectify an a.c. input voltage of 240V r.m.s. The output resistor has resistance 25kΩ.

   a Calculate the peak value of the input voltage.

   b Estimate the average current in the output resistor.

3 From power stations, electrical energy is transmitted over long distances to homes and business, at 400kV. An alternative might be to transmit at the domestic supply voltage of 230V. Explain how the high-voltage option reduces power losses in the cables.

4 An ideal transformer is to be used to step down a 220V alternating supply to 7.5V.

   a Calculate the ratio of the number of turns on the secondary coil to that on the primary coil.

   b Calculate the number of turns on the primary coil given that there are 140 turns on the secondary coil.

   c Suggest why the wire forming the secondary coil would, in this case, be thicker than that for the primary coil.

www.iranino.co

IRANINOVIN

5 **a** State Faraday's law of electromagnetic induction. *[2]*

b The output of an ideal transformer is connected to a bridge rectifier, as shown in Fig. 24.10.

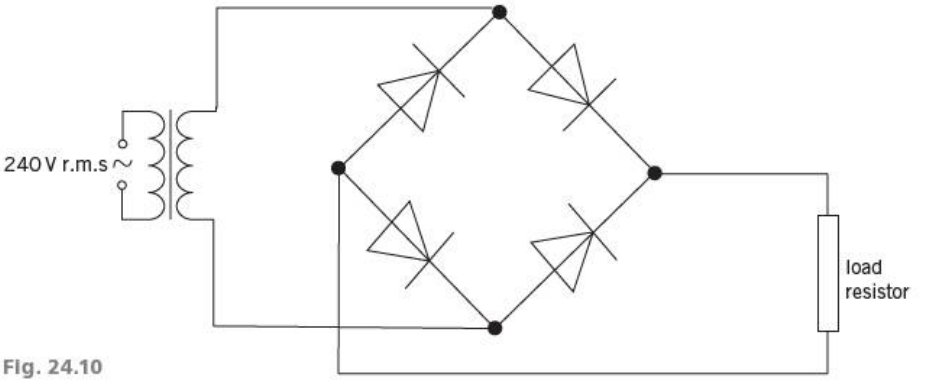


Fig. 24.10

The input to the transformer is 240 V r.m.s. and the **maximum** potential difference across the load resistor is 9.0 V.

i On a copy of Fig. 24.10, mark with the letter P the positive output from the rectifier. *[1]*

ii Calculate the ratio

$$\frac{\text{number of turns on primary coil}}{\text{number of turns on secondary coil}} \quad [3]$$

c The variation with time t of the potential difference V across the load resistor in **b** is shown in Fig. 24.11.

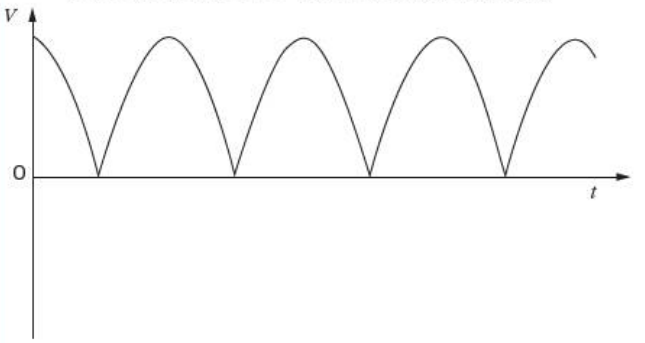


Fig. 24.11

A capacitor is now connected in parallel with the load resistor to produce some smoothing.

i Explain what is meant by *smoothing*. *[1]*

ii On a copy of Fig. 24.11, draw the variation with time t of the smoothed output potential difference. *[2]*

Cambridge International AS and A Level Physics, 9702/42 May/June 2013 Q 6

6 An alternating current supply is connected in series with a resistor R, as shown in Fig. 24.12.

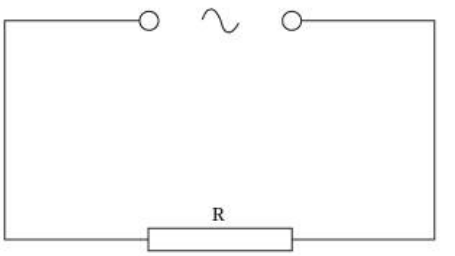


Fig. 24.12

The variation with time t (measured in seconds) of the current I (measured in amperes) in the resistor is given by the expression

$I = 9.9\sin(380t)$.

a For the current in the resistor R, determine

i the frequency, *[2]*

ii the r.m.s. current. *[2]*

b To prevent over-heating, the mean power dissipated in resistor R must not exceed 400 W.

Calculate the minimum resistance of R. *[2]*

Cambridge International AS and A Level Physics, 9702/42 May/June 2011 Q 6

www.iranino.co

IRANINOVIN

7 A simple iron-cored transformer is illustrated in Fig. 24.13.

a i State why the primary and secondary coils are wound on a core made of iron. *[1]*

ii Suggest why thermal energy is generated in the core when the transformer is in use. *[3]*

b The root-mean-square (r.m.s.) voltage and current in the primary coil are V_P and I_P respectively.

The r.m.s. voltage and current in the secondary coil are V_S and I_S respectively.

i Explain, by reference to direct current, what is meant by the *root-mean-square* value of an alternating current. *[2]*

ii Show that, for an ideal transformer, $\frac{V_s}{V_p} = \frac{I_p}{I_s}$. *[2]*

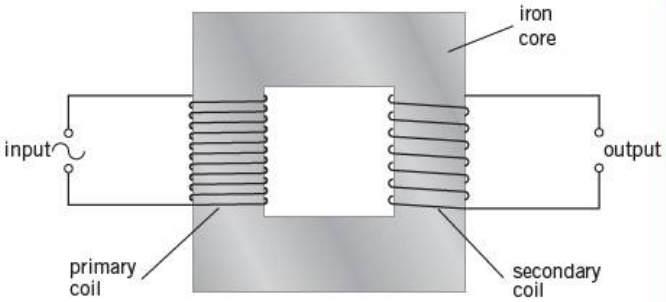


Fig. 24.13

Cambridge International AS and A Level Physics, 9702/42 Oct/Nov 2010 Q 6

www.iranino.co

IRANINOVIN

25 Quantum physics

By the end of this topic, you will be able to:

25.1 (a) appreciate the particulate nature of electromagnetic radiation
(b) recall and use $E = hf$

25.2 (a) understand that the photoelectric effect provides evidence for a particulate nature of electromagnetic radiation while phenomena such as interference and diffraction provide evidence for a wave nature
(b) recall the significance of threshold frequency
(c) explain photoelectric phenomena in terms of photon energy and work function energy
(d) explain why the maximum photoelectric energy is independent of intensity, whereas the photoelectric current is proportional to intensity
(e) recall, use and explain the significance of $hf = \Phi + \frac{1}{2}m_e{v_{max}}^2$

25.3 (a) describe and interpret qualitatively the evidence provided by electron diffraction for the wave nature of particles
(b) recall and use the relation for the de Broglie wavelength $\lambda = h/p$

25.4 (a) show an understanding of the existence of discrete energy levels in isolated atoms (e.g. atomic hydrogen) and deduce how this leads to spectral lines
(b) distinguish between emission and absorption line spectra
(c) recall and solve problems using the relation $hf = E_1 - E_2$

25.5 (a) appreciate that, in a simple model of band theory, there are energy bands in solids
(b) understand the terms valence band, conduction band and forbidden band (band gap)
(c) use simple band theory to explain the temperature dependence of the resistance of metals and of intrinsic semiconductors
(d) use simple band theory to explain the dependence on light intensity of the resistance of an LDR

25.6 (a) explain the principles of the production of X-rays by electron bombardment of a metal target
(b) describe the main features of a modern X-ray tube, including control of the intensity and hardness of the X-ray beam
(c) understand the use of X-rays in imaging internal body structures, including a simple analysis of the causes of sharpness and contrast in X-ray imaging
(d) recall and solve problems by using the equation $I = I_0 e^{-\mu x}$ for the attenuation of X-rays in matter;
(e) understand the purpose of computed tomography or CT scanning
(f) understand the principles of CT scanning
(g) understand how the image of an 8-voxel cube can be developed using CT scanning

Starting points

- A simple model of an atom involves a small massive positively charged nucleus around which orbit electrons.
- The electrons have different energies, dependent on their orbits.
- Electromagnetic waves may undergo interference and diffraction.

25.1 & 25.2 Photoelectric emission of electrons and energy of a photon

Some of the electrons in a metal are free to move around in it. (It is these free electrons that form the electric current when a potential difference is applied across the

www.iranino.co

ends of a metal wire.) However, to remove free electrons from a metal requires energy, because they are held in the metal by the electrostatic attraction of the positively charged nuclei. If an electron is to escape from the surface of a metal, work must be done on it. The electron must be given energy. When this energy is in the form of light energy, the phenomenon is called **photoelectric emission**.

Photoelectric emission is the release of electrons from the surface of a metal when electromagnetic radiation is incident on its surface.

Demonstration of photoelectric emission

A clean zinc plate is placed on the cap of a gold-leaf electroscope. The electroscope is then charged negatively, and the gold leaf deflects, proving that the zinc plate is charged. This is illustrated in Figure 25.1.

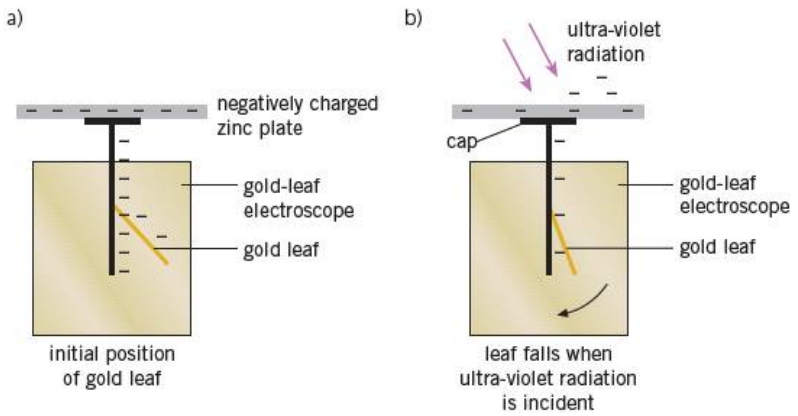


Figure 25.1 Demonstration of photoelectric emission

If visible light of any colour is shone on to the plate, the leaf does not move. Even when the intensity (the brightness) of the light is increased, the leaf remains in its deflected position. However, when ultra-violet radiation is shone on the plate, the leaf falls immediately, showing that it is losing negative charge. This means that electrons are being emitted from the zinc plate. These electrons are called **photoelectrons**. If the intensity of the ultra-violet radiation is increased, the leaf falls more quickly, showing that the rate of emission of electrons has increased.

The difference between ultra-violet radiation and visible light is that ultra-violet radiation has a shorter wavelength and a higher frequency than visible light.

Further investigations with apparatus like this lead to the following conclusions:

- If photoemission takes place, it does so instantaneously. There is no delay between illumination and emission.
- Photoemission takes place only if the frequency of the incident radiation is above a certain minimum value called the **threshold frequency** f_0.
- Different metals have different threshold frequencies.
- Whether or not emission takes place depends only on whether the frequency of the radiation used is above the threshold for that surface. It does not depend on the intensity of the radiation.
- For a given frequency, the rate of emission of photoelectrons is proportional to the intensity of the radiation.

Another experiment, using the apparatus shown in Figure 25.2, can be carried out to investigate the energies of the photoelectrons.

If ultra-violet radiation of a fixed frequency (above the threshold) is shone on to the metal surface A, it emits photoelectrons. Some of these electrons travel from A to B. Current is detected using the microammeter. If a potential difference is applied between A and B, with B negative with respect to A, any electron going from A to B

www.iranino.co

IRANINOVIN

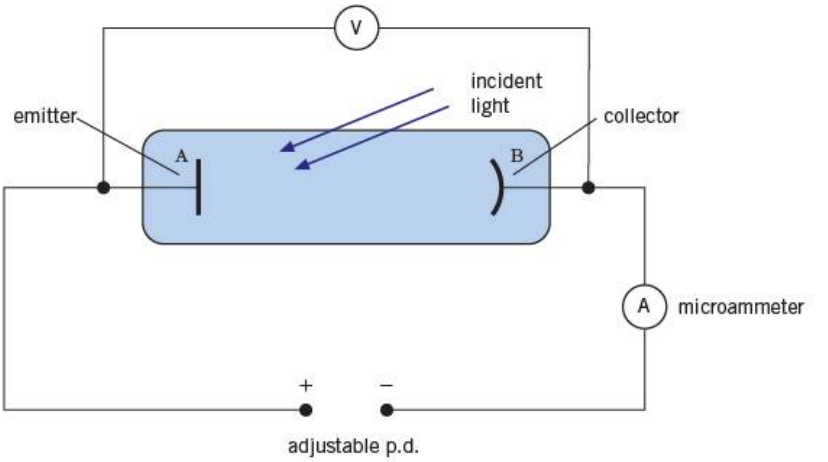


Figure 25.2 Experiment to measure the maximum kinetic energy of photoelectrons

will gain potential energy as it moves against the electric field. The gain in potential energy is at the expense of the kinetic energy of the electron. That is,

loss in kinetic energy = gain in potential energy

= charge of electron × potential difference

If the voltage between A and B is gradually increased, the current registered on the microammeter decreases and eventually falls to zero. The minimum value of the potential difference necessary to stop the electron flow is known as the **stopping potential**. It measures the maximum kinetic energy with which the photoelectrons are emitted. The fact that there is a current in the microammeter at voltages less than the stopping potential indicates that there is range of kinetic energies for these electrons.

If the experiment is repeated with radiation of greater intensity but the same frequency, the maximum current in the microammeter increases, but the value of the stopping potential is unchanged.

The experiment can be repeated using ultra-violet radiation of different frequencies, measuring the stopping potential for each frequency. When the maximum kinetic energy of the photoelectrons is plotted against the frequency of the radiation, the graph of Figure 25.3 is obtained.

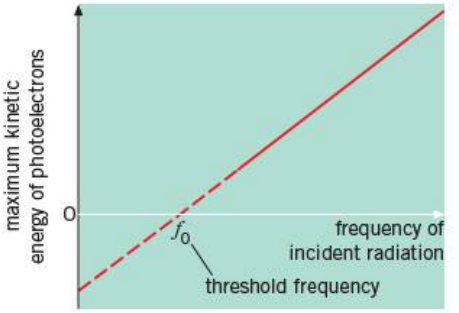


Figure 25.3 Graph of maximum kinetic energy of photoelectrons against frequency of radiation

The following conclusions are drawn from this experiment:

- The photoelectrons have a range of kinetic energies, from zero up to some maximum value. If the frequency of the incident radiation is increased, the maximum kinetic energy of the photoelectrons also increases.
- For constant frequency of the incident radiation, the maximum kinetic energy is unaffected by the intensity of the radiation.
- When the graph of Figure 25.3 is extrapolated to the point where the maximum kinetic energy of the photoelectrons is zero, the minimum frequency required to cause emission from the surface (the threshold frequency) may be found.

www.iranino.co

At the time when the photoelectric effect was first being studied, it was fully accepted that light is a wave motion. Evidence for this came from observed interference and diffraction effects. The conclusions of experiments on photoemission produced doubt as to whether light is a continuous wave. One of the main problems concerns the existence of a threshold frequency.

Classical wave theory predicts that when an electromagnetic wave (that is, light) interacts with an electron, the electron will absorb energy from it. So, if an electron absorbs enough energy, it should be able to escape from the metal. Remember from AS Level Topic 14 that the energy carried by a wave depends on its amplitude and its frequency. Thus, even if we have a low-frequency wave, its energy can be boosted by increasing the amplitude (that is, by increasing the brightness of the light). So, according to wave theory, we ought to be able to cause photoemission using any frequency of light, provided we make it bright enough. Alternatively, we could use less bright light and shine it on the metal for a longer time, until enough energy to cause emission has been delivered. But this does not happen. The experiments we have described above showed conclusively that radiation of frequency below the threshold, no matter how intense or for how long it is used, does not produce photoelectrons. The classical wave theory of electromagnetic radiation leads to the following predictions:

1 Whether an electron is emitted or not should depend on the power of the incident wave; that is, on its intensity. A very intense wave, of any frequency, should cause photoemission.
2 The maximum kinetic energy of the photoelectrons should be greater if the radiation intensity is greater.
3 There is no reason why photoemission should be instantaneous.

These predictions, based on wave theory, do not match the observations. A new approach, based on an entirely new concept, the **quantum theory**, was used to explain these findings.

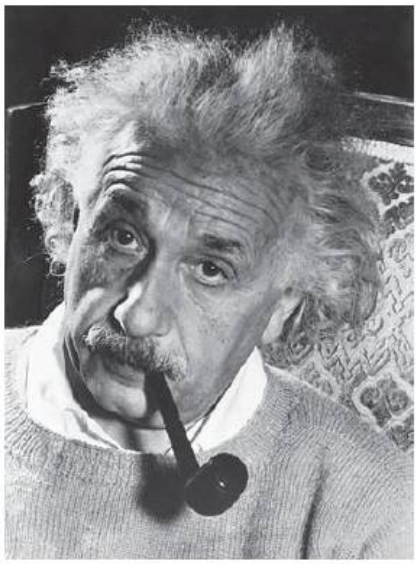


Figure 25.4 Albert Einstein

Einstein's theory of photoelectric emission

In 1901, the German physicist Max Planck had suggested that the energy carried by electromagnetic radiation might exist as discrete packets called **quanta**. The energy E carried in each quantum is given by

$$E = hf$$

where f is the frequency of the radiation and h is a constant called the Planck constant. The value of the Planck constant is 6.63×10^{-34} J s.

In 1905, Albert Einstein developed the theory of quantised energy to explain all the observations associated with photoelectric emission. He proposed that light radiation consists of a stream of energy packets called **photons**.

A photon is the special name given to a quantum of energy when the energy is in the form of electromagnetic radiation.

When a photon interacts with an electron, it transfers all its energy to the electron. It is only possible for a single photon to interact with a single electron; the photon cannot share its energy between several electrons. This transfer of energy is instantaneous.

The photon theory of photoelectric emission is as follows. If the frequency of the incident radiation is less than the threshold frequency for the metal, the energy carried by each photon is insufficient for an electron to escape the surface of the metal. If the photon energy is insufficient for an electron to escape, it is converted to thermal energy in the metal.

The minimum amount of energy necessary for an electron to escape from the surface is called the **work function energy** φ.

Some values for the work function energy φ and threshold frequency f_0 of different metals are given in Table 25.1.

www.iranino.co

IRANINOVIN

Table 25.1 Work function energies and threshold frequencies

metal	Φ/J	Φ/eV	f_0/Hz
sodium	3.8×10^{-19}	2.4	5.8×10^{14}
calcium	4.6×10^{-19}	2.9	7.0×10^{14}
zinc	5.8×10^{-19}	3.6	8.8×10^{14}
silver	6.8×10^{-19}	4.3	1.0×10^{15}
platinum	9.0×10^{-19}	5.6	1.4×10^{15}

Remember: $1\,\text{eV} = 1.6 \times 10^{-19}\,\text{J}$.

If the frequency of the incident radiation is equal to the threshold frequency, the energy carried by each photon is just sufficient for electrons at the surface to escape. If the frequency of the incident radiation is greater than the threshold frequency, surface electrons will escape and have surplus energy in the form of kinetic energy. These electrons will have the maximum kinetic energy. If a photon interacts with an electron below the surface, some energy is used to take the electron to the surface, so that it is emitted with less than the maximum kinetic energy. This gives rise to a range of values of kinetic energy.

Einstein used the principle of conservation of energy to derive the photoelectric equation

photon energy = work function energy + maximum kinetic energy of photoelectron

or

$$hf = \Phi + \tfrac{1}{2}m_e v_{max}^2$$

For radiation incident at the threshold frequency, $\frac{1}{2}m_e v_{max}^2 = 0$ so that $hf_0 = \Phi$. The photoelectric equation can then be written

$$hf = hf_0 + \tfrac{1}{2}m_e v_{max}^2$$

Note

Experimental evidence indicates that the photoelectric current, i.e. the rate of emission of photoelectrons, depends on the intensity of the radiation when the frequency is constant. Increasing intensity gives rise to increasing rate of emission of photoelectrons.

This observation has led many students to believe that rate of emission is independent of frequency. This is incorrect. A beam of radiation has an intensity. This intensity is numerically equal to the power incident normally on unit area of the surface. Since the beam consists of a stream of photons, the intensity is the product of the rate of arrival of photons and the energy of each photon. At constant intensity, the rate at which photons arrive at the metal surface depends on the energy of each photon. So, if the frequency of the radiation increases, the energy of each photon increases and, therefore, for constant intensity, the rate of arrival of photons decreases. Fewer photons per unit time means a smaller rate of emission of electrons.

Example

The work function energy of platinum is 9.0×10^{-19} J. Calculate:

(a) the threshold frequency for the emission of photoelectrons from platinum,

(b) the maximum kinetic energy of a photoelectron when radiation of frequency 2.0×10^{15} Hz is incident on a platinum surface.

(a) Using $hf_0 = \Phi$, $f_0 = \Phi/h$, so

$f_0 = 9.0 \times 10^{-19}/6.6 \times 10^{-34} = \mathbf{1.4 \times 10^{15}\,Hz}$

(b) Using $hf = hf_0 + \frac{1}{2}m_e v_{max}^2$, $hf - hf_0 = \frac{1}{2}m_e v_{max}^2$ and

$\frac{1}{2}m_e v_{max}^2 = 6.6 \times 10^{-34}\,(2.0 \times 10^{15} - 1.4 \times 10^{15}) = \mathbf{4.0 \times 10^{-19}\,J}$

www.iranino.co

IRANINOVIN

Now it's your turn

1 The work function energy of silver is 4.3 eV. Show that the threshold frequency is about 1.0×10^{15} Hz.

2 Electromagnetic radiation of frequency 3.0×10^{15} Hz is incident on the surface of sodium metal. The emitted photoelectrons have a maximum kinetic energy of 1.6×10^{-18} J. Calculate the threshold frequency for photoemission from sodium.

3 Data for the threshold frequency f_0 and the work function energy ϕ of some metal surfaces are shown in the table.

metal	f_0 / 10^{14} Hz	Φ / 10^{-19} J
platinum		9.0
sodium	5.8	3.8
zinc	8.8	5.8

(a) Calculate the threshold frequency for platinum.

(b) A beam of electromagnetic radiation having a continuous range of wavelengths between 320 nm and 550 nm is incident, in turn, on each of the above metals. Determine which metals will emit photoelectrons.

(c) When light of frequency f and intensity I is incident on a certain metal surface, electrons are emitted. State and explain the effect, if any, on the emission of photoelectrons for light of frequency $2f$ and intensity I.

25.3 Wave–particle duality

If light waves can behave like particles (photons), perhaps moving particles can behave like waves?

When a beam of X-rays of a single wavelength is directed at a thin metal foil, a diffraction pattern is produced as shown in Figure 25.5.

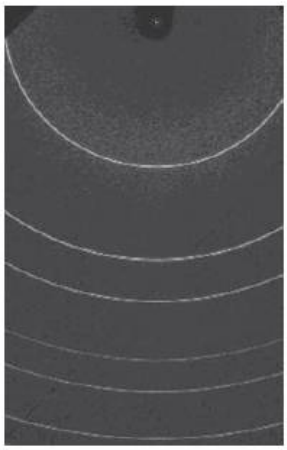


Figure 25.5 X-ray diffraction pattern of a metal foil

This is a similar effect to the diffraction pattern produced when light passes through a diffraction grating (see Topic 15). The foil contains many tiny crystals. The gaps between neighbouring planes of atoms in the crystals act as slits, creating a diffraction pattern.

If a beam of electrons is directed at a graphite film, a similar diffraction pattern is produced, as shown in Figure 25.6.

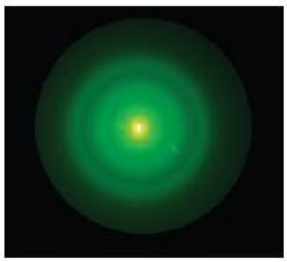


Figure 25.6 Electron diffraction pattern of graphite

The electrons, which we normally consider to be particles, are exhibiting a property we would normally associate with waves. Remember that, to observe diffraction, the wavelength of the radiation should be comparable with the size of the aperture. The separation of planes of atoms in crystals is of the order of 10^{-10} m. The fact that diffraction is observed with electrons suggest that they have a wavelength of about the same magnitude.

In 1924 the French physicist Louis de Broglie suggested that all moving particles have a wave-like nature. Using ideas based upon the quantum theory and Einstein's theory of relativity, he suggested that the momentum p of a particle and its associated wavelength λ are related by the equation

$$\lambda = \frac{h}{p}$$

where h is the Planck constant. λ is known as the **de Broglie wavelength**.

Example

Calculate the de Broglie wavelength of an electron travelling with a speed of 1.0×10^7 m s^{-1}. (Planck constant $h = 6.6 \times 10^{-34}$ J s; electron mass $m_e = 9.1 \times 10^{-31}$ kg)

Using $\lambda = h/p$ and $p = mv$,

$\lambda = 6.6 \times 10^{-34}/9.1 \times 10^{-31} \times 1.0 \times 10^7 =$ **7.3×10^{-11} m**

www.iranino.co

IRANINOVIN

Now it's your turn

4 Calculate the de Broglie wavelength of an electron travelling with a speed of 5.5×10^7 m s^{-1}.
(Planck constant $h = 6.6 \times 10^{-34}$ J s; electron mass $m_e = 9.1 \times 10^{-31}$ kg)

5 Calculate the de Broglie wavelength of an electron which has been accelerated from rest through a potential difference of 100 V. (Planck constant $h = 6.6 \times 10^{-34}$ J s; electron mass $m_e = 9.1 \times 10^{-31}$ kg; electron charge $e = -1.6 \times 10^{-19}$ C)

6 Calculate the speed of a neutron with de Broglie wavelength 1.5×10^{-10} m.
(Planck constant $h = 6.6 \times 10^{-34}$ J s; neutron mass $m_n = 1.7 \times 10^{-27}$ kg)

25.4 Energy levels in atoms and line spectra

When white light from a tungsten filament lamp is passed through a prism, the light is dispersed into its component colours, as illustrated in Figure 25.7.

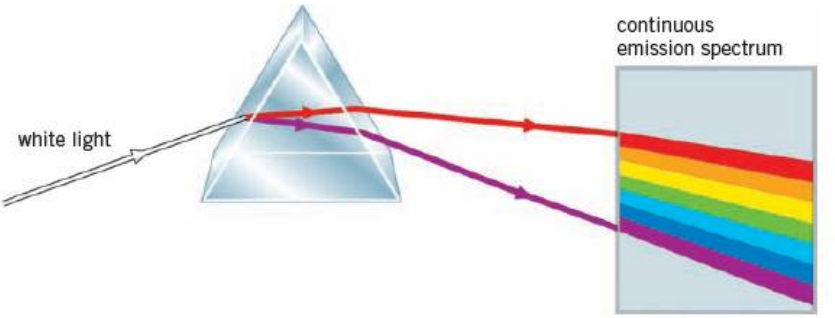


Figure 25.7 Continuous spectrum of white light from a tungsten filament lamp

The band of different colours is called a **continuous spectrum**. A continuous spectrum has all colours (and wavelengths) between two limits. In the case of white light, the colour and wavelength limits are violet (about 400 nm) and red (about 700 nm). Since this spectrum has been produced by the emission of light from the tungsten filament lamp, it is referred to as an **emission spectrum**. Finer detail of emission spectra than is obtained using a prism may be achieved using a diffraction grating.

A discharge tube is a transparent tube containing a gas at low pressure. When a high potential difference is applied across two electrodes in the tube, light is emitted. Examination of the light with a diffraction grating shows that the emitted spectrum is no longer continuous, but consists of a number of bright lines (Figure 25.8).

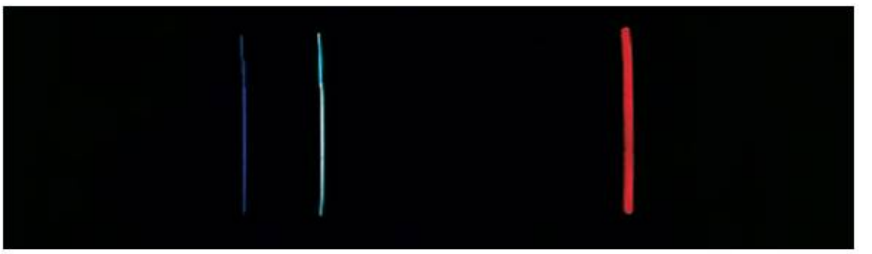


Figure 25.8 Line spectrum of hydrogen from a discharge tube

Such a spectrum is known as a **line spectrum**. It consists of a number of separate colours, each colour being seen as the image of the slit in front of the source. The wavelengths corresponding to the lines of the spectrum are characteristic of the gas which is in the discharge tube.

www.iranino.co

IRANINOVIN

Electron energy levels in atoms

To explain how line spectra are produced we need to understand how electrons in atoms behave. Electrons in an atom can have only certain specific energies. These energies are called the **electron energy levels** of the atom. The energy levels may be represented as a series of lines against a vertical scale of energy, as illustrated in Figure 25.9.

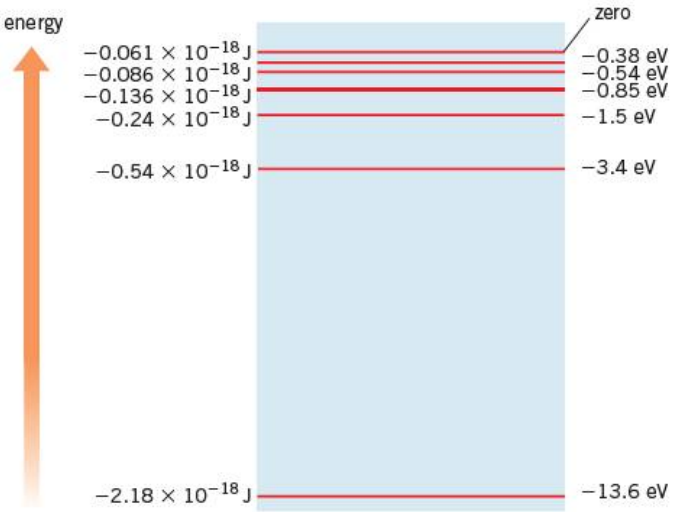


Figure 25.9 Electron energy levels for the hydrogen atom

The electron in the hydrogen atom can have any of these energy values, but cannot have energies between them.

Normally electrons occupy the lowest energy levels available. Under these conditions the atom and its electrons are said to be in the **ground state**. Figure 25.10a represents a hydrogen atom with its single electron in the lowest energy state.

If, however, the electron absorbs energy, perhaps by being heated, or by collision with another electron, it may be promoted to a higher energy level. The energy absorbed is exactly equal to the difference in energy of the two levels. Under these conditions the atom is described as being in an **excited state**. This is illustrated in Figure 25.10b.

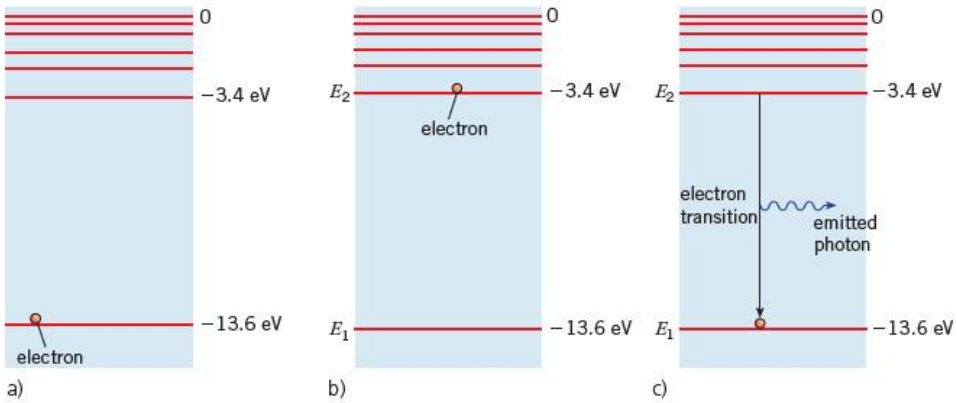


Figure 25.10 Electron in a hydrogen atom a) in its ground state, b) in an excited state, c) returning to its ground state with photon emission

An excited atom is unstable. After a short time, the excited electron will return to a lower level. To achieve this, the electron must lose energy. It does so by emitting a photon of electromagnetic radiation, as illustrated in Figure 25.10c.

www.iranino.co

IRANINOVIN

The energy hf of the photon is given by

$$hf = E_2 - E_1$$

where E_2 is the energy of the higher level and E_1 is that of the lower, and h is the Planck constant. Using the relation between the speed c of light, frequency f and wavelength λ, the wavelength of the emitted radiation is given by

$$\lambda = \frac{hc}{\Delta E}$$

where $\Delta E = E_2 - E_1$. This movement of an electron between energy levels is called an **electron transition**. Note that, the larger the energy of the transition, the higher the frequency (and the shorter the wavelength) of the emitted radiation.

Note that this downward transition results in the **emission** of a photon. The atom can be raised to an excited state by the **absorption** of a photon, but the photon must have just the right energy, corresponding to the difference in energy of the excited state and the initial state. So, a downward transition corresponds to photon emission, and an upward transition to photon absorption.

Figure 25.11 shows some of the possible transitions that might take place when electrons in an excited atom return to lower energy levels.

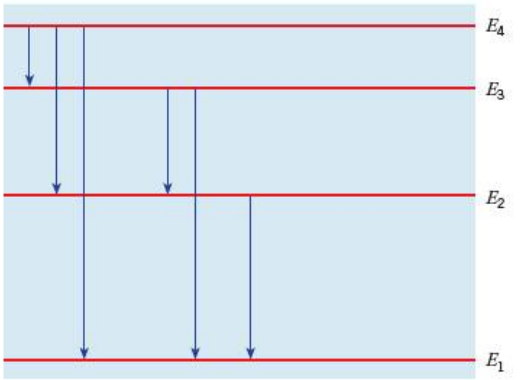


Figure 25.11 Some possible electron transitions

Each of the transitions results in the emission of a photon with a particular wavelength. For example, the transition from E_4 to E_1 results in light with the highest frequency and shortest wavelength. On the other hand, the transition from E_4 to E_3 gives the lowest frequency and longest wavelength.

Because all elements have different energy levels, the energy differences are unique to each element. Consequently, each element produces a different and characteristic line spectrum. Spectra can be used to identify the presence of a particular element. The line spectrum of mercury is shown in Figure 25.12.



Figure 25.12 Mercury line spectrum

The study of spectra is called **spectroscopy**, and instruments used to measure the wavelengths of spectra are **spectrometers**. Spectrometers for accurate measurement of wavelength make use of diffraction gratings (Topic 15) to disperse the light.

www.iranino.co

IRANINOVIN

Continuous spectra

Whilst the light emitted by isolated atoms such as those in low-pressure gases produces line spectra, the light emitted by atoms in a solid, a liquid, or a gas at high pressure produces a continuous spectrum. This happens because of the proximity of the atoms to each other. Interaction between the atoms results in a broadening of the electron energy levels. Consequently, transitions of a wide range of magnitudes of energy are possible, and light of a broad spread of wavelengths may be emitted. This is seen as a **continuous spectrum**.

Absorption spectra

When white light passes through a low-pressure gas and the spectrum of the white light is then analysed, it is found that light of certain wavelengths is missing. In their place are dark lines. This type of spectrum is called an **absorption spectrum**; one is shown in Figure 25.13.



Figure 25.13 An absorption spectrum

As the white light passes through the gas, some electrons absorb energy and make transitions to higher energy levels. The wavelengths of the light they absorb correspond exactly to the energies needed to make the particular upward transitions. When these excited electrons return to lower levels, the photons are emitted in all directions, rather than in the original direction of the white light. Thus, some wavelengths appear to be missing. It follows that the wavelengths missing from an absorption spectrum are those present in the emission spectrum of the same element. This is illustrated in Figure 25.14.

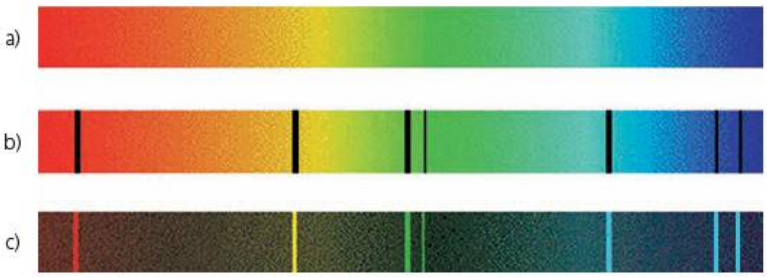


Figure 25.14 Relation between an absorption spectrum and the emission spectrum of the same element: a) spectrum of white light, b) absorption spectrum of element, c) emission spectrum of the same element

Example

Calculate the wavelength of the radiation emitted when the electron in a hydrogen atom makes a transition from the energy level at -0.54×10^{-18} J to the level at -2.18×10^{-18} J.

(Planck constant $h = 6.6 \times 10^{-34}$ J s; speed of light $c = 3.00 \times 10^{8}$ m s^{-1})

Here $\Delta E = E_2 - E_1 = -0.54 \times 10^{-18} - (-2.18 \times 10^{-18}) = 1.64 \times 10^{-18}$ J

Using $\lambda = hc/\Delta E$,

$\lambda = 6.6 \times 10^{-34} \times 3.00 \times 10^{8}/1.64 \times 10^{-18} = 1.21 \times 10^{-7}$ m = **121 nm**

www.iranino.co

IRANINOVIN

Now it's your turn

7 Calculate the wavelength of the radiation emitted when the electron in a hydrogen atom makes a transition from the energy level at -3.4×10^{-18} J to the level at -8.5×10^{-18} J.
(Planck constant $h = 6.6 \times 10^{-34}$ J s; speed of light $c = 3.00 \times 10^{8}$ m s^{-1})

8 The electron in a hydrogen atom makes a transition from the energy level at −13.58 eV to the level at −0.38 eV when a photon is absorbed. Calculate the frequency of the radiation absorbed.
(Planck constant $h = 6.6 \times 10^{-34}$ J s; 1 eV = 1.60×10^{-19} J)

9 The energy required to completely remove an electron in the ground state from an atom is called the ionisation energy. This energy may be supplied by the absorption of a photon, in which case the process is called photo-ionisation. Use information from Figure 25.9 to deduce the wavelength of radiation required to achieve photo-ionisation of hydrogen.
(Planck constant $h = 6.6 \times 10^{-34}$ J s; speed of light $c = 3.00 \times 10^{8}$ m s^{-1})

25.5 Band theory

We have seen that electrons in an isolated atom have specific energy levels. Electrons in an atom can have only these particular energies (see Figure 25.10a). Atoms in a gas at low pressure are, on average, sufficiently far apart for them to behave as isolated atoms. This is not the case for high pressure gases or liquids and solids where the atoms are much closer together.

Where the atoms are close together, the electrons in an atom experience forces due to the charges in neighbouring atoms. The effect is to cause the energy of the electron to change slightly. The energy is not fixed for all electrons and, consequently, the single energy level in an isolated atom broadens out to become an energy band in a solid, as illustrated in Figure 25.15.

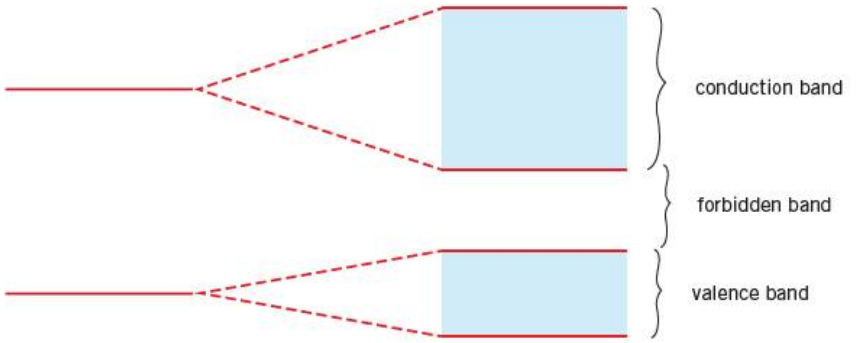


Figure 25.15 Formation of electron energy bands

In semiconductors and insulators, the lower energy bands are filled with electrons. The **valence band** is the uppermost *filled* band and the **conduction band** is the lowest *unfilled* band (see Figure 25.15). Each band consists of a very large number of closely-spaced energy levels. Between the conduction band and the valence band, there exists the **forbidden band**, see Figure 25.15.

The forbidden band is a range of energies that cannot be occupied by an electron.

Conduction in an intrinsic semiconductor

An **intrinsic semiconductor** is a substance such as pure silicon or pure germanium that, at room temperatures, has a typical resistance between that of metals and insulators.

In an intrinsic semiconductor at zero kelvin (0 K),
- the valence band is fully occupied
- the conduction band is empty
- the forbidden band has a width of about 1 eV (1 electronvolt).

www.iranino.co

IRANINOVIN

The valence band being fully occupied means that the electrons cannot move under the action of an electric field. Since these charge carriers cannot move in the valence band and there are no electrons in the conduction band, then there is no electric current and the resistance of the intrinsic semiconductor is very large.

Above zero kelvin, the electrons in the valence band will have some energy due to higher temperature. Electrons will have a distribution of energies with the mean energy increasing as the temperature rises. At any temperature above 0K, some electrons will have sufficient energy to jump across the forbidden band and enter the conduction band. These electrons in the conduction band are negative charge carriers and can migrate under the action of an electric field.

The electrons that enter the conduction band leave behind a vacancy or **hole** in the valence band. This vacancy or hole is the result of an absence of a negative charge and, thus, it behaves as a positive charge. Under the action of an electric field, a neighbouring electron fills the hole, creating the hole in its old position. Thus, the hole migrates through the solid and acts as a positive charge carrier.

As a model for hole migration, think about a bubble rising through fizzy water. The bubble is a hole in the water. The water moves from above the hole to below it and so the bubble appears to rise. What, in fact, is happening is that the centre of gravity of the water is sinking very slightly.

The electric current in an intrinsic semiconductor above 0K is due to the movement of electrons (negative charge carriers) in the conduction band and holes, positive charge carriers, moving in the opposite direction in the valence band.

The rise in temperature of the solid has a further important consequence. In the solid, the atoms are held in a lattice and, above absolute zero, the atoms in this lattice vibrate. These vibrations hinder the progress of the charge carriers through the solid. As the temperature rises, so the amplitude of the lattice vibrations increases. However, any effect on the increase of the current due to the increase in the number of charge carriers outweighs any possible reduction due to an increase in lattice vibrations. Overall, the resistance decreases with rise in temperature and, thus, current increases. This is illustrated in the graph of Figure 25.16.

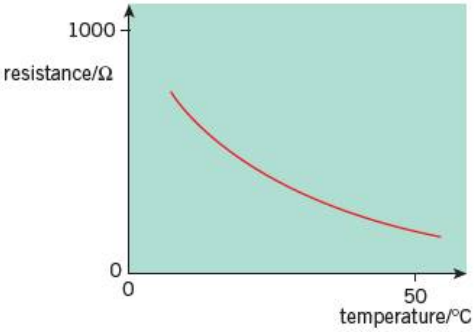


Figure 25.16 Change of resistance with temperature of a sample of an intrinsic semiconductor

For an intrinsic semiconductor above zero kelvin (0K),

- some electrons are able to jump the forbidden band to enter the conduction band
- electrons in the conduction band act as negative charge carriers
- the vacancies left behind in the valence band by the electrons entering the conduction band are known as holes
- holes behave as positive charge carriers in the valence band
- lattice vibrations increase as the temperature rises
- any effect due to increase in the number of charge carriers outweighs any effect due to increase in lattice vibrations
- overall, resistance decreases as temperature rises
- the number density, the number per unit volume (see Topic 19) for electrons in the conduction band, is equal to the number density of holes in the valence band.

www.iranino.co

IRANINOVIN

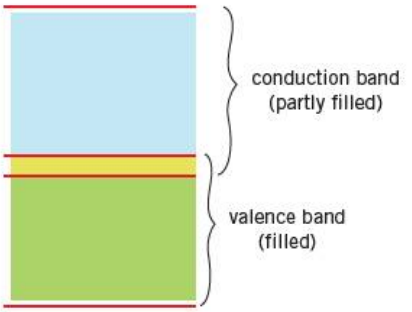


Figure 25.17 Energy bands in a metal

Conduction in a metal

In a metal
- the valence band overlaps the conduction band
- the forbidden band does not exist.

At all temperatures, the electrons can move under the action of an electric field. The outer electrons behave as if they are free and, thus, act as negative charge carriers. The number of these charge carriers does not increase with temperature because there is no forbidden band (see Figure 25.17).

Lattice vibrations do increase with rise in temperature. Thus, these vibrations cause an increase in the hindrance to the movement of the electrons. The resistance increases with temperature and, hence, the current decreases. In general, this change in resistance with temperature is comparatively small (see Figure 25.18).

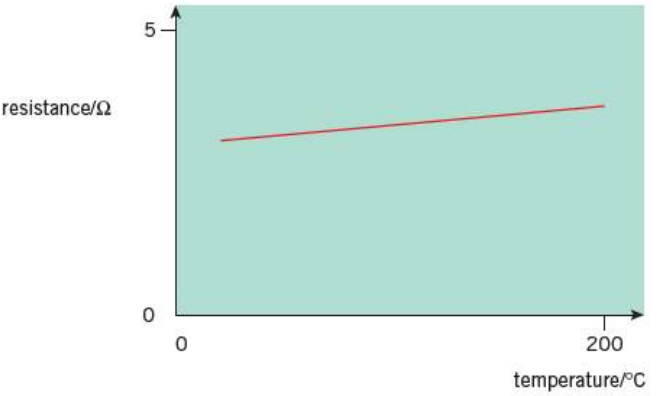


Figure 25.18 Change of resistance with temperature for a sample of a metal

Conduction in an insulator

In an insulator
- the valence band is fully occupied
- the conduction band is empty
- the forbidden band is very wide.

The wide forbidden band means that the very great majority of electrons in the valence band do not have sufficient energy to enter the conduction band. Therefore, there are very few charge carriers and the resistance is very large. When an insulator is heated to high temperatures, some electrons may have sufficient energy to enter the conduction band, then electrical breakdown of its resistance will occur.

Light-dependent resistors

Thermal energy is not the only form of energy that can causes electrons in the valence band to move to the conduction band. Semiconductor materials can be produced with a forbidden band such that photons of light promote electrons from the valence band to the conduction band. Increase in light intensity gives rise to an increase in the number of photons per unit time and an increase in the number of electrons promoted to the conduction band. Thus, the resistance of these materials will decrease as the intensity of light increases – the light-dependent resistor (LDR).

Example

Explain, using band theory, why the resistance of an intrinsic semiconductor decreases with temperature rise.

As the temperature rises, more electrons have sufficient energy to cross the forbidden band and to enter the conduction band.

www.iranino.co

IRANINOVIN

Electrons in the conduction band act as negative charge carriers under the action of an electric field.

The vacancies left in the valence band by the electrons (holes) act as positive charge carriers.

Increase in temperature increases the lattice vibrations which hinder movement of the charge carriers.

Effect due to increase in number of charge carriers outweighs effect due to the increase in lattice vibrations, so the resistance decreases with rise in temperature.

Now it's your turn

10 State and explain what is meant by a hole in the valence band of an intrinsic semiconductor.

11 Explain, using band theory, why the resistance of many metals increases with temperature rise.

25.6 The production and use of X-rays



Figure 25.19 Modern X-ray machine

Whenever a charged particle is accelerated, electromagnetic radiation is emitted. This radiation is known as Bremmstrahlung radiation or 'braking (slowing down)' radiation. The frequency of the radiation depends on the magnitude of the acceleration. The larger the acceleration (or deceleration), the greater is the frequency of the emitted photon.

X-ray photons may be produced by the bombardment of metal targets with high-speed electrons. The electrons are first accelerated through a potential difference of many thousands of volts so that they have high energy and high speed. This acceleration is, however, not sufficient for X-ray radiation to be emitted. The high-speed electrons strike a metal target, which causes the electrons to change direction and to lose kinetic energy very rapidly. Large decelerations are involved that give rise to the emission of X-ray photons. It should be remembered that not all of the energy of the electrons is emitted as X-ray photons. The majority is transferred to thermal energy in the target metal.

A simplified design of an X-ray tube is shown in Figure 25.20.

A metal filament is heated using a low-voltage supply and electrons are emitted from the filament – this is called the thermionic effect. The electrons are then accelerated towards the target, constructed of metal with a high melting point. The electrons are accelerated through a large potential difference of the order of 20–90 kV. When the electrons are stopped in the metal target, X-ray photons (Bremmstrahlung radiation) are emitted and the X-ray beam passes out of the tube through a window that is transparent to X-rays. The tube is evacuated and is made of a material that is opaque to X-rays.

www.iranino.co

IRANINOVIN

This reduces background radiation around the tube. The anode is earthed and the cathode is maintained at a high negative potential, so that the equipment is at earth potential. Since the majority of the energy of the electrons is transferred to thermal energy in the anode, the anode is either cooled or is made to spin rapidly so that the thermal energy is dispersed over a larger area of the target.

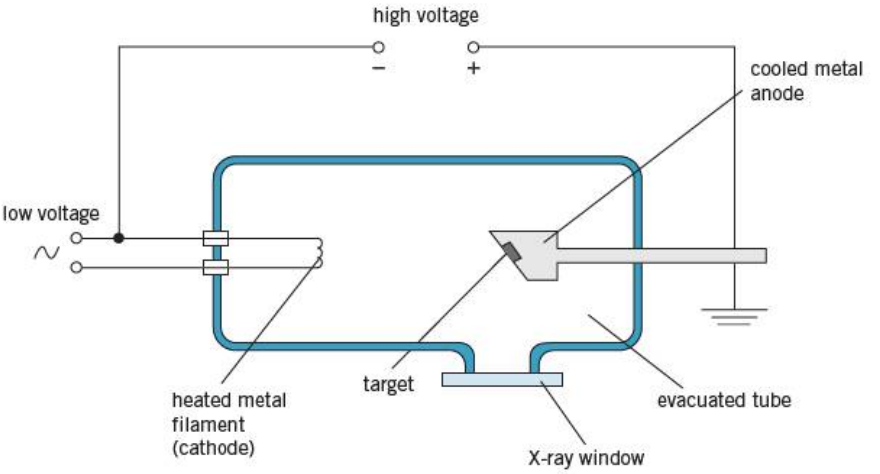


Figure 25.20 Design of an X-ray tube

Control of the X-ray beam

In order that the optimum X-ray image may be obtained, both the intensity and the hardness of the X-ray beam need to be controlled.

- The **intensity** is the wave power per unit area, and this affects the degree of blackening of the image.
- The **hardness** is the penetration of the X-ray beam, which determines the fraction of the intensity of the incident beam that can penetrate the part of the body being X-rayed. In general, the shorter the wavelength of the X-rays, the greater their penetration.

A typical X-ray spectrum showing the variation with wavelength of the intensity of an X-ray beam is shown in Figure 25.21.

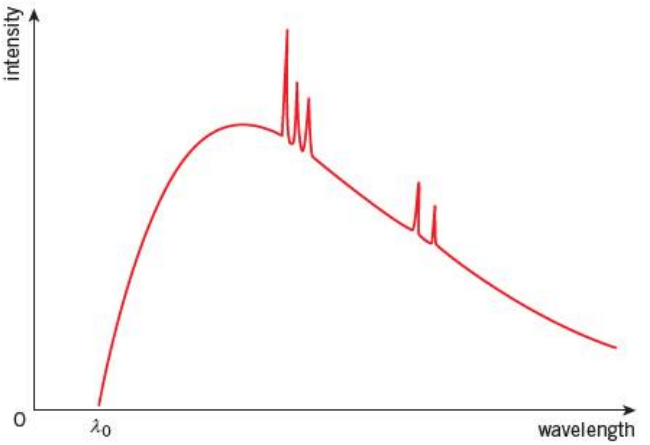


Figure 25.21 Typical X-ray spectrum

The spectrum has two distinct components. First, there is a continuous distribution of wavelengths with a sharp cut-off at the shortest wavelength λ_0. Second, sharp peaks may be observed. These sharp peaks correspond to the emission line spectrum of the target metal and are, therefore, a characteristic of the target.

www.iranino.co

IRANINOVIN

The continuous distribution comes about because the electrons, when incident on the metal target, will not all have the same deceleration but will, instead, have a wide range of values. Since the wavelength of the emitted radiation is dependent on the deceleration, there will be a distribution of wavelengths. The cut-off wavelength corresponds to an electron that is stopped in one collision in the target so that all of its kinetic energy is given up as one X-ray photon.

The kinetic energy E_K of an electron is equal to the energy gained by the electron when it is accelerated from the cathode to the anode.

$$E_K = eV$$

where e is the charge on the electron and V is the accelerating potential difference.

The energy E of a photon of wavelength λ is given by

$$E = \frac{hc}{\lambda}$$

where h is the Planck constant (see page 353). Thus, at the cut-off wavelength,

$$eV = \frac{hc}{\lambda_0}$$

and so

$$\lambda_0 = \frac{hc}{eV}$$

The accelerating potential V thus determines the cut-off wavelength λ_0. The larger the potential difference, the shorter the wavelength. The hardness of the X-ray beam is, therefore, controlled by variation of the accelerating potential difference between the cathode and the anode.

The continuous distribution of wavelengths implies that there will be X-ray photons of long wavelengths that would not penetrate the person being investigated and so would not contribute towards the X-ray image. Such long-wavelength photons would add to the radiation dose received by the person without serving any useful purpose. For this reason, the X-ray beam emerging from the X-ray tube frequently passes through aluminium filters that absorb these long-wavelength photons.

The intensity of the X-ray beam depends on the number of photons emitted per unit time and, hence, the number of electrons hitting the metal target per unit time. Since the electrons are produced by thermionic emission, then increasing the heater current in the cathode will increase the rate of production of electrons and, hence, increase the intensity of the X-ray beam.

Example

The accelerating potential difference between the cathode and the anode of an X-ray tube is 30 kV. Given that the Planck constant is 6.6×10^{-34} J s, the charge on the electron is 1.6×10^{-19} C and the speed of light in free space is 3.0×10^{8} m s^{-1}, calculate the minimum wavelength of photons in the X-ray beam.

For the minimum wavelength,

energy gained by electron = energy of photon

$$eV = \frac{hc}{\lambda_0}$$

$$1.6 \times 10^{-19} \times 30 \times 10^{3} = \frac{(6.6 \times 10^{-34} \times 3.0 \times 10^{8})}{\lambda_0}$$

$$\lambda_0 = \mathbf{4.1 \times 10^{-11}\ m}$$

Now it's your turn

12 State what is meant by the *hardness* of an X-ray beam.

13 The Planck constant is 6.6×10^{-34} J s, the charge on the electron is 1.6×10^{-19} C and the speed of light in free space is 3.0×10^{8} m s^{-1}. Calculate the minimum wavelength of photons produced in an X-ray tube for an accelerating potential difference of 45 kV.

www.iranino.co

IRANINOVIN

The X-ray image

An X-ray image is shown in Figure 25.22. This is not really an image in the sense of the real image produced by a lens. Rather, the image is like a shadow, as illustrated in Figure 25.23.



Figure 25.22 X-ray image showing a fractured collar bone

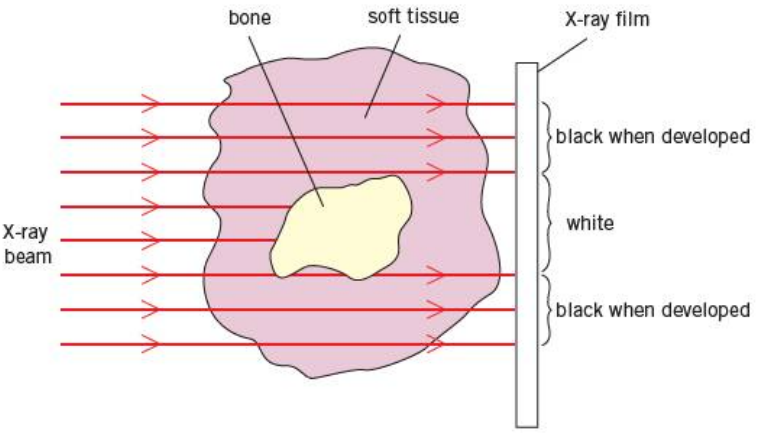


Figure 25.23 How an X-ray image is produced

The X-ray beam is incident on the body part of the patient. The X-ray beam can penetrate soft tissues (skin, fat, muscle, etc.) with little loss of intensity and so photographic film, after development, will show a dark area corresponding to these soft tissues. Bone, however, causes a greater attenuation (reduces the intensity by a greater extent) than the soft tissues and, therefore, the photographic film will be lighter in colour in areas corresponding to the positions of bones. What is produced on the film is a two-dimensional shadow of the bone and the surrounding tissues.

The quality of the shadow image produced depends on its sharpness and contrast. **Sharpness** is related to the ease with which the edges of structures can be determined. A shadow image where the bones and other organs are clearly outlined is said to be a 'sharp image'. Although an image may be sharp, it may still not be clearly visible because there is little difference in the degree of blackening between, say, the bone and the surrounding tissue. An X-ray image having a wide range of degrees of blackening in different regions is said to have good **contrast**.

A sharp image requires a parallel X-ray beam. This can be achieved by reducing the area of the target anode in the X-ray tube, limiting the size of the aperture through which the X-ray beam passes, and reducing scattering of the emergent beam. Figure 25.24 illustrates the effect of changing the area of the target anode.

www.iranino.co

IRANINOVIN

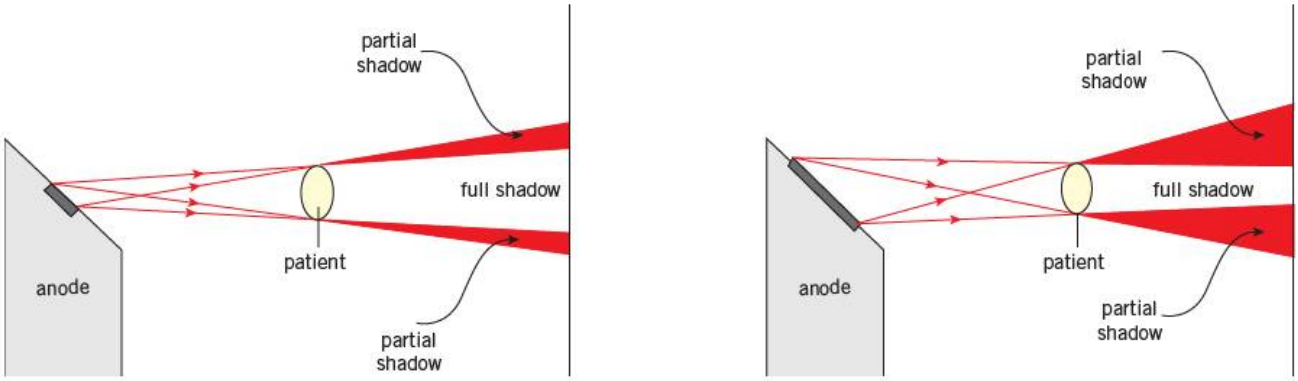


Figure 25.24 Effect of changing the area of the target anode on sharpness

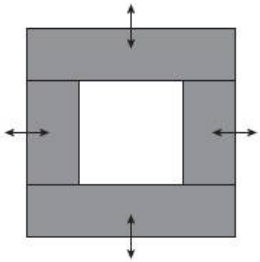


Figure 25.25 Limiting aperture size

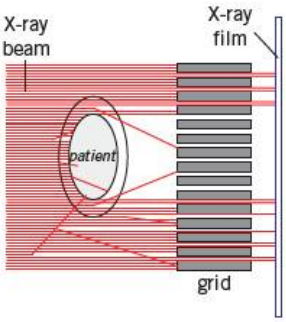


Figure 25.26 Use of a grid



Figure 25.27 X-ray of stomach after a barium meal

The full shadow produces an area that is white on development of the photographic film. Where there is no shadow, the image will be black. There is a region of partial shadow, or greyness, where the image gradually changes from white to black. If the image is to be sharp, then this area of greyness must be reduced as much as possible. The area of the target anode should be kept to a minimum.

A reduction in the grey area at the edge of the image can also be achieved by limiting the size of the aperture through which the X-ray beam passes. This is achieved by using overlapping metal sheets, as shown in Figure 25.25. The aperture size may be varied by sliding the metal sheets over one another.

The emerging parallel X-ray beam should then give rise to a sharp image. However, as a result of interactions between photons and any substance through which the beam passes (this could even be air), some photons will be scattered. Scattered photons will result in a loss of sharpness. These scattered photons may be absorbed in a metal grid placed in front of the X-ray film (Figure 25.26).

Good contrast is achieved when neighbouring body organs and tissues absorb the X-ray photons to very different extents. This is usually the case for bone and muscle. This is, however, not the case where, for example, the stomach is to be investigated. The patient is then asked to swallow a solution of barium sulfate – a 'barium meal' (Figure 25.27).

Barium is a good absorber of X-ray photons. As a result, when the barium sulfate solution coats the inside of the stomach, the outline of the stomach will show up clearly on the image. Similarly, blood vessels can be made visible by injecting a radio-opaque dye into the bloodstream.

Contrast also depends on other factors, such as exposure time. Contrast may be improved by backing the X-ray film with a fluorescent material.

The attenuation of X-rays

When a beam of X-ray photons passes through a medium, absorption processes occur that reduce the intensity of the beam. The intensity of a parallel beam is reduced by the same fraction each time the beam passes through equal thicknesses of the medium. Consequently, the variation of the percentage of the intensity transmitted with thickness of absorber may be shown as in Figure 25.28.

It can be seen that the same thickness of medium is always required to reduce the transmitted beam intensity by 50%, no matter what starting point is chosen. This thickness of medium is called the **half-value thickness (HVT)** and is denoted by the symbol $x_{\frac{1}{2}}$.

The decrease in transmitted intensity is an exponential decrease. Consider a parallel beam having an incident intensity I_0. The medium (the absorber) has thickness x and the transmitted intensity is I, as illustrated in Figure 25.29.

www.iranino.co

IRANINOVIN

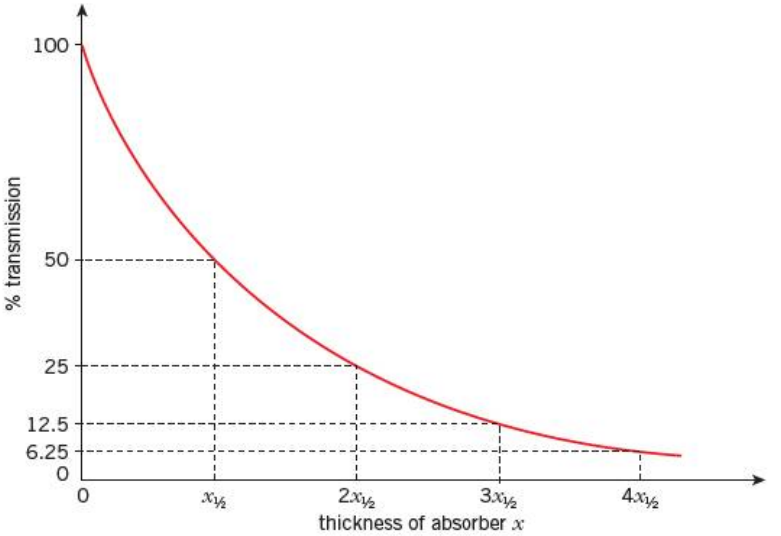


Figure 25.28 The percentage transmission of X-rays in a medium

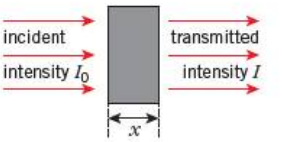


Figure 25.29 Absorption of X-rays in a medium

The transmitted intensity is given by the expression

$$I = I_0 e^{-\mu x} \text{ or } I = I_0 \exp(-\mu x)$$

where μ is a constant that is dependent on the medium and on the energy of the X-ray photons, known as the **linear attenuation coefficient** or the **linear absorption coefficient** of the medium. The unit of μ is mm^{-1} or cm^{-1}.

For a thickness $x_{½}$ (the half-value thickness or HVT) of the medium, the intensity I will be equal to $½I_0$. Hence,

$$\tfrac{1}{2}I_0 = I_0 \exp(-\mu x_{½})$$

which gives

$$\ln 2 = \mu x_{½}$$

Approximate values of the linear attenuation (absorption) coefficient μ for some substances are given in Table 25.2.

Note that the expression $I = I_0 \exp(-\mu x)$ applies to a parallel beam. If the beam is not parallel, then there will be further changes in intensity without any absorption. For example, the intensity of a divergent beam decreases with distance from the source.

Table 25.2 Some approximate values of linear attenuation (absorption) coefficient

substance	μ/cm^{-1}
copper	7
water	0.3
bone	3
fat	0.9

Example

The linear absorption coefficient of copper is 0.693 mm^{-1}. Calculate

(a) the thickness of copper required to reduce the incident intensity by 50% (the half-value thickness or HVT),

(b) the fraction of the incident intensity of a parallel beam that is transmitted through a copper plate of thickness 1.2 cm.

(a) $I = I_0 e^{-\mu x}$ or $I = I_0 \exp\left(-\mu x_{\frac{1}{2}}\right)$

$$I/I_0 = 0.50 = \exp\left(-0.693 x_{\frac{1}{2}}\right)$$

$$\ln 0.50 = -0.693 x_{½}$$

$$x_{½} = \mathbf{1.0\,mm}$$

(b) $I/I_0 = \exp(-0.693 \times 12)$

$I/I_0 = \mathbf{2.4 \times 10^{-4}}$

www.iranino.co

IRANINOVIN

Now it's your turn

14 **(a)** State what is meant by the *sharpness* of an X-ray image.
(b) Explain how the sharpness of an X-ray image may be improved.

15 For one particular energy of X-ray photons, water has a linear attenuation (absorption) coefficient of $0.29\,cm^{-1}$. Calculate the depth of water required to reduce the intensity of a parallel beam of these X-rays to 1.0×10^{-3} of its incident intensity.

16 The linear attenuation (absorption) coefficients of bone and of the soft tissues surrounding the bone are $2.9\,cm^{-1}$ and $0.95\,cm^{-1}$ respectively. A parallel beam of X-rays is incident, separately, on a bone of thickness 3.0 cm and on soft tissue of thickness 5.0 cm. Calculate the ratio:

$$\frac{\text{intensity transmitted through bone}}{\text{intensity transmitted through soft tissue}}$$

Computed tomography (CT scanning)

The image produced on X-ray film, as outlined earlier, is a 'shadow' or 'flat' image. There is little, if any, indication of depth. That is, the position of an organ within the body is not apparent. Also, soft tissues lying behind structures that are very dense cannot be detected. Tomography is a technique whereby a three-dimensional image or 'slice' through the body may be obtained. The image is produced by **computed tomography** using what is known as a CT scanner (Figure 25.30). An example of such an image is shown in Figure 25.31.



Figure 25.30 Radiologist moving a patient into a CT scanner

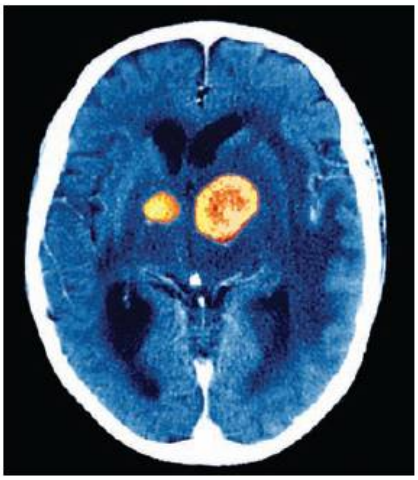


Figure 25.31 CT scan through the head of a patient with a cerebral lymphoma

In this technique, a series of X-ray images is obtained. Each image is taken through the section or slice of the body from a different angle, as illustrated in Figure 25.32.

Data for each individual X-ray image and angle of viewing is fed into a high-power computer. A two-dimensional image of the slice is computed. This is then repeated for successive slices. The computer enables the images of each slice to be combined so that a complete three-dimensional image of the whole object is obtained, which can then be viewed from any angle.

The basic principles of CT scanning may be illustrated using a simple cubic shape as shown in Figure 25.33.

The aim of CT scanning is to produce an image of a section through the cube from measurements made about its axis. The section, or 'slice' through the cube is divided up into a series of small units, called **voxels**. Each voxel will absorb the X-ray beam to a different extent. The intensity transmitted through each voxel alone can be given a number, referred to as a **pixel**. The various pixels are built up from measurements of the X-ray intensity along different directions through the section or slice.

www.iranino.co

IRANINOVIN

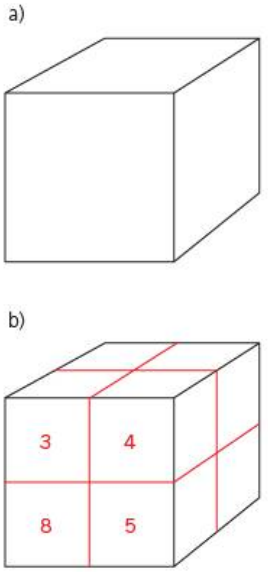


Figure 25.33 The first section showing the pixels

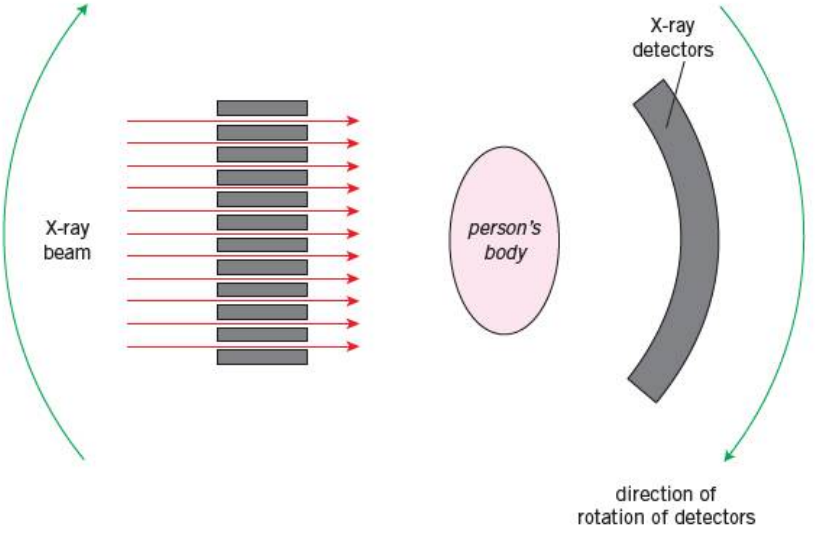


Figure 25.32 Arrangement for a CT scan

Suppose that the cube in Figure 25.33a is divided into eight voxels. The cube can be thought to consist of two slices or sections. For the first section, let the pixels be as shown in Figure 25.33b. The purpose of the CT scan is to reproduce these pixels in their correct positions.

When the X-ray beam is directed at the section from the left, as shown in Figure 25.34, the detectors will give readings of 7 and 13. The voxels will be partially completed as shown.

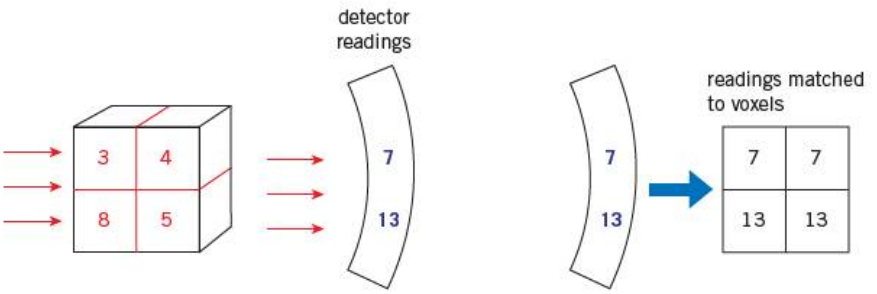


Figure 25.34 The first set of detector readings

The X-ray tube and the detectors are now rotated through 45°. The new detector readings are 4, 8 and 8. These readings are added to the readings already in the voxels, as shown in Figure 25.35.

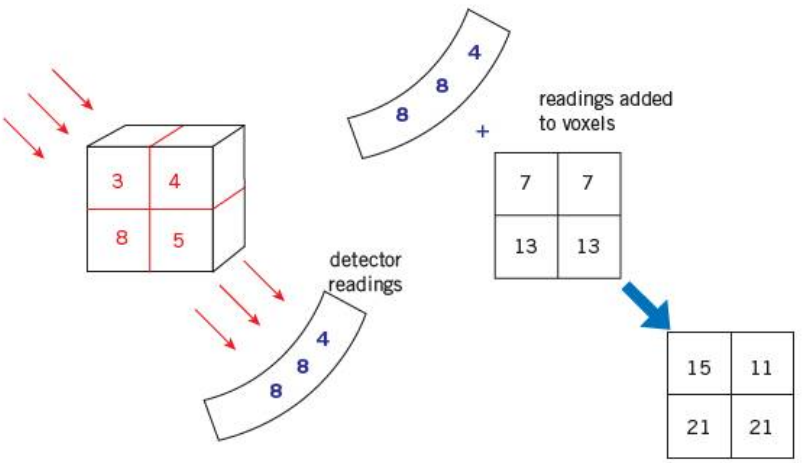


Figure 25.35 The second set of detector readings

www.iranino.co

IRANINOVIN

After rotation through a further 45°, a third set of detector readings is taken. These readings are added to the voxel readings. The result is shown in Figure 25.36.

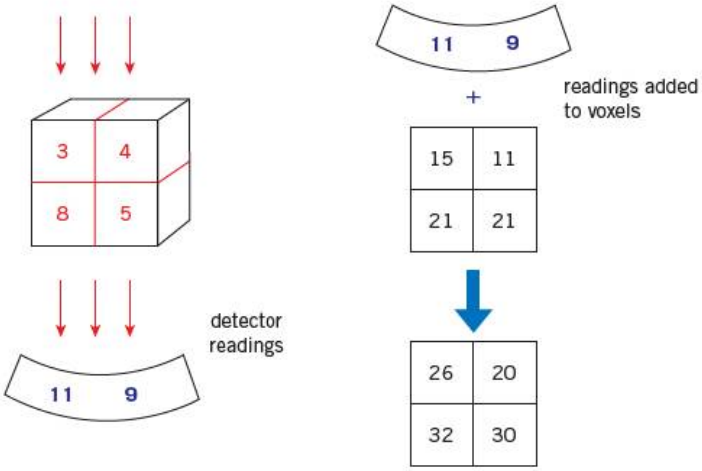


Figure 25.36 The third set of detector readings

After rotation through a further 45° a final set of readings is taken. Once again the readings are added to those already in the voxels, giving the result shown in Figure 25.37.

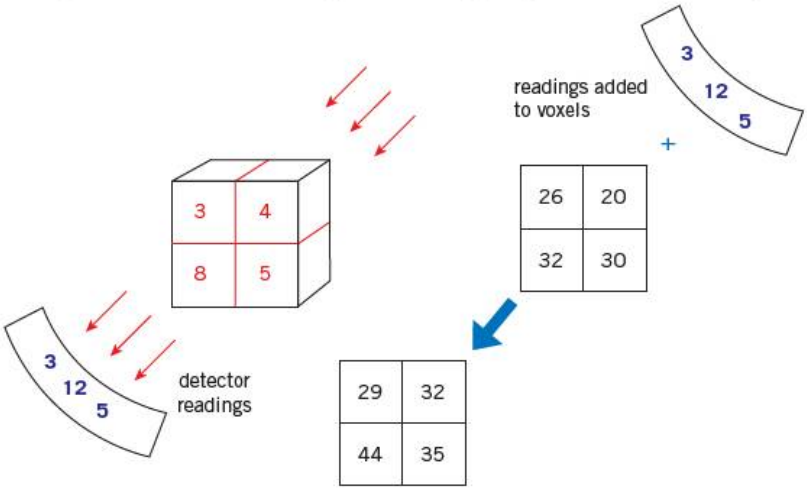


Figure 25.37 The fourth set of detector readings

The resulting pattern of the pixels is shown in Figure 25.38a.

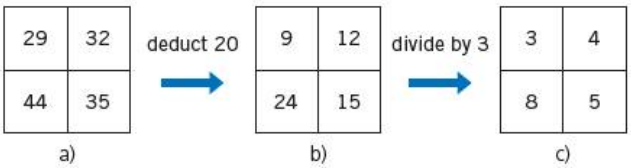


Figure 25.38 The final result

Having summed all the measurements, it remains to reduce these measurements to that of the original. This is achieved in two stages.

1. The background intensity must be removed. This background is equal to the sum of the detector readings for any one position. In this case, the sum is 20. This background is deducted from each pixel, as shown in Figure 25.38b.
2. Allowance must now be made for the fact that more than one view was made of the section. In this example, there were four sets of readings and consequently each pixel reading is divided by 3.

www.iranino.co

IRANINOVIN

The final result is shown in Figure 25.38c – and, note, this is the same as Figure 25.33b.

Once the pattern of pixels for one section has been obtained, the CT scanner is moved relative to the object so that the next neighbouring section is analysed. This procedure is repeated until the whole of the object to be analysed has been scanned.

The analysis above is a very simple example. In practice, the image of each section is built up from a large number of units or voxels. The larger the number of voxels, the better the definition. This is similar, in principle, to the digital camera. The use of a large number of voxels implies that measurements must be made from a large number of different angles. The storage of the data for each angle and its construction into a final image on a screen requires a powerful computer. The reconstruction of the intensity in each voxel will involve more than one million separate computations. All the data for all the sections is stored in the computer memory so that a three-dimensional image of the whole object is formed. This enables sections of the image to be viewed from many different angles. Finally, the computer enables the brightness and contrast of the image to be varied so that the optimum image may be obtained.

Example

Compare the image produced during an X-ray investigation and that produced in CT scanning.

An X-ray image is a two dimensional projection onto a flat screen of a three-dimensional object. A CT scan is a three-dimensional image. The computer in which the data for the image is stored enables different sections to be viewed at different angles.

Now it's your turn

17 The principles of CT scanning have been understood for some time. However, scanners could not be developed until large powerful computers were available. By reference to the image produced in a CT scan, suggest why such a computer is necessary.

18 A simple object consists of four voxels. The object is scanned from four different directions, each at 45° to the next. The detector measurements for each individual voxel are summed and the result is shown in Figure 25.39.
The total of the readings of the detectors in any one position is 22. Determine the pattern of the pixels in the voxels.

40	49
34	31

Figure 25.39

19 (a) Outline how X-ray images are used to build up the image produced in a CT scan.
(b) Explain why the radiation dose received during a CT scan is greater than that for an X-ray 'photograph'.

Summary

Energy of a photon

- A photon is a quantum (or packet) of energy of electromagnetic radiation having energy equal to the product of the Planck constant and the frequency of the radiation, $E = hf$.

Photoelectric emission of electrons

- Electrons may be emitted from metal surfaces if the metal is illuminated by electromagnetic radiation. This phenomenon is called photoelectric emission.
- Photoelectric emission cannot be explained by the wave theory of light. It is necessary to use the quantum theory, in which electromagnetic radiation is thought of as consisting of packets of energy called photons.
- The work function energy ϕ of a metal is the minimum energy needed to free an electron from the surface of the metal.
- The Einstein photoelectric equation is: $hf = \phi + \frac{1}{2}m_e v_{max}^2$
- The threshold frequency f_0 is given by: $hf_0 = \phi$

Wave–particle duality

- Moving particles show wave-like properties.
- The de Broglie wavelength is given by $\lambda = h/p$, where p is the momentum of the particle and h is the Planck constant.

www.iranino.co

IRANINOVIN

Energy levels in atoms and line spectra

- Electrons in isolated atoms can have only certain energies. These energies may be represented in an energy level diagram.
- Electrons in a given energy level may absorb energy and make a transition to a higher energy level.
- Excited electrons may return to a lower level with the emission of a photon. The frequency f of the emitted radiation is given by $E_2 - E_1 = hf$, where E_2 and E_1 are the energies of the upper and lower levels and h is the Planck constant; the wavelength λ is given by $\lambda = c/f$, where c is the speed of light.
- When an electron absorbs energy from white light and moves to a higher energy level, a line absorption spectrum is produced.

Band theory

- Current is due to a movement of negative charge carriers (electrons) in the conduction band and to a movement of positive charge carriers (holes) in the valence band.
- The forbidden band in insulators is very wide. In intrinsic semiconductors it is approximately 1 eV in width.
- In metals, the valence band and the conduction band overlap.
- Lattice vibrations increase in amplitude as temperature rises.
- In intrinsic semiconductors, effects due to increased numbers of charge carriers outweigh effects due to increased lattice vibrations.
- In metals, the number of charge carriers does not increase significantly and increased lattice vibrations cause an increase in resistance with temperature rise.

Production and use of X-rays

- Remote sensing enables investigations to be made where there is no contact with the object under investigation.
- X-rays are produced when high-speed electrons are stopped by a metal target.
- An X-ray image is a 'shadow' of structures in which the X-ray beam is attenuated.
- The attenuation in the intensity of an X-ray beam is given by: $I = I_0 e^{-\mu x}$
- Computed tomography (CT scanning) enables an image of a section through the body to be obtained by combining many X-ray images, each one taken from a different angle.

Examination style questions

1 A zinc plate is placed on the cap of a gold-leaf electroscope and charged negatively. The gold leaf is seen to deflect. Explain fully the following observations.

- **a** When the zinc plate is illuminated with red light, the gold leaf remains deflected.
- **b** When the zinc plate is irradiated with ultra-violet radiation, the leaf collapses.
- **c** When the intensity of the ultra-violet radiation is increased, the leaf collapses more quickly.
- **d** If the zinc plate is initially charged positively, the gold leaf remains deflected regardless of the nature of the incident radiation.

2 A beam of monochromatic light of wavelength 630 nm transports energy at the rate of 0.25 mW. Calculate the number of photons passing a given cross-section of the beam each second.

(Planck constant $h = 6.6 \times 10^{-34}$ J s; speed of light $c = 3.0 \times 10^{8}$ m s^{-1})

3 The work function energy of a certain metal is 4.0×10^{-19} J.

- **a** Calculate the longest wavelength for which photoemission is obtained.
- **b** This metal is irradiated with ultra-violet radiation of wavelength 250 nm. Calculate, for the emitted electrons:
  - **i** the maximum kinetic energy,
  - **ii** the maximum speed.

  (Planck constant $h = 6.6 \times 10^{-34}$ J s; speed of light $c = 3.0 \times 10^{8}$ m s^{-1}; electron mass $m_e = 9.1 \times 10^{-31}$ kg

4 Calculate the de Broglie wavelengths of:

- **a** a ball of mass 0.30 kg moving at 50 m s^{-1},
- **b** a bullet of mass 50 g moving at 500 m s^{-1},
- **c** an electron of mass 9.1×10^{-31} kg moving at 3.0×10^{7} m s^{-1},
- **d** a proton of mass 1.7×10^{-27} kg moving at 3.0×10^{6} m s^{-1}.

(Planck constant $h = 6.6 \times 10^{-34}$ J s)

www.iranino.co

IRANINOVIN

5 Atoms in the gaseous state (for example, low-pressure gas in a discharge tube) produce an emission spectrum consisting of a series of separate lines. The wavelengths of these lines are characteristic of the particular atoms involved. Hot atoms in the solid state (for example, a hot metal filament in an electric light bulb) produce a continuous emission spectrum which is characteristic of the temperature of the filament rather than of the atoms involved. Suggest reasons for this difference.

6 When the visible spectrum emitted by the Sun is observed closely it is noted that light of certain frequencies is missing and in their place are dark lines.

a Explain how the cool outer gaseous atmosphere of the Sun could be responsible for the absence of these frequencies.

b Suggest how an analysis of this spectrum could be used to determine which gases are present in the Sun's atmosphere.

7 An explanation of the photoelectric effect includes the terms photon energy and work function energy.

a Explain what is meant by

i *a photon,* [2]

ii *work function energy.* [1]

b In an experiment to investigate the photoelectric effect, a student measures the wavelength λ of the light incident on a metal surface and the maximum kinetic energy E_{max} of the emitted electrons. The variation with E_{max} of $\frac{1}{\lambda}$ is shown in Fig. 25.40.

i The work function energy of the metal surface is Φ. State an equation, in terms of λ, Φ and E_{max}, to represent conservation of energy for the photoelectric effect. Explain any other symbols you use. [2]

ii Use your answer in **i** and Fig. 25.40 to determine

- the work function energy Φ of the metal surface, [2]
- a value for the Planck constant. [3]

Cambridge International AS and A Level Physics, 9702/43 Oct/Nov 2011 Q 7

8 By reference to band theory, explain why

a the resistance of an intrinsic semiconductor decreases with temperature rise, [4]

b the resistance of a metal increases with temperature rise. [3]

9 a Explain how the hardness of an X-ray beam is controlled by the accelerating voltage in the X-ray tube. [2]

b The attenuation of a parallel beam of X-ray radiation is given by the expression

$$\frac{I}{I_0} = e^{-\mu x}$$

where μ is the linear attenuation (absorption) coefficient and x is the thickness of the material through which the beam passes.

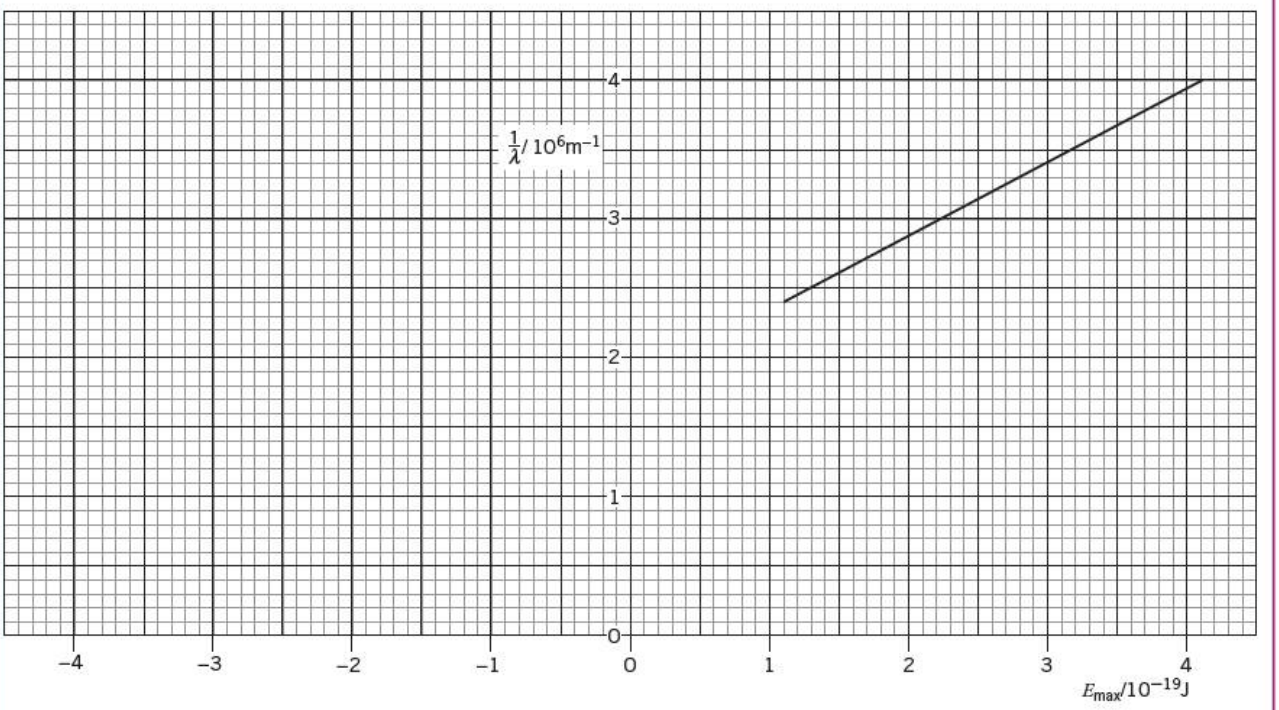


Fig. 25.40

www.iranino.co

IRANINOVIN

i State
- what is meant by *attenuation*, [1]
- why the expression applies only to a parallel beam. [2]

ii The linear attenuation coefficients for X-rays in bone and in soft tissue are $2.9\,cm^{-1}$ and $0.95\,cm^{-1}$ respectively.
Calculate, for a parallel X-ray beam, the ratio [2]

$$\frac{\text{fraction } \frac{I}{I_0} \text{ of intensity transmitted through bone of thickness } 2.5\,cm}{\text{fraction } \frac{I}{I_0} \text{ of intensity transmitted through soft tissue of thickness } 6.0\,cm}$$

Cambridge International AS and A Level Physics, 9702/42 May/June 2013 Q 11

www.iranino.co

IRANINOVIN

26 Nuclear physics

By the end of this topic, you will be able to:

26.3 (a) show an appreciation of the association between energy and mass as represented by $E = mc^2$ and recall and use this relationship
(b) understand the significance of the terms mass defect and mass excess in nuclear reactions
(c) represent simple nuclear reactions by nuclear equations of the form ${}^{14}_{7}\mathrm{N} + {}^{4}_{2}\mathrm{He} \rightarrow {}^{17}_{8}\mathrm{O} + {}^{1}_{1}\mathrm{H}$
(d) define and understand the terms mass defect and binding energy
(e) sketch the variation of binding energy per nucleon with nucleon number
(f) explain what is meant by nuclear fusion and nuclear fission
(g) explain the relevance of binding energy per nucleon to nuclear fusion and to nuclear fission

26.4 (a) infer the random nature of radioactive decay from the fluctuations in count-rate
(b) show an appreciation of the spontaneous and random nature of nuclear decay
(c) define the terms activity and decay constant and recall and solve problems using $A = \lambda N$
(d) infer and sketch the exponential nature of radioactive decay and solve problems using the relationship $x = x_0 e^{-\lambda t}$, where x could represent activity, number of undecayed particles or received count rate
(e) define half-life
(f) solve problems using the relation $\lambda = \frac{0.693}{t_{\frac{1}{2}}}$

Starting points

- An atom may be modelled as a massive, but very small, positively charged nucleus surrounded by negatively charged electrons.
- A nucleus contains protons and neutrons.

26.3 Mass defect and nuclear binding energy

Mass defect

At a nuclear level, the masses we deal with are so small that it would be very clumsy to measure them in kilograms. Instead, we measure the masses of nuclei and nucleons in **atomic mass units (u)**.

One atomic mass unit (1 u) is defined as being equal to one-twelfth of the mass of a carbon-12 atom. 1 u is equal to 1.66×10^{-27} kg.

Using this scale of measurement, to six decimal places, we have

proton mass $m_p = 1.007276$ u
neutron mass $m_n = 1.008665$ u
electron mass $m_e = 0.000549$ u

Because all atoms and nuclei are made up of protons, neutrons and electrons, we should be able to use these figures to calculate the mass of any atom or nucleus.

www.iranino.co

IRANINOVIN

For example, the mass of a helium-4 nucleus, consisting of two protons and two neutrons, should be

$(2 \times 1.007276)\text{u} + (2 \times 1.008665)\text{u} = 4.031882\text{u}$.

However, the actual mass of a helium nucleus is 4.001508u.

The difference between the expected mass and the actual mass of a nucleus is called the **mass defect** of the nucleus. In the case of the helium-4 nucleus, the mass defect is $4.031882 - 4.001508 = 0.030374\text{u}$.

The mass defect of a nucleus is the difference between the total mass of the separate nucleons and the combined mass of the nucleus.

Example

Calculate the mass defect for a carbon-14 ($^{14}_{6}C$) nucleus. The measured mass is 14.003240u.

The nucleus contains 6 protons and 8 neutrons, of total mass
$(6 \times 1.007276) + (8 \times 1.008665) = 14.112976\text{u}$.
The mass defect is $14.112976 - 14.003240 =$ **0.109736u**.

Now it's your turn

1 Calculate the mass defect for a nitrogen-14 ($^{14}_{6}N$) nucleus. The measured mass is 14.003070u.

Mass–energy equivalence

In 1905, Albert Einstein proposed that there is an equivalence between mass and energy. The relationship between energy E and mass, or change in mass m, is

$$E = mc^2$$

where c is the speed of light. E is measured in joules, m in kilograms and c in metres per second.

Using this relation, we can calculate that 1.0kg of matter is equivalent to $1.0 \times (3.0 \times 10^8)^2 = 9.0 \times 10^{16}\text{J}$.

The mass defect of the helium nucleus, calculated previously as 0.030374u, is equivalent to

$0.030374 \times 1.66 \times 10^{-27} \times (3.00 \times 10^8)^2 = 4.54 \times 10^{-12}\text{J}$.

(Note that the mass in u must be converted to kg by multiplying by 1.66×10^{-27}.)

The joule is an inconveniently large unit to use for nuclear calculations. A more convenient energy unit is the mega electronvolt (MeV), as many energy changes that take place in the nucleus are of the order of several MeV. One mega electronvolt is the energy gained by one electron when it is accelerated through a potential difference of one million volts.

Since *electrical energy* = *charge* × *potential difference*, and the electron charge is $1.60 \times 10^{-19}\text{C}$,

$1\,\text{MeV} = 1.60 \times 10^{-19} \times 1.00 \times 10^6$

or

$$1\,\text{MeV} = 1.60 \times 10^{-13}\text{J}$$

The energy equivalent of the mass defect of the helium nucleus is thus $4.54 \times 10^{-12}/1.60 \times 10^{-13} = 28.4\,\text{MeV}$.

If mass is measured in u and energy in MeV,

1 u is the equivalent of 931 MeV

Binding energy

Within the nucleus there are strong forces which bind the protons and neutrons together. To completely separate all these nucleons requires energy. This energy is referred to as the

www.iranino.co

IRANINOVIN

binding energy of the nucleus. Stable nuclei, those which have little or no tendency to disintegrate, have large binding energies. Less stable nuclei have smaller binding energies.

Similarly, when protons and neutrons are joined together to form a nucleus, this binding energy must be released. The binding energy is the energy equivalent of the mass defect.

We have seen that the binding energy of the helium-4 nucleus is 28.4 MeV. 28.4 MeV of energy is required to separate, to infinity, the two protons and the two neutrons of this nucleus.

Binding energy is the energy equivalent of the mass defect of a nucleus. It is the energy required to separate to infinity all the nucleons of a nucleus.

Example

Calculate the binding energy, in MeV, of a carbon-14 nucleus with a mass defect of 0.109736 u.

Using the equivalence 1 u = 931 MeV, 0.109736 u is equivalent to 102 MeV. Since the binding energy is the energy equivalent of the mass defect, the binding energy = **102 MeV**.

Now it's your turn

2 Calculate the binding energy, in MeV, of a nitrogen-14 nucleus with a mass defect of 0.108517 u.

Nuclear equations

The nucleus of any nuclide can be represented using nuclide notation (AS Level Topic 26). This notation is useful when we wish to consider a nuclear reaction. For example, when a helium nucleus bombards a nitrogen nucleus, the reaction can be represented by

$$^{14}_{7}\text{N} + ^{4}_{2}\text{He} \rightarrow ^{17}_{8}\text{O} + ^{1}_{1}\text{H}$$

For this reaction to take place, then three conditions must be met.

1 Conservation of proton number (7 + 2 = 8 + 1)
2 Conservation of nucleon number (14 + 4 = 17 + 1)
3 Conservation of mass–energy.

mass of $^{14}_{7}\text{N}$ = 14.003074 u
mass of $^{4}_{2}\text{He}$ = 4.002604 u
mass of $^{17}_{8}\text{O}$ = 16.99913 u
mass of $^{1}_{1}\text{H}$ = 1.007825 u

The change in mass is 18.005678 u – 18.006955 u = (–)1.277 × 10^{-3} u

This change in mass is equivalent to 1.2 MeV.

Note that the mass of the products is greater than the mass of the reacting nuclei. There is a mass excess.

For this reaction to take place then, by conservation of mass–energy, the helium nucleus must have kinetic energy of at least 1.2 MeV when it bombards the nitrogen nucleus.

For a reaction to occur spontaneously, there must be a mass defect so that the products of the reaction have some kinetic energy and mass–energy is conserved.

Stability of nuclei

A stable nucleus is one which has a very low probability of decay (see AS Level Topic 26). Less stable nuclei are more likely to disintegrate. A useful measure of stability is the **binding energy per nucleon** of the nucleons in the nucleus.

Binding energy per nucleon is defined as the total energy needed to completely separate all the nucleons in a nucleus divided by the number of nucleons in the nucleus.

Figure 26.1 shows the variation with nucleon number of the binding energy per nucleon for different nuclides.

www.iranino.co

IRANINOVIN

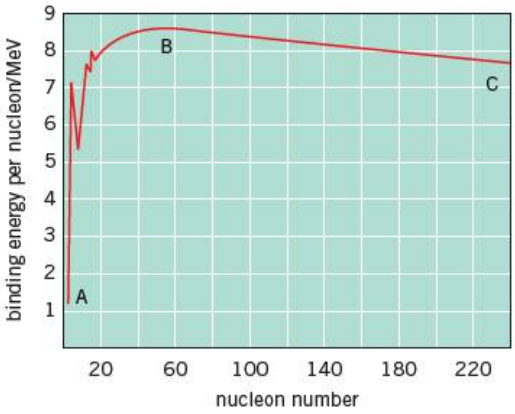


Figure 26.1 Binding energy per nucleon against nucleon number

The most stable nuclides are those with the highest binding energy per nucleon. That is, those near point B on the graph. Iron ($^{56}_{26}$Fe) is one of these stable nuclides. Typically, very stable nuclides have binding energies per nucleon of about 8 MeV.

Light nuclei, between A and B on the graph, may combine or fuse to form larger nuclei with larger binding energies per nucleon. This process is called **nuclear fusion**. For the process to take place, conditions of very high temperature and pressure are required, such as in stars like the Sun.

Heavy nuclei, between B and C on the graph, when bombarded with neutrons, may break into two smaller nuclei, again with larger binding energy per nucleon values. This process is called **nuclear fission**.

Figure 26.2 highlights nuclides which may undergo fusion (blue part of curve) or fission (red part of curve) in order to increase their binding energy per nucleon.

When nuclear fusion or fission takes place, the nucleon numbers of the nuclei involved change. A higher binding energy per nucleon is achieved, and this is accompanied by a release of energy. This release of energy during fission reactions is how the present generation of nuclear power stations produce electricity.

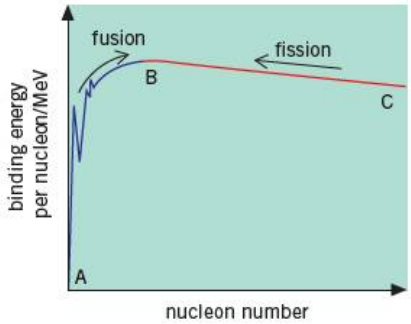


Figure 26.2

Example

The binding energy of a helium-4 nucleus is 28.4 MeV. Calculate the binding energy per nucleon.

The helium-4 nucleus has 4 nucleons. The binding energy per nucleon is thus 28.4/4 = **7.1 MeV per nucleon**.

Now it's your turn

3 The binding energy of a carbon-14 nucleus is 102 MeV. Calculate the binding energy per nucleon.

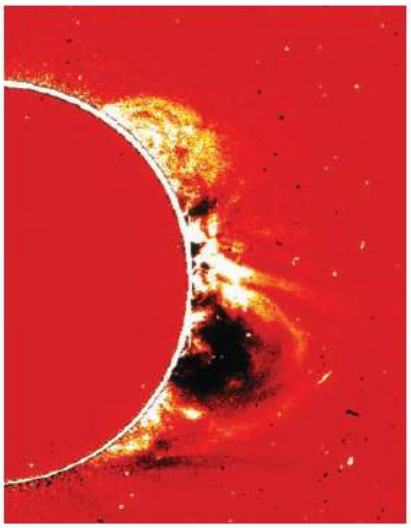


Figure 26.3 The Sun's surface

www.iranino.co

Nuclear fusion

Most of the energy on Earth comes from the Sun, where it is produced by nuclear fusion reactions. Light nuclei, such as isotopes of hydrogen, join together to produce heavier, more stable nuclei, and in doing so release energy.

Figure 26.3 shows eruptions from the Sun's surface caused by particularly energetic fusion reactions.

Nuclear fusion occurs when two light nuclei combine to form a nucleus of greater mass.

One such fusion reaction is

$$^{2}_{1}\text{H} + ^{2}_{1}\text{H} \rightarrow ^{3}_{2}\text{H} + ^{1}_{0}\text{n} + \text{energy}$$

IRANINOVIN

From the binding energy per nucleon curve (Figure 26.1) we see that the binding energy per nucleon for light nuclei, such as hydrogen, is low. But if two light nuclei are made to fuse together, they may form a new heavier nucleus which has a higher binding energy per nucleon. It will be more stable than the two lighter nuclei from which it was formed. Because of this difference in stability, a fusion reaction such as this will release energy.

Although fusion reactions are the source of solar energy, we are, at present, unable to duplicate this reaction in a controlled manner on Earth. This is because the nuclei involved in fusion have to be brought very close together. Conditions of extremely high temperature and pressure, similar to those found at the centre of the Sun, are required. Reactions requiring these conditions are called **thermonuclear reactions**. Some fusion reactions involving hydrogen isotopes have been made to work in the Joint European Torus (JET), although not yet in a controlled, sustainable manner. In 2006, an international consortium agreed to undertake the ITER (International Tokamak Engineering Research) project, which is designed to produce up to 500 MW of fusion power sustained for over 400 seconds by the fusion of a ^{2}H–^{3}H mixture. Construction on a site in southern France will take several years.

Nuclear fission

Within the nucleus of an atom, the nucleons experience both attractive and repulsive forces. The attractive force is called the **strong nuclear force**. This acts like a 'nuclear glue' to hold the nucleons together. The repulsive forces are the electric (Coulomb-law) forces between the positively charged protons. Gravitational forces of attraction exist, but are negligible in comparison to the other forces. Stable nuclei have much larger attractive forces than repulsive forces. Stable nuclides generally have approximately the same number of neutrons and protons in the nucleus. That is, the neutron-to-proton ratio is close to one. In heavy nuclei such as uranium and plutonium, there are far more neutrons than protons, giving a neutron-to-proton ratio of more than one. For example, uranium-235 has 92 protons and 143 neutrons, giving a neutron-to-proton ratio of 1.55. This leads to a much lower binding energy per nucleon compared with iron, and such nuclides are less stable. Any further increase in the number of neutrons in such nuclei is likely to cause the nucleus to undergo **nuclear fission**.

Nuclear fission is the splitting of a heavy nucleus into two lighter nuclei of approximately the same mass.

When a uranium-235 nucleus absorbs a neutron, it becomes unstable and splits into two lighter, more stable nuclei. There are many possible nuclear reactions, one of which is

$$^{235}_{92}U + {}^{1}_{0}n \rightarrow {}^{236}_{92}U \rightarrow {}^{141}_{56}Ba + {}^{92}_{36}Kr + 3{}^{1}_{0}n + \text{energy}$$

This process is called *induced* nuclear fission, because it is started by the capture of a neutron by the uranium nucleus.

Each of the fission reactions described by this equation results in the release of three neutrons. Other possible fission reactions release two or three neutrons. If these neutrons are absorbed by other uranium-235 nuclei, these too may become unstable and undergo fission, thereby releasing even more neutrons. The reaction is described as being a **chain reaction** which is accelerating. This is illustrated in Figure 26.4. If this type of reaction continues uncontrolled, a great deal of energy is released in a short time, and a nuclear explosion results.

If the number of neutrons which take part in the chain reaction is controlled so that the number of fissions per unit time is constant, rather than increasing, the rate of release of energy can be controlled. This situation is illustrated in Figure 26.5.

These conditions apply in the reactor of a modern nuclear power station, where some of the neutrons released in fission reactions are absorbed by control rods in order to limit the rate of fission reactions.

Two of the neutrons produced by fission are absorbed by control rods. The third neutron induces further nuclear fission.

www.iranino.co

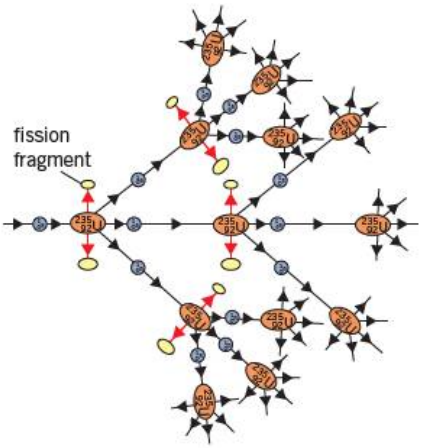


Figure 26.4 Accelerating chain reaction

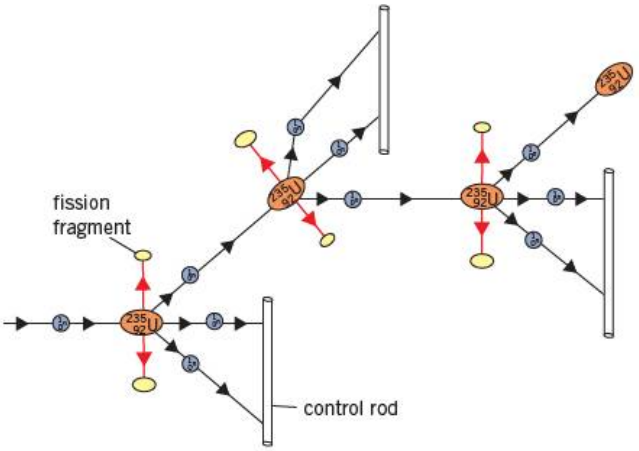


Figure 26.5 Controlled chain reaction.

26.4 The spontaneous and random nature of radioactive decay

Some elements have nuclei which are unstable. In order to become more stable, they emit particles and/or electromagnetic radiation (see AS Level Topic 26). The nuclei are said to be **radioactive**.

Detection of the count-rate of radioactive sources shows that the emission of radiation is both spontaneous and random.

Radioactive decay is a spontaneous process because it is not affected by any external factors, such as temperature or pressure.

Decay is random in that it is not possible to predict which nucleus in a sample will decay next.

Radioactive decay is a random process in that it cannot be predicted which nucleus will decay next. There is a constant probability that a nucleus will decay in any fixed period of time.

Random decay

We now look at some of the consequences of the random nature of radioactive decay. If six dice are thrown simultaneously (Figure 26.6), it is likely that one of them will show a six.

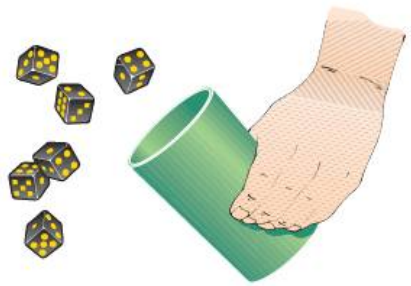


Figure 26.6 The dice experiment

If 12 dice are thrown, it is likely that two of them will show a six, and so on. While it is possible to predict the likely number of sixes that will be thrown, it is impossible to say which of the dice will actually show a six. We describe this situation by saying that the throwing of a six is a **random process**.

www.iranino.co

IRANINOVIN

In an experiment similar to the one just described, some students throw a large number of dice (say 6000). Each time a six is thrown, that die is removed. The results for the number of dice remaining after each throw are shown in Table 26.1.

Table 26.1 Results of dice-throwing experiment

number of throws	number of dice remaining	number of dice removed
0	6000	
1	5000	1000
2	4173	827
3	3477	696
4	2897	580
5	2414	483
6	2012	402
7	1677	335
8	1397	280

Figure 26.7 is a graph of the number of dice remaining against the number of throws.

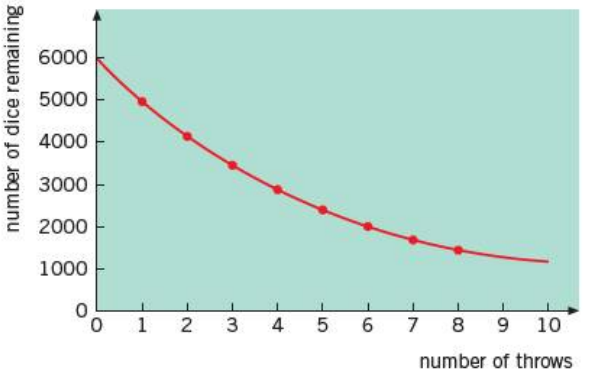


Figure 26.7

This kind of graph is called a **decay curve**. The rate at which dice are removed is not linear, but there is a pattern. After between 3 and 4 throws, the number of dice remaining has halved. Reading values from the graph shows that approximately 3.8 throws would be required to halve the number of dice. After another 3.8 throws the number has halved again, and so on.

We can apply the dice experiment to model radioactive decay. The 6000 dice represent radioactive nuclei. To score a six represents radioactive emission. All dice scoring six are removed, because once a nucleus has undergone radioactive decay, it is no longer available for further decay. Thus, we can describe how rapidly a sample of radioactive material will decay.

A graph of the number of undecayed nuclei in a sample against time has the typical decay curve shape shown in Figure 26.8.

It is not possible to state how long the entire sample will take to decay (its 'life'). However, after 3 minutes the number of undecayed nuclei in the sample has halved. After a further 3 minutes, the number of undecayed atoms has halved again. We describe this situation by saying that this radioactive isotope has a **half-life** $t_{\frac{1}{2}}$ of 3 minutes.

The half-life of a radioactive nuclide is the time taken for the number of undecayed nuclei to be reduced to half its original number.

The half-lives of different isotopes have a very wide range of values. Examples of some radioactive isotopes and their half-lives are given in Table 26.2.

www.iranino.co

IRANINOVIN

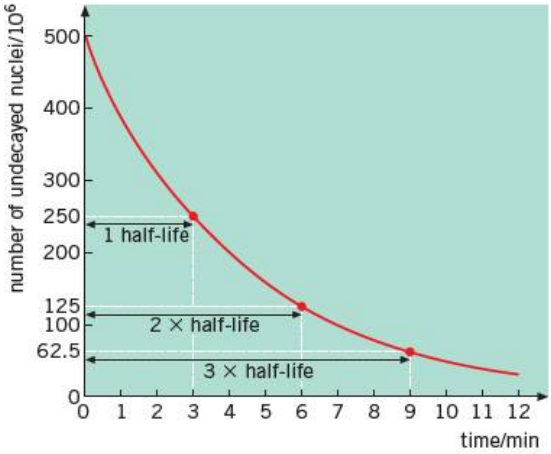


Figure 26.8 Radioactive decay curve

Table 26.2 Examples of half-life

radioactive isotope	half-life
uranium-238	4.5×10^9 years
radium-226	1.6×10^3 years
radon-222	3.8 days
francium-221	4.8 minutes
astatine-217	0.03 seconds

We shall see later that half-life may also be expressed in terms of the activity of the material.

If you measure the count rate from an isotope with a very long half-life, you might expect to obtain a constant value. In fact, the count rate fluctuates about an average value. This demonstrates the random nature of radioactive decay.

Example

The half-life of francium-221 is 4.8 minutes. Calculate the fraction of a sample of francium-221 remaining undecayed after a time of 14.4 minutes.

The half-life of francium-221 is 4.8 min, so after 4.8 min half of the sample will remain undecayed. After two half-lives (9.6 min), $0.5 \times 0.5 = 0.25$ of the sample will remain undecayed. After three half-lives (14.4 min), $0.5 \times 0.25 = 0.125$ will remain undecayed. So the fraction remaining undecayed is **0.125** or **1/8**.

Now it's your turn

4 Using the half-life values given in Table 26.2, calculate:
- (a) the fraction of a sample of uranium-238 remaining undecayed after 9.0×10^9 years,
- (b) the fraction of a sample of astatine-217 remaining undecayed after 0.30 s,
- (c) the fraction of a sample of radium-226 that has decayed after 3200 years,
- (d) the fraction of a sample of radon-222 that has decayed after 15.2 days.

In carrying out experiments with radioactive sources, it is important to take account of background radiation. In order to determine the count-rate due to the radioactive source, the background count-rate must be subtracted from the total measured count-rate. Allowance for background radiation gives the corrected count-rate.

Mathematical descriptions of radioactive decay

Activity and decay constant

As we saw in the dice experiment, increasing the number of dice increases the number of sixes that appear with each throw. Similarly, if we investigate the decay of a sample of radioactive material, we find that the greater the number of radioactive nuclei in the sample the greater the rate of decay.

www.iranino.co

IRANINOVIN

If there are N nuclei in a sample at time t, then at time $(t + \mathrm{d}t)$ some nuclei will have decayed so that the number remaining is $(N - \mathrm{d}N)$.

For radioactive decay, the probability of decay per unit time is constant, known as the **decay constant** λ.

For radioactive decay, the decay constant λ is the probability per unit time of the decay of a nucleus.

In this case,

probability of decay = $\mathrm{d}N/N$

and

probability of decay per unit time = $\lambda = -(\mathrm{d}N/\mathrm{d}t)N$.

This gives

$$\frac{\mathrm{d}N}{\mathrm{d}t} = -\lambda N$$

Note that a negative sign has been included. This is because as time t increases, the number N decreases.

$\mathrm{d}N/\mathrm{d}t$ is the rate at which the number of nuclei in the sample is changing, so $\mathrm{d}N/\mathrm{d}t$ represents the rate of decay. $-\mathrm{d}N/\mathrm{d}t$ is known as the **activity** A of the source, and is measured in **becquerels**.

The activity of a radioactive source is the number of nuclear decays occurring per unit time in the source. Activity is measured in becquerels (Bq), where 1 becquerel is 1 decay per second.

$1\,\mathrm{Bq} = 1\,\mathrm{s}^{-1}$

Combining $A = -\mathrm{d}N/\mathrm{d}t$ and $\mathrm{d}N/\mathrm{d}t = -\lambda N$, we have

$A = \lambda N$

The decay constant λ has the units s^{-1}, yr^{-1} and so on.

This is an important equation because it relates a quantity we can measure ($\mathrm{d}N/\mathrm{d}t$, the rate of decay or activity) to a quantity which cannot, in practice, be determined (N, the number of undecayed nuclei). We shall see later that the decay constant λ is directly related to the half-life $t_{\frac{1}{2}}$ and this can be obtained by experiment. This opens the way to calculating the number of undecayed nuclei in a sample. Trying to count nuclei when they are decaying is similar to counting sheep in a field while some are escaping through a gap in the hedge!

Example

Calculate the number of phosphorus-32 nuclei in a sample which has an activity of 5.0×10^6 Bq. (decay constant of phosphorus-32 = $5.6 \times 10^{-7}\,\mathrm{s}^{-1}$)

From $\mathrm{d}N/\mathrm{d}t = -\lambda N$, $N = (-\mathrm{d}N/\mathrm{d}t)/\lambda = -5.0 \times 10^6/5.6 \times 10^{-7} = -8.9 \times 10^{12}$.

The minus sign in this answer arises because $\mathrm{d}N/\mathrm{d}t$ is the rate of decay. The quantity measured by a ratemeter is the rate of decay, and so should be negative, but it is always displayed as a positive quantity. Similarly, activities are always quoted as positive. So don't be worried about discarding the minus sign here! The number of phosphorus-32 nuclei = $\mathbf{8.9 \times 10^{12}}$.

Note: In this type of calculation, because the activity is measured in becquerel (s^{-1}), the decay constant λ must be measured in consistent units. If λ had been quoted as $4.8 \times 10^{-2}\,\mathrm{day}^{-1}$, it would have been necessary to convert to s^{-1}.

Now it's your turn

5 Calculate the activity of the following samples of radioactive materials:

(a) 6.7×10^{21} atoms of strontium-90,
(decay constant of strontium-90 = $8.3 \times 10^{-10}\,\mathrm{s}^{-1}$)

(b) 2.0 mg of uranium-238. 0.238 kg of uranium-238 contains 6.0×10^{23} atoms.
(decay constant of uranium-238 = $5.0 \times 10^{-13}\,\mathrm{s}^{-1}$)

www.iranino.co

To solve the equation $dN/dt = -\lambda N$ requires mathematics beyond the scope of A/AS Physics. However, it is important to know the solution, in order to find the variation with time of the number of nuclei remaining in the sample. The solution is

$$N = N_0 e^{-\lambda t} \text{ or } N = N_0 \exp(-\lambda t)$$

where N_0 is the initial number of undecayed nuclei in the sample, and N is the number of undecayed nuclei at time t.

The equation represents an exponential decay. The decay curve of N against t is as shown in Figure 26.8. Since activity A is proportional to N ($A = \lambda N$), the curve of A against t is the same shape, and we can write

$$A = A_0 e^{-\lambda t}$$

Example

A sample of phosphorus-32 contains 8.6×10^{12} nuclei at time $t = 0$. The decay constant of phosphorus-32 is 4.8×10^{-2} day^{-1}. Calculate the number of undecayed phosphorus-32 nuclei in the sample after 10 days.

From $N = N_0 e^{-\lambda t}$, we have $N = 8.6 \times 10^{12} \times e^{-0.048 \times 10}$, so

$N = \mathbf{5.3 \times 10^{12}}$.

(Again, it is important to measure λ and t in consistent units. Here λ is in day^{-1} and t is in days, so there is no problem.)

Now it's your turn

6 A sample of phosphorus-32 contains 8.6×10^{12} nuclei at time $t = 0$. The decay constant of phosphorus-32 is 4.8×10^{-2} day^{-1}. Calculate the number of undecayed phosphorus-32 nuclei in the sample after
(a) 20 days,
(b) 40 days.

Decay constant and half-life

Using the equation $N = N_0 e^{-\lambda t}$, we can derive an equation which relates the half-life to the decay constant. For any radioactive isotope, the number of undecayed nuclei after one half-life is, by the definition of half-life, equal to $\frac{1}{2}N_0$, where N_0 is the original number of undecayed nuclei. Using the radioactive decay equation

$$N = N_0 e^{-\lambda t}$$

we have, at time $t = t_{\frac{1}{2}}$

$$\tfrac{1}{2}N_0 = N_0 \exp(-\lambda t_{\frac{1}{2}})$$

and, dividing each side of the equation by N_0,

$$\tfrac{1}{2} = \exp(-\lambda t_{\frac{1}{2}})$$

or

$$2 = \exp(\lambda t_{\frac{1}{2}})$$

Taking natural logarithms of both sides,

$$\ln 2 = \lambda t_{\frac{1}{2}}$$

So that

$$t_{\frac{1}{2}} = \ln 2/\lambda$$

or

$$t_{\frac{1}{2}} = 0.693/\lambda$$

www.iranino.co

IRANINOVIN

Example

Calculate the half-life of radium-226, which has a decay constant of $1.42 \times 10^{-11}\ s^{-1}$.

Using $t_{½} = 0.693/\lambda$, we have $t_{½} = 0.693/1.42 \times 10^{-11} = \mathbf{4.88 \times 10^{10}\ s}$

Now it's your turn

7 Calculate the half-lives of the following radioactive nuclides:
   (a) bismuth-214, which has a decay constant of $4.3 \times 10^{3}\ s^{-1}$,
   (b) carbon-14, which has a decay constant of $4.1 \times 10^{-12}\ s^{-1}$.

8 Calculate the decay constants of the following radioactive nuclides:
   (a) helium-5, which has a half-life of 6.0×10^{-20} s,
   (b) sodium-24, which has a half-life of 15 h.

Summary

- The mass defect of a nucleus is the difference between the total mass of the separate nucleons and the mass of the nucleus.
- Einstein's mass–energy equivalence relation $E = mc^2$
- The binding energy of a nucleus is the energy needed to separate completely all its constituent nucleons.
- The binding energy per nucleon is a measure of the stability of a nucleus. A high binding energy per nucleon means the nucleus is stable.
- Nuclear fusion is the joining together of light nuclei to form a larger, heavier nucleus.
- Nuclear fission is the splitting of a heavy nucleus into two smaller, lighter nuclei of approximately equal mass.
- Radioactive decay is a spontaneous, random process.
- The half-life $t_{½}$ of a radioactive nuclide is the time taken for the number of undecayed nuclei to be reduced to half the original number.
- The activity of a radioactive source is the number of nuclei that decay per unit time. The unit of activity is the becquerel (Bq). 1 becquerel = $1\ s^{-1}$.
- The activity $-dN/dt$ of a source is related to the number N of undecayed nuclei by the equation
  $dN/dt = -\lambda N$
  where λ is the decay constant.
- The decay constant is defined as the probability of decay per unit time of a nucleus.
- The number N of undecayed nuclei in a radioactive sample at time t is given by the equation
  $N = N_0\, e^{-\lambda t}$
  where N_0 is the number of undecayed nuclei at time $t = 0$.
- The half-life $t_{½}$ and the decay constant λ are related by the equation $\ln 2 = \lambda t_{½}$ or $t_{½} = 0.693/\lambda$.

Examination style questions

1 The table below shows the variation with time t of the activity of a sample of a radioactive nuclide X. The average background count during the experiment was $36\ min^{-1}$.

t/hour	0	1	2	3	4	5	6	7	8	9	10
activity/ min^{-1}	854	752	688	576	544	486	448	396	362	334	284

   a Plot a graph to show the variation with time of the corrected count-rate.

   b Use the graph to determine the half-life of the nuclide X.

2 Calculate the mass defect, the binding energy of the nucleus, and the binding energy per nucleon for the following nuclei:

   a hydrogen-3, $^{3}_{1}H$; nuclear mass 3.01605 u,

   b zirconium-97, $^{97}_{40}Zr$; nuclear mass 97.09801 u,

   c radon-222, $^{222}_{86}Rn$; nuclear mass 222.01754 u.

   (proton mass = 1.00728 u; neutron mass = 1.00867 u)

www.iranino.co

IRANINOVIN

3 One possible reaction taking place in the core of a reactor is

$^{235}_{92}U + {}^{1}_{0}n \rightarrow {}^{95}_{42}Mo + {}^{139}_{57}La + 2{}^{1}_{0}n + 7{}^{0}_{-1}e$

For this reaction, calculate:

a the mass on each side of the equation,

b the change in mass after fission has taken place,

c the energy released per fission of uranium-235,

d the energy available from the complete fission of 1.00 g of uranium-235,

e the mass of uranium-235 used by a 500 MW nuclear power station in one hour, assuming 30% efficiency.

(masses: $^{235}_{92}U$, 235.123 u; $^{95}_{42}Mo$, 94.945 u; $^{139}_{57}La$, 138.955 u; proton, 1.007 u; neutron, 1.009 u. 0.235 kg of uranium-235 contains 6.0×10^{23} atoms.)

4 Two fusion reactions which take place in the Sun are described below.

a A hydrogen-2 (deuterium) nucleus absorbs a proton to form a helium-3 nucleus.

b Two helium-3 nuclei fuse to form a helium-4 nucleus plus two free protons.

For each reaction, write down the appropriate nuclear equation and calculate the energy released.

(masses: $^{2}_{1}H$, 2.01410 u; $^{3}_{2}He$, 3.01605 u; $^{4}_{2}He$, 4.00260 u; $^{1}_{1}p$, 1.00728 u; $^{1}_{0}n$, 1.00867 u)

5 Calculate the mass of caesium-137 that has an activity of 4.0×10^{5} Bq.

The number of atoms in 0.137 kg of caesium-137 is 6.0×10^{23} atoms. The half-life of caesium-137 is 30 years.

6 **a** The activity of a radioactive source X falls from 5.0×10^{10} Bq to 1.0×10^{10} Bq in 5.0 hours. Calculate the half-life.

b The activity of a certain mass of carbon-14 is 5.00×10^{9} Bq. The half-life of carbon-14 is 5570 years.

Calculate the number of carbon-14 nuclei in the sample.

7 The half-life of a radioactive isotope of sodium used in medicine is 15 hours.

a Determine the decay constant for this nuclide.

b A small volume of a solution containing this nuclide has an activity of 1.2×10^{4} disintegrations per minute when it is injected into the bloodstream of a patient. After 30 hours, the activity of 1.0 cm^3 of blood taken from the patient is 0.50 disintegrations per minute. Estimate the volume of blood in the patient. Assume that the solution is uniformly diluted in the blood, that it is not taken up by the body tissues, and that there is no loss by excretion.

8 **a** State what is meant by the *binding energy* of a nucleus. *[2]*

b Show that the energy equivalence of 1.0 u is 930 MeV. *[3]*

c Data for the masses of some particles and nuclei are given in Table 26.3.

Table 26.3

	mass/u
proton	1.0073
neutron	1.0087
deuterium ($^{2}_{1}H$)	2.0141
zirconium ($^{97}_{40}Zr$)	97.0980

Use data from Table 26.3 and information from **b** to determine, in MeV,

i the binding energy of deuterium, *[2]*

ii the binding energy **per nucleon** of zirconium. *[3]*

Cambridge International AS and A Level Physics, 9702/42 May/June 2011 Q 8

9 **a** Explain why the mass of an α-particle is less than the total mass of two individual protons and two individual neutrons. *[2]*

b An equation for one possible nuclear reaction is

$^{4}_{2}He + {}^{14}_{7}N \rightarrow {}^{17}_{8}O + {}^{1}_{1}p.$

Data for the masses of the nuclei are given in Table 26.4.

Table 26.4

	mass/u
proton $^{1}_{1}p$	1.00728
helium-4 $^{4}_{2}He$	4.00260
nitrogen-14 $^{14}_{7}N$	14.00307
oxygen-17 $^{17}_{8}O$	16.99913

i Calculate the mass change, in u, associated with this reaction. *[2]*

ii Calculate the energy, in J, associated with the mass change in **i**. *[2]*

iii Suggest and explain why, for this reaction to occur, the helium-4 nucleus must have a minimum speed. *[2]*

Cambridge International AS and A Level Physics, 9702/42 May/June 2013 Q 8

10 **a** State what is meant by a *nuclear fusion reaction*. *[2]*

b One nuclear reaction that takes place in the core of the Sun is represented by the equation

$^{2}_{1}H + {}^{1}_{1}H \rightarrow {}^{3}_{2}He + \text{energy}.$

Data for the nuclei are given in Table 26.5.

Table 26.5

	mass/u
proton $^{1}_{1}H$	1.00728
deuterium $^{2}_{1}H$	2.01410
helium $^{3}_{2}He$	3.01605

i Calculate the energy, in joules, released in this reaction. *[3]*

ii The temperature in the core of the Sun is approximately 1.6×10^{7} K.

Suggest why such a high temperature is necessary for this reaction to take place. *[2]*

Cambridge International AS and A Level Physics, 9702/43 May/June 2013 Q 8

www.iranino.co

IRANINOVIN

Answers

Topic 1

Now it's your turn

1 $10\,800\,cm^2$ or $1.08 \times 10^4\,cm^2$
2 1.0×10^9
3 (a) $6.8 \times 10^{-12}\,F$
(b) $3.2 \times 10^{-5}\,C$
(c) $6.0 \times 10^{10}\,W$
4 800
5 4.6×10^4
6 $kg\ m^{-3}$
7 $kg\ m^{-1}\ s^{-2}$
8 (a) yes
(b) yes
9 $m^2\ s^{-2}$
10 (a) $kg\ m^2\ s^{-2}$
(b) $m^2\ s^{-2}\ K^{-1}$
11 $kg\ m^{-1}\ s^{-2}$
12 (a) 100–150 g
(b) 50–100 kg
(c) 2–3 m
(d) 0.5–1.0 cm
(e) $0.5\,cm^3$
(f) $4 \times 10^{-3}\,m^3$
(g) $220\,m\ s^{-1}$
(h) 310 K
13 (a) scalar
(b) vector
(c) scalar
14 (a) vector
(b) scalar
(c) vector
15 Velocity has direction, speed does not. Velocity is defined in terms of displacement which is a vector, speed is defined in terms of distance which is a scalar.
16 Student is correct. Weight is a force which acts vertically downwards.
17 Direction of arrow gives direction of vector. Length of arrow drawn to scale represents magnitude of vector quantity.
18 (a) 690 N
(b) 210 N
(c) 510 N at an angle of 28° to the 450 N force
19 upstream at 78° to the bank
20 120 N at an angle of 25° to the 50 N force in an anticlockwise direction
21 (a) $7\,km\ h^{-1}$
(b) $1\,km\ h^{-1}$
22 11 N at an angle to the 6.0 N force of 56° in an anticlockwise direction
23 (a) $250\,km\ h^{-1}$
(b) $180\,km\ h^{-1}$
24 (a) $1.0\,m\ s^{-1}$
(b) $9.1\,m\ s^{-1}$

Examination style questions

2 $kg\ m^2$
3 (a) (i) $kg\ m^2\ s^{-2}$
(ii) $kg\ m^2\ s^{-2}$
4 (b) (i) 92 N
(ii) 77 N
(c) (i) 59 N
(ii) 59 N
5 (a) (i) $18\,m\ s^{-1}$
(ii) 29° above horizontal
(b) (i) $10\,m\ s^{-1}$
(ii) 33°
6 (a) 380 pm
(b) 0.086 Ms
(c) 8.3 min
(e) (ii) $230\,m\ s^{-1}$
7 (a) (i) m^3
(b) $kg\ m^{-1}\ s^{-1}$
8 (a) 20 Hz → 20 kHz
(b) 10 nm → 400 nm
(c) 10 g → 100 g
(d) $0.1\,kg\,m^{-3}$ → $10\,kg\,m^{-3}$
9 (a) length, current, temperature
(b) (i) $\boldsymbol{F}$ kg ms^{-2} $\boldsymbol{\rho}$ kg m^{-3} $\boldsymbol{v}$ ms^{-1}
(ii) 2

Topic 2

Now it's your turn

1 12.52 mm
2 15 ± 1 g
3 10
4 38.8 ± 0.5 °C
5 (a) 28 mV
(b) 9.9 mV
7 (a) ± 0.06 A
(b) 2.01 ± 0.09 A

www.iranmo.com

8 **(a)** **(i)** micrometer screw gauge
(ii) ohm-meter or voltmeter and ammeter
(iii) cathode ray oscilloscope or a.c. voltmeter
(b) $2.7 \pm 0.2\,g\,cm^{-3}$
9 **(a)** **(i)** 1.6%
(ii) 1.1%
(b) 4.3%
(c) **(ii)** $9.8 \pm 0.4\,ms^{-2}$
10 **(a)** end of rule damaged; measure length from 10 cm mark
(b) parallax error; place pencil in contact with scale, over the graduations and define the ends with set-squares
11 **(a)** micrometer screw gauge
(b) zero error on drum
(c) averaging reduces random errors; spiral readings allow for a non-circular cross-section, and moving along the length of the wire allows for any taper
12 **(a)** 1.2 **(b)** ± 0.1 cm
13 $(870 \pm 40)\,cm^3$

Examination style questions

7 **(a)** micrometer screw gauge
8 **(b)** $2.7 \pm 0.2\,g\,cm^{-3}$
9 **(a)** **(i)** 1.6%
(ii) 1.1%
(b) 4.3%
(c) **(ii)** $9.8 \pm 0.4\,m\,s^{-2}$

Topic 3

Now it's your turn

1 5.3×10^{-11} m
2 3200 s
3 $6\,m\,s^{-1}$
4 $6\,m\,s^{-2}$
5 3.3 s
6 30 km
7 $180\,m\,s^{-1}$
9 3.6 h 610 km h^{-1}
10 $-5.0\,m\,s^{-2}$
12 $2.5\,m\,s^{-2}$
13 7.5 s 15 m s^{-1}
14 8.2 m s^{-1} upwards
15 **(a)** $1.8\,m\,s^{-2}$
(b) 770 m
16 $6.1\,m\,s^{-1}$
17 20 m
18 75 m
19 $140\,m\,s^{-1}$
20 $47\,m\,s^{-1}$
21 **(a)** B
(b) A
22 3.5°

Examination style questions

2 8
3 **(a)** 61° 2.8 m
(b) 3.9 m s^{-1}
(c) 1.4 s
4 **(a)** $9.4\,m\,s^{-1}$
(b) 10%
6 **(a)** **(ii)** 10.3 m
(iii) 11.9 m
7 **(b)** **(i)** 10.6
(ii) $0.21\,m\,s^{-2}$
8 **(a)** **(i)** $14.6\,m\,s^{-1}$
(ii) 2.0 s
9 **(b)** **(ii)** $30\,m\,s^{-1}$
(iv) 10 800 N

Topic 4

Now it's your turn

1 $1.7\,m\,s^{-2}$
2 5.3 N
3 **(a)** 61 N
(b) 94 N
4 6.8×10^{-24} Ns (kg m s^{-1})
5 $1.2\,m\,s^{-1}$
6 600 N
7 **(a)** $\frac{1}{2}u_A$
(b) 0.5

Examination style questions

1 50 kg
2 **(a)** $7.7\,m\,s^{-1}$
(b) $45.3\,m\,s^{-2}$
(c) 2040 N
4 $T/2$
5 3.3×10^4 Ns (kg m s^{-1})
6 3.6×10^7 N
7 0.27 N
9 $1.03 \times 10^5\,m\,s^{-1}$
10 Heavy particle's speed is practically unchanged; light particle moves with a speed $2u$ in the same direction as the incident heavy particle.
11 the lighter body
12 **(a)** 0.54 kg m s^{-1}
(b) 0.54 Ns
(c) 36 N
13 **(a)** **(i)** 2.16 kg m s^{-1}
(ii) $1.07\,m\,s^{-1}$
(iii) 1.16 J
(b) 0.059 m
14 $3080\,m\,s^{-1}$
16 **(a)** $3u$ to the left
(b) **(i)** $3u$ to the right
(ii) $3t_1$
(c) $3u/2$
17 **(a)** $3m$
(b) 0.25
18 **(b)** **(i)** $F_A = -F_B$; $t_A = t_B$
19 **(a)** **(i)** $7.5\,m\,s^{-1}$
(ii) $13\,m\,s^{-1}$
(b) **(i)** 8.6 m
(c) **(i)** (–)0.73 N s

www.iranino.co

IRANINOVIN

20 **(a)** **(i)** $4.0\,m\,s^{-1}$
(b) $7.4\,m\,s^{-1}$ at 33° to the vertical
(c) **(i)** $4.4\,m\,s^{-1}$
(ii) **1** $0.36\,kg\,m\,s^{-1}$ **2** 3.0 N

Topic 5

Now it's your turn

1 5.6 N m
2 4.5 N
3 14 (13.5) N
4 **(b)** 29 N
5 1.4×10^4 Pa

Examination style questions

1 **(a)** 4.2 N m
(b) 9.1 N
2 **(a)** 2.9 N m
(b) 8.0 N
3 **(a)** 48 N
(b) 72 N
4 67 N
5 1.5×10^5 Pa
7 **(b)** **(i)** 23 N
(iv) 1.9 kg
8 **(c)** **(ii)** 660 N
9 **(b)** **(i)** 250 N
10 **(b)** **(i)** 12 N m
(c) **(i)** 0.90 N
(ii) 1.5 N
11 **(c)** **(i)** $2500\,kg\,m^{-3}$
(ii) 2.2×10^4 Pa

Topic 6

Now it's your turn

1 **(a)** 180 J
(b) 30 J
2 **(a)** 15.5 J
(b) 7.6 J
3 1800 J
4 75 J
5 **(a)** kinetic energy at its lowest point → potential energy at its highest point → kinetic energy at its lowest point, etc (kinetic and potential energy at points in between)
(b) potential energy of compressed gas → kinetic energy of spray droplets → heat when droplets have stopped moving(energy transferred to surrounding atmosphere)
(c) kinetic energy when thrown → potential and kinetic energy at highest point of motion → internal and heat and sound energy when clay hits the ground
6 3.1×10^5 J
7 **(a)** 1000 J
(b) 3000 J
8 **(a)** 1.3×10^4 J gain
(b) 12 J gain
(c) 2.4×10^9 J loss

www.iranmo.co

9 All the electrical energy is changed to internal energy of the heater and the heater gets hot. The heater releases all this energy to the surroundings so the process is 100% efficient.
10 58%
11 4.3×10^5 J
12 1.2×10^5 N
13 1.5 kW
14 6.7 cents
15 0.96 cents
16 5.3×10^{10}

Examination style questions

1 90 J; 0; 78 J; 0.15 m; 36°; 2.6×10^3 N
2 **(a)** 7.9 N
(b) 0.19 J
3 87 kJ
4 **(a)** chemical
(b) nuclear
(c) gravitational potential
(d) kinetic of the wind
(e) kinetic of gases, sound, light, potential energy
(f) internal potential
5 **(a)** 0.51 m
(b) $180\,J\,s^{-1}$
6 **(a)** 11 J
(b) $14\,m\,s^{-1}$
7 **(a)** 16 MJ
(b) 110 MJ
(c) 210 MJ
8 15%
10 2100 cents
11 **(a)** **(i)** 16.5 kW
(ii) 3200 N
(iii) 96 kW
(b) **(i)** 1.32 MJ
(ii) 3.84 MJ
12 **(b)** **(i)** **1** 8.3 J **2** 13 m
(ii) 3
13 **(b)** 1090 N
(c) **(ii)** 76 m
(iii) **1** 2.7×10^5 J **2** 3.5×10^5 J
14 **(c)** **(ii)** $3.8 \times 10^3\,kg\,s^{-1}$
15 **(a)** **(ii)** 17.6 m
(b) **(i)** 140 N up
(ii) 17.6 m

Topic 9

Now it's your turn

2 $400\,N\,m^{-1}$
3 **(a)** $3.0 \times 10^4\,N\,m^{-1}$
(b) 54 N
4 5.4×10^{-2} J
5 4.2×10^{-2} m
6 **(a)** 1.29×10^7 Pa
(b) 1.17×10^{-4}
(c) 0.16 mm
7 5.3×10^6 Pa

Examination style questions

1 **(a)** 500 N m^{-1}
(b) 13.1 cm
2 7.9 J
3 0.36 mm
5 **(c)** **(ii)** **1** 0.057 J **3** 0.020 J
6 **(b)** **(i)** 4.6×10^{-3} m
(ii) 4×
7 **(b)** **(ii)** 78 N m^{-1}
(iii) 0.26 J
(c) **(i)** 0.12 m
(ii) 53 N m^{-1}

Topic 14

Now it's your turn

1 0.40 m s^{-1}
2 0.68 m
3 1.4
4 **(a)** 1.2×10^{-3} s
(b) 830 Hz
5 270 Hz
6 4.6×10^{14} Hz
7 6.3
8 0.038 m

Examination style questions

1 **(a)** 200 Hz
(b) 5.0×10^{-3} s
2 **(a)** 7.5×10^{14} Hz to 4.3×10^{14} Hz
(b) 1.2 m
3 ¼ (0.25)
4 **(a)** 5.7×10^{-8} J
(b) 4.9×10^{-5} W m^{-2}
5 2.0×10^{9} W m^{-2}
6 **(b)** **(i)** 4 mm
(ii) 0.12 m s^{-1}
(iii) 180° or π rad
7 **(b)** **(i)** 0.6 s
(ii) 4.0 cm
(iii) 6.7 cm s^{-1}
(c) **(ii)** 3.0
8 **(b)** **(ii)** 6.0×10^{14} Hz
9 **(b)** **(i)** 2.8 cm; 135°
(ii) 4.0 cm
10 **(a)** **(i)** 7.6 mm
(ii) 180°
(iii) 12 m s^{-1}

Topic 15

Now it's your turn

1 6.25×10^{-7} m (625 nm)
2 63 m
3 880 Hz; 1320 Hz
4 0.38 m
5 340 Hz
6 128 Hz
7 56°
8 22 cm
9 652 nm
10 28.2°; 70.7°
11 2 (3 if zero order included)

Examination style questions

2 **(c)** 3.5 mm
5 **(b)** 8.5 cm
6 **(a)** 108 m s^{-1}
(b) 45 Hz
7 **(a)** 2.5×10^{-6} m
(b) 10.2°
(c) 5; 3
9 0.73° (1.26×10^{-2} rad)
10 **(b)** 180° (π rad)
(c) 5:1
(d) 224 m s^{-1}
11 26.6 cm
13 **(b)** 212 mm
(c) 449 mm
14 **(c)** 180 Hz
15 **(c)** 2
16 **(b)** 2.1×10^{-3} m
18 **(a)** **(i)** 0.8 m
(b) **(ii)** 1.2 m

Topic 17

Now it's your turn

3 **(a)** right
(b) left
(c) left
(d) right
4 **(a)** 5.0×10^{4} V m^{-1}
(b) 2.4×10^{-12} C
5 **(a)** The sphere moves downwards
(b) 1.6 m s^{-2}
6 2.3×10^{-3} m, up

Examination style questions

1 1.03×10^{6} m s^{-1}
2 1.46×10^{-9} s
3 **(a)** 1.92×10^{-16} J
(b) 1.92×10^{-16} J
(c) 2.05×10^{7} m s^{-1}
4 **(a)** **(ii)** 3.0×10^{4} V m^{-1}
(b) **(ii)** 7.2×10^{-17} N m
5 **(a)** **(i)** 1.4×10^{4} N C^{-1} or V m^{-1}
(b) **(i)** 2.5×10^{15} m s^{-2}
(ii) 4.5×10^{-9} s
6 **(b)** 1.4×10^{4} V m^{-1}
(c) **(i)** 2.5×10^{15} m s^{-2} up
(ii) 2.4×10^{-9} s
(d) deflects 7 mm so emerges

www.iranino.co

IRANINOVIN

Topic 19

Now it's your turn

1 2.0 A
2 4.8×10^5 s
3 **(a)** 60 C
(b) 3.8×10^{20}
4 3.9×10^{28} m^{-3}
5 1.6×10^{-13} J
6 **(a)** 0.25 C
(b) 2.2 J
7 100 Ω
8 current in both is 0.42 A
9 **(a)** 9.2 A
(b) 26 Ω
10 18 m
11 0.97 mm

Examination style questions

1 **(a)** 760 C
(b) 1000 W
(c) 57 Ω
2 **(a)** **(i)** 0.20 A
(ii) 0.60 W
(b) 5400 J
3 3.5×10^6 J
4 2.0 kW
5 6.7 m
6 **(b)**

I / A	0.20	0.40	0.60	0.80	1.00	1.20	1.40
R / Ω	0.95	1.20	2.45	3.65	4.56	5.47	6.21

7 **(a)** 0.62 Ω
(b) 4.3×10^{-7} Ω m
8 **(a)** **(i)** 0.25 A
(b) 0.14 m
9 **(b)** 15 Ω
(c) **(i)** Repeat Fig 19.13 completed with the following values:

top row:	1.5	15	30
bottom row:	3.0	20	10

(d) **(i)** 3
10 **(a)** **(i)** 19 Ω
(ii) 3.5 m
(b) **(i)** 630 W
(ii) ¼ length or 4 times greater area
11 **(b)** **(i)** 3200 W
(ii) 13 A
(iii) 3.5×10^7 C
(iv) 8.4×10^{19}

Topic 20

Now it's your turn

1 **(a)** EI
(b) I^2R or EI
2 0.56 Ω
3 **(a)** 3.0 A, when the cell is short circuited
(b) 1.1 W, when load resistance equals the internal resistance
4 0.25 A
5 2.0 A towards the junction
6 4.3 Ω
7 25 Ω
8 **(a)** 12 V
(b) 0.57 V
(c) 5.0 kΩ
9 **(a)** 0.82 V
(b) 7.7 V

Examination style questions

1 **(a)** 0.05 Ω
(b) 0.3 Ω
2 **(a)** 0.25 A
(b) 1.6 Ω
(c) 12 J
4 **(a)** 169 Ω
(b) 13 Ω
5 **(a)** 5 Ω
(b) 3.0 A
6 **(a)** 25 Ω
7 **(a)** 4.5 V
(b) **(i)** 50 Ω
(ii) 0.09 A
(iii) 0.90 V
9 10 resistors each of resistance 12 kΩ, power rating 0.5 W connected in parallel
10 **(a)** 4 Ω
(b) 8 Ω
(c) 3 Ω
(d) 1.0 A
11 **(a)** **(i)** 1.6×10^{-2} Ω
(ii) 1.1×10^{-3} Ω
(iii) 27 W
(b) **(i)** 4.4 s
(ii) 4.4×10^{21}
(c) **(i)** 11.7 V
(ii) 307 W
12 **(a)** 1.02 V, 1.22 W
(b) **(ii)** 7.53 m
(iii) 1.41 W
13 **(b)** **(i)** 0.29 A
(ii) 1.03 V
(iii) 1.03 V
14 **(b)** 1.29 V
(c) 0.75 V
15 **(a)** **(i)** **1** 0.16 Ω **2** (14 – E) V
(ii) 7.3 V
(b) **(i)** 1.8×10^5 C
(ii) 2.5×10^6 J
(iii) 3.6×10^5 J
(c) 86%
16 **(c)** **(i)** 5.7 V
(ii) 1.7 Ω
(d) **(i)** 4.64 W
(ii) 51%

www.iraninovin.com

Topic 26

Now it's your turn

1 19, 40, 21

3 5.8×10^{-13} J

Examination style questions

3 ${}^{231}_{90}Y + {}^{4}_{2}He$, ${}^{231}_{91}Z$, ${}^{0}_{0}\gamma$

4 **(a)** 2.3×10^{7} m s^{-1}

(b) 2.7×10^{5} m s^{-1}

5 **(a)** ${}^{224}_{88}Ra \rightarrow {}^{220}_{86}Rn + {}^{4}_{2}He$

(c) 3.0×10^{5} m s^{-1}

6 **(a)** 3300 mm^{-1}

7 **(b)** **(i)** ${}^{4}_{2}\propto$ ${}^{1}_{1}\rho$ or ${}^{1}_{1}H$

(ii) **2** 7.3 x 10^{6} m s^{-1}

8 **(b)** **(i)** 1 to 3 $\times 10^{-10}$ m

(ii) 1 to 7 $\times 10^{-15}$ m

9 **(b)** **(ii)** **1** 36 **2** 3

10 **(b)** 3 and 0 for superscript values
2 and –1 for subscript values

(c) 4.5×10^{7} m s^{-1}

11 **(b)** **(i)** ${}^{236}_{92}U \rightarrow {}^{232}_{9}Th + {}^{4}_{2}He$

(ii) **1** correct position for U at $Z = 92$ and $N = 145$
2 correct position for Np relative to U at $Z + 1$ and $N - 1$

www.iranino.co

IRANINOVIN

Answers

Topic 7

Now it's your turn

1 42 m
2 (a) 0.22 m s^{-1}
(b) 0.63 rad s^{-1}
(c) 0.14 m s^{-2}
3 4.7 N
4 (a) 590 m
(b) 4200 N

Examination style questions

2 (a) 3.3 rad s^{-1}
(b) 0.53 revolutions per second
3 (a) 7.8 km s^{-1}
(b) 5.3×10^3 s
4 (a) (ii) 1 3 *mg* 2 4 *mg*
(b) (i) 6.4 rad s^{-1}
(ii) 4.6 m s^{-1}

Topic 8

Now it's your turn

1 4.6×10^{16} N
2 1.4 hours
3 25 N kg^{-1}
4 -9.4×10^9 J; 1.1×10^4 m s^{-1}

Examination style questions

1 5.5×10^3 kg m^{-3}
2 7.78×10^8 km
3 −0.50%; −0.25%
4 5.7×10^{-11} N
5 (b) (ii) 8.86×10^4 km
6 (c) (ii) 54 R_E
7 (b) (i) 6.46×10^{23} kg; 23.6 hours

Topic 10

Now it's your turn

1 3.3×10^{-9} m
2 1.29 kg m^{-3}
3 1.1 mol; 6.7×10^{23}; 2.1×10^{25} m^{-3}
4 2.5×10^{-5} m^3
5 8.7×10^{-6} m^3
6 1.0×10^{-20} J
7 3.1×10^3 m s^{-1}

www.iranino.co

Examination style questions

1 212 kPa
2 1.28
3 6100 m s^{-1}
4 (b) 0.72 mol
5 (c) 580 m s^{-1}

Topic 11

Now it's your turn

1 103.6 Ω
2 1.001×10^5 Pa

Examination style questions

1 −43.9 °C
2 133
3 16 °C

Topic 12

Now it's your turn

1 3×10^{-10} m
2 (a) 1400 J gain
(b) 41 J loss
3 4200 J kg^{-1} K^{-1}
4 (a) (i) 2.5×10^5 J
(ii) 4.6×10^4 J
(b) 0.84
5 1100 J K^{-1}
6 (a) (i) 16.5 kJ
(ii) 113 kJ
(b) 6.8
7 480 °C
8 41 g
9 200 J

Examination style questions

1 −20 °C
2 (a) 40 J K^{-1}
(b) 38 °C
3 (a) 22.6 g
4 (a) 2020 J kg^{-1} K^{-1}
(b) 91 s
5 0, 250 J
6 30 J increase
7 (b) 400 W

8 **(b)** **(i)** 0
(ii) 240 J
(iii) 240 J −360 J
720 J
−840 J −360 J

Topic 13

Now it's your turn

1 **(a)** $x = 0.10 \cos \pi t$
(b) 0.20 s
2 0.25 m
3 2.5 s
4 **(a)** 0.87 J
(b) 1.97 J
(c) 2.84 J

Examination style questions

1 **(a)** 220 Hz
(b) 1.4×10^3 rad s^{-1}
(c) 42 m s^{-1}
(d) 5.9×10^4 m s^{-2}
(e) 7.1 m s^{-1}
2 2.8 m s^{-1}
3 9.826 m s^{-2}
4 **(a)** 5.8 N m^{-1}
(b) 15 mm
(c) 0.58 s
(d) −8.2 mm
5 3.2
6 **(a)** $0.87x_0$
(b) $0.71x_0$
7 **(a)** 2.0 Hz
8 **(b)** **(i)** $2A\rho g\Delta x$
(ii) $2g\Delta x/L$
(iv) $1/2\pi \times \sqrt{(2g/L)}$
9 **(a)** 1.3 s
(b) 16 N
10 **(a)** 1.28×10^{14} Hz
11 **(a)** **(i)** 1.7 cm; 2.8 Hz
(c) 1.2 cm
12 **(b)** **(i)** 3.3 Hz
13 **(a)** 2.19 m s^{-2}
14 **(a)** **(ii)** 16 mm; 4.6 Hz
(b) **(i)** 13.4 m s^{-2}
(ii) 2.0 N

Topic 14.6

Now it's your turn

1 **(b)** 9.4×10^{-3}
2 0.22

Examination style questions

2 0.908*I* incident on fat–muscle boundary
0.904*I* transmitted through fat–muscle boundary
0.332*I* incident on muscle–bone boundary
0.21*I* transmitted into bone

www.iranino.co

3 **(a)** **(ii)** 1.7×10^3 kg m^{-3}
(b) **(ii)** 1.04×10^{-3}; 0.999

Topic 16

Now it's your turn

1 constant frequency 750 kHz, amplitude variation 3.5 V–4.0 V, frequency of variation of amplitude 3 kHz
2 **(a)** 1.0×10^5 Hz
(b) 1.0×10^4 Hz
4 **(a)** 25
(b) 10
5 **(a)** 42 dB
(b) 22 dB
6 **(a)** 1.4 dB km^{-1}
(b) 94 km

Examination style questions

7

type	wavelength / m	range
sky wave	20	worldwide
ground (surface) wave	> 100	up to 1000 km
space wave	< 10	line of sight

8 **(b)** **(i)** 330 m
(ii) 18 kHz
(iii) 9000 Hz
9 **(b)** **(i)** 0111
(ii) 0110
10 **(a)** 2.3×10^{-4} W
(b) 16 dB
(c) 0.19 dB km^{-1}

Topic 17

Now it's your turn

1 2.3×10^{-8} N
2 1.0×10^6 N C^{-1}
3 0.45 J

Examination style questions

1 **(a)** 2.1×10^{21} N C^{-1}
(b) 650 N
(c) 1.7×10^7 V
(d) 5.2×10^{-12} J
2 1.51 m to the left of the 2.4 μc charge
3 **(a)** 2.56×10^{-17} N
(c) 1.02×10^{-22} J
4 2.2×10^{39}
5 **(b)** **(ii)** 11.4 cm; 7.9 cm

Topic 18

Now it's your turn

1 1.0 mF
2 1.4×10^{-9} F

IRANINOVIN

3 (a) 9.4 mC
(b) (ii) 940 μF
(iii) 10 V
(iv) 4.7 mC
4 0.203 J
5 (a) 6.0×10^{-5} C
(b) 2.7×10^{-5} C
(c) 5.4 V

Examination style questions

1 (a) 1.1×10^{-10} F
(b) (i) 1.3×10^{-9} C
(ii) 2.4×10^{4} V m^{-1}
2 (a) 2.0 μF
(b) 1.2 μF
3 (a) (i) 4.1×10^{-4} J
(ii) 9.0×10^{-5} C
(b) (i) 6.0×10^{-5} A
(ii) 1.5 s
(iii) 0.77 s
4 (b) (i) 8 μF
(iii) 6.0 V; 36 μC
5 (b) (i) 1800 μF
(ii) ≈ 0.039 J

Topic 19.4 & 20.3

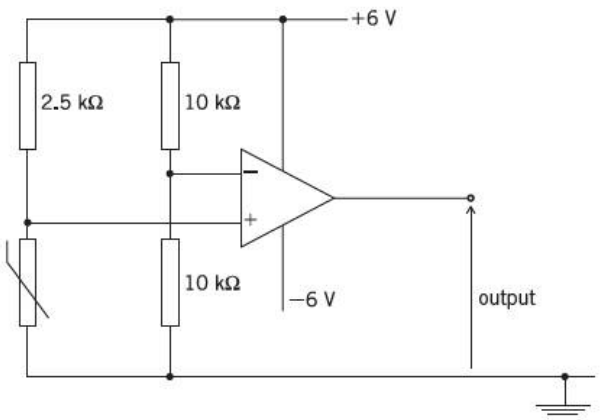
Now it's your turn

3 Connections are across variable resistor
$R = 750\,\Omega$

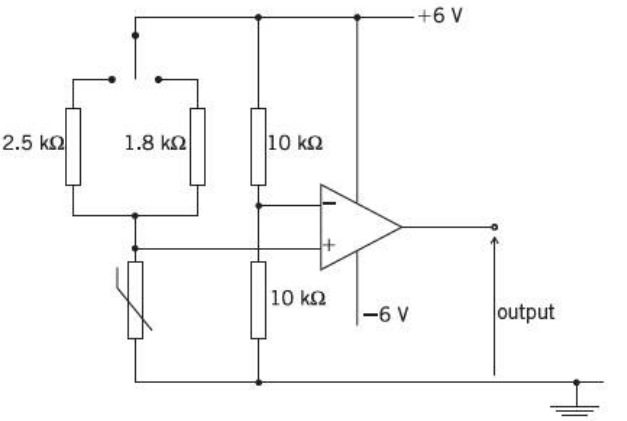
Topic 21

Now it's your turn

1



2



3 (b) 10
(c) 0.9 V

Examination style questions

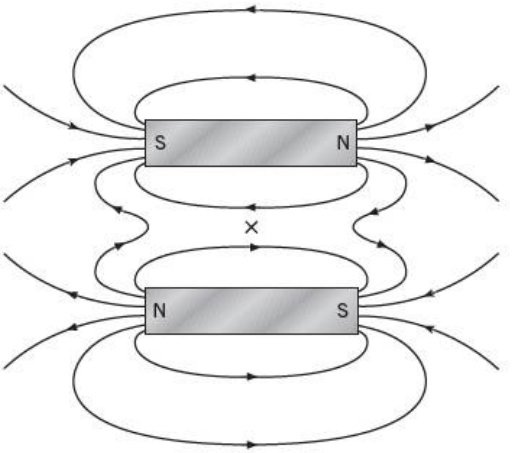
1 (c) (i) −19
(ii) 0.47 V
2 (a) 2.25 kΩ
3 (b) (i) 71
(ii) 1.7 kΩ
(c) ±9 V
4 (b) (ii) 4.0 V
(iii) 0.13 V
5 (b) (ii) 860 Ω

Topic 22

Now it's your turn

1 Diagram should be as Figure 22.4 but with the direction of field lines reversed.

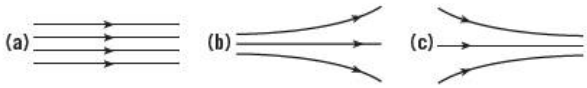
2



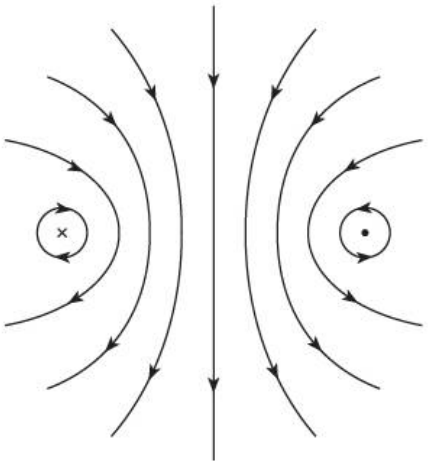
3 3.4×10^{-7} N
4 (a) 2.8×10^{-2} N m^{-1}
(b) 2.0×10^{-2} N m^{-1}
5 (a) 8.8×10^{6} m s^{-1}
(b) 9.3 cm
6 4.2×10^{6} m s^{-1}
7 0.52 μV

www.iranino.co

8



9



Examination style questions

1 **(a)** **(ii)** 2.5×10^{4} A
2 **(a)** **(i)** 1.5×10^{-4} T
(ii) 2.3×10^{-3} N m^{-1}
3 **(a)** 0.15 T
6 **(c)** **(ii)** 0.499 T
8 **(b)** **(ii)** 2.8×10^{-5} T
(iii) 2.6×10^{-4} N m^{-1}

Topic 23

Now it's your turn

1 **(a)** **(i)** 1600 m^2
(ii) 0.064 Wb
(iii) 0.064 V
2 0.043 V
3 **(a)** 2.7 m^2
(b) 0.41 Wb
(c) 0.41 V

Examination style questions

1 **(a)** **(i)** 1.4 T
(ii) 4.0 mWb
(iii) 1.6 Wb
(b) 150 V across the switch
2 **(a)** 0
(b) 50 mV
(c) 29 mV
3 **(a)** **(i)** 4.5 mT
(ii) 1.1×10^{-4} Wb
(b) 7.1 mV
4 $E = 0$ when current constant. Spikes in opposite directions at times t_{on} and t_{off}
5 **(b)** 1.3 Hz
6 **(b)** **(i)** 6.6×10^{-3} T
(ii) 3.0×10^{-3} Wb
(c) **(ii)** 7.4×10^{-3} V

Topic 24

Now it's your turn

1 **(a)** 1.4 kW
(b) 2.9 kW, 0 kW
2 **(a)** 45
(b) 56 mA
3 **(a)** 30
(b) 7500
(c) 450 mA

Examination style questions

1 20 V, 2.5 A
2 **(a)** 340 V
(b) 4.8 mA
4 **(a)** 0.034
(b) 4100
5 **(b)** **(ii)** 37.7
6 **(a)** **(i)** 60 Hz
(ii) 7.0 A
(b) 8.2 Ω

Topic 25

Now it's your turn

2 5.8×10^{14} Hz
3 **(a)** 1.4×10^{15} Hz
(b) sodium and zinc
4 1.3×10^{-11} m
5 1.2×10^{-10} m
6 2.6×10^{3} m s^{-1}
7 3.9×10^{-8} m
8 3.2×10^{15} Hz
9 9.1×10^{-8} m
13 2.8×10^{-11} m
15 24 cm
16 0.019
18

6	9
4	3

Examination style questions

2 8.0×10^{14} m^{-2}
3 **(a)** 495 nm
(b) **(i)** 3.9×10^{-19} J
(ii) 9.3×10^{5} m s^{-1}
4 **(a)** 4.4×10^{-35} m
(b) 2.6×10^{-35} m
(c) 2.4×10^{-11} m
(d) 1.3×10^{-13} m
7 **(b)** **(ii)** 3.5×10^{-19} J; 6.3×10^{-34} J s
9 **(b)** **(ii)** 0.21

www.iranino.co

IRANINOVIN

Topic 26

Now it's your turn

1 0.109906 u
2 101 MeV
3 7.3 MeV per nucleon
4 (a) $\frac{1}{4}$
(b) $\frac{1}{1024}$
(c) $\frac{3}{4}$
(d) $\frac{15}{16}$
5 (a) 5.6×10^{12} Bq
(b) 2.5×10^{6} Bq
6 (a) 3.3×10^{12}
(b) 1.3×10^{12}
7 (a) 1.6×10^{-4} s
(b) 1.7×10^{11} s
8 (a) $1.2 \times 10^{19}\,s^{-1}$
(b) $4.6 \times 10^{-2}\,h^{-1}$

Examination style questions

1 (b) 6 hours
2 (a) 0.00857 u, 7.98 MeV, 2.66 MeV per nucleon
(b) 0.687 u, 640 MeV, 6.60 MeV per nucleon
(c) 1.79 u, 1660 MeV, 7.50 MeV per nucleon
3 (a) 236.132 u, 235.918 u
(b) 0.214 u
(c) 199 MeV
(d) 8.2×10^{10} J
(e) 74 g
4 (a) 4.96 MeV
(b) 14 MeV
5 1.2×10^{-10} kg
6 (a) 2.2 h
(b) 1.3×10^{21}
7 (a) $0.046\,h^{-1}$
(b) $6000\,cm^3$
8 (c) (i) 1.77 MeV
(ii) 6.61 MeV per nucleon
9 (b) (i) 7.4×10^{-4} u
(ii) 1.1×10^{-13} J
10 (b) (i) 7.9×10^{-13} J

www.iranino.co

IRANINOVIN

Index

absolute zero 213
absorption spectra 359
accelerating potential difference, X-ray tubes 363, 365
acceleration 43–4
  centripetal 185–6
  of free fall 24–5, 44, 45–6, 197
  kinematic equations 44–5, 46–7
  Newton's second law of motion 56
acceleration–time graphs 49
accuracy 31
activity of a radioactive source 384
adiabatic change 226
aircraft, banking 187
air resistance 60
alpha (α)-particles 171–2, 174
  Rutherford's α-scattering experiments 177–8
alternating current (a.c.) 341–3
  rectification 345–7
  use for electrical energy transmission 345
ammeters 27–9
ampere (amp, A) 145, 317
amplifier gain 268–9, 302
  inverting amplifier 303
  non-inverting amplifier 304
  see also operational amplifier (op-amp)
amplitude modulation (AM) 259, 260
  comparison with FM transmissions 261–2
amplitude of an oscillation 231
amplitude of a wave 103
analogue meters 27–8
  as output devices 307
analogue signals 262
analogue-to-digital converters (ADCs) 263–5
angles, measurement of 23–4
angular frequency of an oscillation 232
angular speed 184
angular velocity 185
antimatter particles 179
  antiquarks 180
  in β-decay 172–3
antinodes
  in a closed pipe 123, 124
  on a vibrating string 121
antiphase 103
  interference 115–16
Archimedes' principle 77
A-scans, ultrasound 252
atomic mass unit (u) 376
atomic number (proton number, *Z*) 169
atomic structure 168–70, 176–7
  electron diffraction patterns 178
  fundamental particles 178–81
  Rutherford's α-scattering experiments 177–8
attenuation of signals 268–70
attenuation of X-rays 367–8
attenuation per unit length 269
avalanche effect, Geiger–Müller tube 175
average acceleration 43
average speed 40–1
average velocity 41–2
Avogadro constant 204
background radiation 176
balances 21–3, 55, 59
band theory of electrons 360
bandwidth 260
  op-amp 299
barium, use in X-ray imaging 367
Barton's pendulums 239–40
base units 3
  and homogeneous equations 5
batteries, internal resistance 158–9
becquerel (Bq) 384
bel (B) 268
beta (β)-particles, beta-decay 172–3, 174, 180–1
binary numbers (digital numbers) 263
binding energy 377–8
binding energy per nucleon 378–9
bits 263
blue shift 109
boiling 218
Boltzmann constant 204
Boyle's law 202
Brahe, Tycho 2, 194
Bremmstrahlung radiation 363
bridge rectifier circuit 346
Brownian motion 206
B-scans, ultrasound 252–3
bubble chamber 319
buoyancy force (upthrust) 71, 76–7
calibration errors 16, 35
calibration of instruments 15–16
  micrometer screw gauge 18
  thermocouple thermometer 27
capacitance 281–2
  influencing factors 282–3
capacitor–resistor circuits, time constant 288
capacitors 282
  charging and discharging 286–9
  energy storage 285–6
  series and parallel connections 283–5
  smoothing of rectifier output 346–7
  uses of 289
  variable 283
carrier waves 258–60
cars
  cornering 187
  suspension system 238
cathode-ray oscilloscope (c.r.o.)
  measurement of the frequency of sound 107
  measurement of frequency 25
  measurement of voltage 30
cellulose acetate, electric charge 136–7
Celsius scale 214
centre of gravity 75
centrifuge 188
centripetal acceleration 185–6
centripetal force 185–6
  examples 186–8
  in vertical circles 188–9
chain reaction, nuclear fission 380–1
charge *see* electric charge
charged particles
  motion in an electric field 141–3
  motion in a magnetic field 318–21
  velocity selection 321–2
Charles' law 202–3
circuits *see* electrical circuits
circular motion
  angular speed and angular velocity 184–5
  centripetal acceleration and centripetal force 185–6
  examples 186–8
  radians 184
  in vertical circles 188–9
circular orbits 193–5
closed triangle, vectors 10
coaxial cables 256
  comparison with wire-pairs 265–6
coherent wave sources 117–18
collisions 62–3
  elastic and inelastic 64–6
communication
  analogue and digital signals 262–3
  modulation 257–60
  radio-wave frequencies 260–1
  relative advantages of AM and FM transmissions 261–2
  signal attenuation 268–70
  transmission of a signal 263–5
communication channels
  coaxial cables 256

www.iran-archive.com

IRANINOVIN

microwaves 256–7
optic fibres 257
radio waves 256
relative merits 265–8
wire-pairs 255–6
communication satellites 266–8
comparator circuits 299–301
components of vectors 11
compressive deformation 94
computed tomography (CT scanning) 369–72
conduction band, electrons 360
conduction of electric current 147
insulators 362
intrinsic semiconductors 360–1
metals 362
conductors
Ohm's law 150–1
resistance 149–50
conservation of electrical charge (Kirchhoff's first law) 160
conservation of energy 81
first law of thermodynamics 226–7
kinetic energy 64
Kirchhoff's second law 160–1
Lenz's law 335
conservation of momentum 61–3
constant-volume gas thermometer 212–13
constructive interference 115
continuous spectra 356, 358–9
contrast, X-ray images 366, 367
controlled chain reactions 380–1
conventional current 147
coplanar vectors 11
cosine rule 14
coulomb (C) 145, 146
Coulomb's law 275
couples 72–3
critical damping 238
cross-talk (cross-linking), wire-pairs 265
current balance 22–3, 315
cut-off wavelength, X-rays 364, 365
damping of oscillations 238, 239–40
daughter nuclides 171
de Broglie wavelength 355
decay constant 384
relationship to half-life 385–6
decay curves 382–3, 385
decibel (dB) 268
deformation 94
Hooke's law 94–5
strain energy (elastic potential energy) 95–6
Young modulus 96–8
density 76
derived units 4–5
destructive interference 115
dielectrics 282
relative permittivity of 283
differentiation 335
diffraction 126
electron diffraction patterns 178
Huygens' explanation 127–8
Young's double-slit experiment 118–19
X-ray diffraction 355
diffraction gratings 129–30
digital meters 29
as output devices 307
digital signals 262–3
digital timers 24–5
digital-to-analogue converters (DACs) 264–5
diodes 151
displacement 41–2
of an oscillating mass 231
of a wave particle 102
and work 82, 83
displacement–distance graphs of waves 102–3
displacement–time graphs 43, 47–8
of oscillations 231
of waves 103
Doppler effect 108–9
double-slit experiment, Thomas Young 118–19
down (d) quarks 180
Earth, mass of 197
eddy current damping 336
efficiency 87–8
Einstein, Albert 353, 377
elastic (strain) potential energy 86, 95–6
elastic change 94
elastic collisions 64–5
elastic limit 94
electrical circuits 156
circuit symbols 157
Kirchhoff's first and second laws 160–1
parallel circuits 160
potential dividers and potentiometers 162–4
resistors in series and in parallel 161–2
series circuits 159
electrical energy transmission 344–5
electrical heating 150
electrical power 149
electric charge 136–7, 145
see also point charges
electric current 145–6
conventional current 147
magnetic effect 313
magnetic fields 324–5
measurement of 27–9
scalar nature 7
see also alternating current
electric fields 137–8
comparison with gravitational fields 192, 278, 325
effect on radioactive emissions 174
induced fields and charges 138–9
motion of charged particles 141–3
uniform 139
electric field strength 139–40
due to a point charge 276
and electric potential 277
electricity meters 90
electric potential 277
electric potential energy 276–7
storage in a capacitor 285–6
electrolytic capacitors 282
electromagnetic force 314–15
electromagnetic induction 332–5
eddy current damping 336
e.m.f. induced between two coils 337
electromagnetic spectrum 109–10
electromagnetic waves 109
see also gamma rays; light; microwaves; radio waves; X-rays
electromagnets 325
electromotive force (e.m.f.) 157
see also electromagnetic induction
electron charge 141, 145, 146
electron diffraction patterns 178
electron energy bands 360
electronic sensors 292
processing circuits *see* operational amplifier
use of potential dividers 295–6
see also output devices; sensing devices
electron microscope 320
electrons 168–9
determination of specific charge 319–21
energy levels 357–8
wave-particle duality 355
electron transition 357–8
electron-volts (eV) 173
emission spectra 356
relationship to absorption spectra 359
empirical scales of temperature 212
energy 80
conservation of 81–2, 226–7
electrical 148–9
forms of 81
kinetic 85–6
mass–energy equivalence 377
potential 86–7
storage in a capacitor 285–6
strain energy (elastic potential energy) 95–6
thermal 219–24
typical energy values 81
energy changes 81–2, 226–7
efficiency 87–8
in simple harmonic motion 235–7
equations, homogeneous (balanced) 5
equilibrium, thermal 212
equilibrium of forces 73–5
errors 32
balances 21
digital meters 29
metre rule 16–17

www.iranino.co

IRANINOVIN

micrometer screw gauge 18, 19
systematic and random 34–6
see also uncertainties
evaporation 218
excited states, atoms 357
exponential changes
capacitor discharge 287
radioactive decay curves 383, 385
X-ray attenuation 367–8
extreme value method of uncertainty estimation 38
farad (F) 281–2
Faraday, Michael, law of electromagnetic induction 334
feedback 302
fields of force 71
comparison of effects 325–6
see also electric fields; gravitational fields; magnetic fields
film badge dosimeter 175
fine-beam tube 320
fixed points, temperature scales 212
Fleming's left-hand rule 314–15
Fleming's right-hand rule 334
forbidden band, electrons 360
force constant (spring constant) 95
forced vibrations 238
forces 55
centripetal 185–6
couples 72–3
electromagnetic 314–15
electromotive 157
equilibrium 73–5
impulse 63
moments 71–2
and motion 56
Newton's laws of motion 56–7, 58, 60–1
types of 71
weight 58–60
and work 82–3
forward bias, diodes 151
fossil fuels 81
fractional uncertainty 37–8
free (mobile) electrons 145
free-body diagrams 60–1
free fall acceleration 44, 45–6, 197
determination of 24–5
free oscillations 237
frequency
measurement using a cathode-ray oscilloscope 25, 107
of an oscillation 230
of a wave 103
frequency modulation (FM) 259, 260
comparison with AM transmissions 261–2
friction 56
production of electric charge 136–7
full-wave rectification 345
fundamental mode of vibration 121
fundamental particles 178–9
hadrons, quark model of 180
leptons 180–1
fusion, latent heat of 218, 222–3
gain of an amplifier 268–9, 299, 302
inverting amplifier 303
non-inverting amplifier 304
Galileo Galilei 46, 50, 56
galvanometer 28, 333
gamma (γ) rays 110, 173–4
gases
kinetic model 218
kinetic theory 207–9
microscopic model of 205–7
structure of 218
see also ideal gases
gas expansion, work done 83–5
gas laws 202–4
gas pressure measurement 20
Gay-Lussac's law (law of pressures) 203
Geiger counter 175
geostationary orbit 195, 267
graphs
decay curves 382–3, 385
of kinematic equations 47–9
of motion 42–3
labelling conventions 6
of oscillations 231
resonance curves 239
of waves 102–4
gravitational constant 192
gravitational fields 191
acceleration of free fall 24–5, 44, 45–6, 197
circular orbits 193–5
comparison with electric fields 192, 278, 325
Newton's law of gravitation 191–2
of a point mass 196–7
gravitational field strength 192–3
gravitational potential 198
gravitational potential energy 86–7, 198–9
ground state of an atom 357
hadrons 179
quark model 180
see also neutrons; protons
half-lives of radioactive isotopes 382–3
relationship to decay constant 385–6
half-value thickness (HVT) 367
half-wave rectification 345
Hall effect, Hall voltage 322–3
Hall probe 31
hardness of X-rays 364, 365
harmonic oscillators 231
harmonics
in a closed pipe 123–4
of a stretched string 121
heat (thermal) capacity 221–2
heat energy exchanges 224
Helmholtz coils 325
holes, intrinsic semiconductors 361
homogeneous (balanced) equations 5
Hooke's law 94–5
Huygens, Christian, explanation of diffraction 127–8
hydrogen, line spectrum 356
ice, determination of latent heat of fusion 223
ideal gases
Boyle's law 202
Charles' law 202–3
combined gas law 203–4
definition 205
Gay-Lussac's law (law of pressures) 203
see also gases
impulse 63
induced nuclear fission 380
induction of electric charge 137, 138–9
inelastic collisions 65
inertia 56
infra-red radiation 110
input impedance, op-amp 299
instantaneous acceleration 43
instantaneous velocity 42–3
insulators 362
intensity of a wave 104
intensity of X-rays 364, 365
intensity reflection coefficient 249–50
interference 114–16
production of stationary waves 120, 122–3
interference patterns 115, 116
production with light waves 117–19
production with sound waves 117
Young's double-slit experiment 118–19
internal energy 225
first law of thermodynamics 226–7
internal resistance 158
effect on power from a battery 159
intrinsic semiconductors 360–1
inverse square law of force (Coulomb's law) 275
inverting amplifier 302–4
ion microscope 176
ions 169
positive 146
velocity selection 322
isothermal change 227
isotopes 170
radioactive 382–3
Joule heating 150
kelvin (K) 214
Kepler, Johannes, third law of planetary motion 194
kilowatt-hours 90
kinematic equations 44–5
examples of use 46–7
graphical representation 47–9

www.iranino.com

IRANINOVIN

kinetic energy 85–6
conservation of 64
of photoelectrons 352
in simple harmonic motion 235–6
of subatomic particles 173
kinetic model of matter 217–18
kinetic theory 206, 207–9
Kirchhoff, Gustav
first law 160
second law 160–1
Larmor frequency 327
latent heat of fusion 218, 222–3
latent heat of vaporisation 218, 222–4
lead-214, β-decay 172
least significant bit (LSB) 263
length, measurement of 16–20
Lenz's law 335
leptons 179, 180–1
see also electrons; neutrinos
lever balance 21–2
light
continuous emission spectrum 356
Doppler effect 109
electromagnetic spectrum 110
interference patterns 117–19
nature of 353–5
light-dependent resistor (LDR) 293, 362
light-emitting diodes (LEDs) 306
linear attenuation (absorption) coefficient
for ultrasound 250–1
for X-rays 368
linear momentum 57–8
line spectra 356, 358
liquid-in-glass thermometer 26, 27, 215
liquids
kinetic model 217–18
structure of 218
lodestone 311
logarithms to the base 10 269
longitudinal waves 102
lost volts 158
magnetic compass 311
magnetic fields 312–13, 325–6
due to currents 324–5
effect on radioactive emissions 174
force between parallel conductors 317
force on a moving charged particle 318–21
Hall effect 322–3
motor effect 314–16
production by an electric current 313
velocity selection of charged particles 321–2
magnetic field strength 315
magnetic flux 332
magnetic flux density 316
measurement of 31, 317–18
magnetic flux linkage 333
magnetic resonance imaging (MRI) 326–8
magnets, law of 312
manometer 20
mass 198
measurement of 21–3
Newton's second law of motion 56
mass defect 377
mass–energy equivalence 377
mass number (nucleon number, *A*) 169
measurements 1, 15–16
accuracy and precision 31–2
of angles 23–4
choice of instrument 33–4
of current and potential difference 27–30
of force 55
of gas pressure 20
of length 16–20
of magnetic flux density 31
of mass 21–3
SI units 2–4
of temperature 26–7
of time 24–5
uncertainty 22, 26, 29, 32–3
meniscus 20
mercury, line spectrum 358
mercury-in-glass thermometer 26, 27, 215
metals
conduction 362
free electrons 145
metal wire strain gauge 295
metre rule 16–17, 19
micrometer screw gauge 17–18, 19
microphones 294
microwaves 110
stationary waves 124
use for communication 256–7, 266
Millikan, Robert A., measurement of electron charge 141, 145
modes of vibration
in a closed pipe 123–4
of a stretched string 121
modulation 257–60
molar gas constant (universal gas constant) 204
mole 204
moment of a force 71–2
moments, principle of 73–4
momentum 57–8
and impulse 63
principle of conservation 61–3
most significant bit (MSB) 263
motion
and force 56
graphical representation 47–9
kinematic equations 44–5, 46–7
kinetic energy 85–6
Newton's laws 56–7, 58, 60–1
non-uniform 60
two-dimensional (projectile) 49–51
see also acceleration; circular motion; displacement; velocity
motor effect 314–16
multimeters 29
multipliers, use with a galvanometer 28
natural frequency 238
negative feedback 302
neutral point, magnetic fields 313
neutrinos and antineutrinos 172
neutrons 168–9
Newton, Isaac
law of gravitation 191–2, 194
laws of motion 56–7, 58, 60–1
nodes
in a closed pipe 123, 124
on a vibrating string 121
noise
analogue and digital signals 262–3
signal-to-noise ratio 268
wire-pairs 265
non-inverting amplifier 304
normal contact forces 59
nuclear equations 378
nuclear fission 379–81
nuclear fusion 379–80
nuclear magnetic resonance 326–7
magnetic resonance imaging 327–8
nuclear power stations 380
nucleon number (mass number, *A*) 169
nucleons 169
nucleus of an atom 168, 169
binding energy 377–8
isotopes 170
radioactive emissions 170–4
stability of 378–9
nuclides, conventional notation 169–70
Oersted, Hans Christian 313
ohm (Ω) 149
Ohm's law 150–1
open-loop gain 299
operational amplifier (op-amp) 298–9
feedback 302
ideal properties 299
inverting 302–4
non-inverting 304
use as a comparator 299–301
optic fibres 257
advantages of 266
orbits 186–7, 193–5
orders of diffraction 128, 130
orders of magnitude 6–7
oscillations 230–1
angular frequency 232
displacement–time graphs 231
free and damped 237–8
resonance 238–40
see also simple harmonic motion
output devices 305
digital and analogue meters 307
light-emitting diodes (LEDs) 306
relays 305–6
output impedance, op-amp 299

www.iranipo.co

IRANINOVIN

overdamping 238
overtones
in a closed pipe 123–4
on a vibrating string 121
parabolic paths 50
parallax error 16–17, 21, 35
parallel circuits 160
capacitors 284–5
resistors 161–2
parallel conductors, force between 317
parent nuclides 171
path difference, interference 116
pendulums 234–5
Barton's 239–40
period of an oscillation 230
period of a wave 103
period of planetary orbits 194–5
permittivity of free space (permittivity of a vacuum) 275
phase difference, definition 103
phase of waves 103–4
and interference 115–16
phosphorus-30, β-decay 172
photoelectric emission 350–1
demonstration of 351–3
photon theory 353–4
relationship to frequency of radiation 354
photoelectric equation 354
photoelectrons 351
energies of 351–2
photomultiplier 175
photon emission and absorption, electron transition 357–8
photons 353
physical quantities 2
conventions 6
piezo-electric transducer 247–9, 294
pixels 369–72
Planck, Max 353
Planck constant 353
planetary orbits 193–5
plastic deformation 94
point charges 274
electric field strength 276
force between (Coulomb's law) 275
polar satellites 267–8
polythene, electric charge 136–7
position–time graphs 42–3
positive ions 145, 146
positrons (β-particles) 172–3
potential difference (p.d.) 148–9, 157
measurement of 27–30
terminal potential difference 158
potential dividers 162–4, 295–6
potential energy 86–7
electric 276–7, 285–6
gravitational 198–9
variation in simple harmonic motion 236
potentiometers 163
power 89–90
electrical 149
power loss
transformers 344
transmission lines 345
precession 326
precision 31–2
pressure 76
pressure of a gas 20, 83–5
kinetic theory 207–8
relationship to temperature (Gay-Lussac's law) 203
relationship to volume (Boyle's law) 202
probing matter 176–7
electron diffraction patterns 178
Rutherford's α-scattering experiments 177–8
progressive waves 101
differences from standing waves 122
projectile motion 49–51
proportionality limit 94
proton number (atomic number, *Z*) 169
protons 168–9
protractor 23–4
Pythagoras' theorem 14
quanta 353
quantum theory 353
quarks 180–1
quartz crystal, piezo-electric effect 247–9
radians 184
radioactive decay 170
background radiation 176
β-particles and β-decay 172 3
detection 175–6
γ-radiation 173–4
mathematical descriptions 383–6
nuclear fission 379–81
α-particles and α-decay 171–2
spontaneous and random nature 176, 381–3
summary of radioactive emissions 174
radioactive decay equation 385
radioactive decay series 174–5
radio waves 110, 256, 266
comparison of AM and FM transmissions 261–2
frequencies 260–1
random decay 381, 382–3
random processes 381–2
random uncertainties (errors) 35–6
parallax error 16–17
reaction times 35
rectification of alternating current 345–7
rectifiers 151
red shift 109
reflection
of ultrasound 249–51
of water waves 106
refraction 106–7
relative atomic mass (Ar) 204
relative molecular mass (Mr) 204
relative permittivity 283
relaxation time, nuclear magnetic resonance 327
relays 305–6
resistance 149–50
internal 158–9
relationship to temperature 151, 361
resistivity 152–3
resistors 149
connection in parallel 161–2
connection in series 161
light-dependent 362
potential dividers and potentiometers 162–4
resolution of vectors 11
resonance 238–40
closed pipes 124
vibrating strings 120
resonance curves 239
resonance tubes, measuring the speed of sound 124–5
resonant frequency 238
restoring force 232
resultant force 55
resultant of vectors 8–11
reverse bias, diodes 151
right-hand grip rule 324–5
ripple tank demonstrations 105–7
root-mean-square (r.m.s.) speed of molecules 208–9
root-mean-square values, alternating current 342
rotational equilibrium 74
rotational kinetic energy 86
Rutherford, Ernest, α-scattering experiments 177–8
sampling frequency (sampling rate) 263–5
satellite communication 266–8
satellite orbits 186–7
geostationary 195
saturation of an amplifier 299, 300
scalars 7–8
scientific method 1
scientific notation 2
scintillation counter 175–6
semiconductors, intrinsic 360–1
sensing devices 292–3
light-dependent resistor (LDR) 293
metal wire strain gauge 295
piezo-electric transducer 294
thermistor 293–4
series circuits 159
capacitors 283–4, 285
resistors 161
sharpness, X-ray images 366–7
shunts, use with a galvanometer 28

www.iranino.co

IRANINOVIN

SI (Système International) units 2–4
derived units 4–5
sidebands (sideband frequencies) 260
signal attenuation 268–70
signal-to-noise ratio 268
simple harmonic motion (s.h.m.) 231–2
kinetic energy changes 235–6
of a mass on a helical spring 234
potential energy changes 236
simple pendulum 234–5
solution of equation for 232–3
total energy 237
sine rule 14
slew rate, op-amp 299
smoothing of rectifier output 346–7
solenoids, magnetic field 324–5
solids
kinetic model 217
structure of 218–19
sound
determination of the frequency 107
Doppler effect 108–9
speed of, measurement using stationary waves 124–6
sound waves, interference patterns 117
specific acoustic impedance (*Z*) 249, 250
specific charge 319
of an electron 319–21
specific heat capacity 219
determination of 219–21
specific latent heat 222–3
spectra 356
absorption 359
continuous 358–9
of X-rays 364
spectrometers 130, 358
spectroscopy 358
speed 40–1
angular 184
of a wave 104
speed of sound, measurement using stationary waves 124–6
spherical waves 104
'spin' of atomic nuclei 326–7
spontaneous processes, radioactive decay 381
spring balance (newton balance) 21, 55, 59
spring constant (force constant) 95
springs
energy storage 80, 95–6
Hooke's law 94–5
longitudinal waves 102
simple harmonic motion 234
stationary (standing) waves 120
static electricity 136
stationary (standing) waves 119–20
in air 123–4
demonstration using microwaves 124
differences from progressive waves 122
explanation by interference 122–3
measuring the speed of sound 124–6
on strings 120–2
step-down transformers 344
step-up transformers 343
stopclocks and stopwatches 24–5
stopping potential, photoelectrons 352
strain 96
strain energy (elastic potential energy) 95–6
strain gauges 295
strange (s) quarks 180
stress 96
strings, vibrating 120–2
strong nuclear force 177, 179, 380
superposition of waves 116
see also interference
symbols, conventions 6
systematic uncertainties (errors) 35, 36
zero error 16
tables, labelling conventions 6
Tacoma Narrows bridge disaster 240
temperature 211–12
effect on resistance 151, 361
measurement of 26–7
thermodynamic 214
temperature of an ideal gas
kinetic theory 208–9
relationship to pressure (Gay-Lussac's law) 203
relationship to volume (Charles' law) 202–3
temperature scales 212–14
tensile deformation 94
tensile stress 96
terminal potential difference 158
terminal velocity 60
tesla (T) 316
theodolite 2
theories 1
thermal (heat) capacity 221–2
thermal energy
exchanges of heat energy 224
specific heat capacity 219–21
specific latent heat 222–3
thermal equilibrium 212
thermionic effect 363
thermistors 163, 215–16, 293–4
thermocouple thermometer 27, 216
thermodynamics, first law of 226–7
thermodynamic temperature 214
thermoelectric effect 216
thermometers 26–7, 212, 215–16
thermometric substances, thermometric properties 212
threshold frequency, photoelectric emission 351, 354
time, measurement of 24–5
time constant, capacitor–resistor circuits 288
top-pan balance 21
torque 73
transducers 294
piezo-electric 247–9
transformers 337, 343–4
translational kinetic energy 86
transmutation 171
transverse waves 101–2
graphical representation 102–4
trigonometry 14
and two-dimensional motion 51
two-dimensional (projectile) motion 49–51
ultrasound
diagnostic use 251–3
generation of 247–9
reflection and absorption 249–51
ultra-violet radiation 110
uncertainties 22, 26, 29, 32–3
combination of 36–8
systematic and random 34–6
units
conventions 6
derived 4–5
SI (Système International) 2–4
universal gas constant (molar gas constant) 204
up (u) quarks 180
upthrust (buoyancy force) 71, 76–7
uranium-234, α-decay 171–2
uranium-235, nuclear reactions 380
uranium-238, decay series 175
valence band, electrons 360
vaporisation, latent heat of 218, 222–4
variable capacitors 283
vector polygons 11
vector representation 8
vectors 7–8
addition of 8–11
resolution of 11
vector triangles 9–10
laboratory demonstration of 10
velocity 41–2
angular 185
kinematic equations 44–5, 46–7
position–time graphs 42–3
velocity selectors 321–2
velocity–time graphs 43, 48, 49
vernier caliper 18–19
vertical circular motion 188–9
vibrating strings 120–2
virtual earth 303
viscosity 71
viscous force 71
volt (V) 148
voltage (potential difference) 148–9, 157
measurement of 27–30
terminal potential difference 158
voltmeters 27–9

www.iranino.co

IRANINOVIN

volume of an ideal gas
relationship to pressure (Boyle's law) 202
relationship to temperature (Charles' law) 202–3
voxels 369–72
water
determination of latent heat of vaporisation 223–4
triple point of 214
wave equation 104
wavefronts 105
wavelength 103
wave motion 101–2
properties 105–7
wave–particle duality 355
waves
carrier waves 258
diffraction 126–30
Doppler effect 108–9
electromagnetic 109
graphical representation 102–4
interference 114–19
microwaves 256–7
principle of superposition 116
progressive 102
radio waves 256, 260–1
sound, determination of frequency 107
spherical 104
stationary (standing) 119–24
transverse 101–2
see also ultrasound
weak force (weak interaction) 180
weber (Wb) 332
weight 58–60, 71, 198
centre of gravity 75
weightlessness 196
white light, diffraction 130
white-light fringes 119
wire-pairs, use in communication 255–6, 265
work 80, 82–3
expanding gases 83–5
work function energy, photoelectric emission 353–4
X-ray diffraction 355
X-ray images 366–7
X-rays 110
attenuation of 367–8
computed tomography 369–72
control of X-ray beams 364–5
production of 363–4
Young, Thomas, double-slit experiment 118–19
Young modulus 96–8
zero error 35
balances 21
digital meters 29
metre rule 16
micrometer screw gauge 18, 19

www.iranino.co